

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1	9	0	0	0	0	In all, the authors of Chapter 9 should be commended on a very substantial effort and potentially a major achievement. However, it is longer than ideal. [Government of Australia]	Taken into account, thank you. Length has been reduced during editing.
9-2	9	0	0			General remark chapter 9: Due to lack of time I have only scanned/commented the executive summary (page 9-3 till 9-5), thereby focusing more on on clarity and logic than on content.I haven't reviewed the remaining part of this chapter [Hans Visser, The Netherlands]	noted
9-3	9	0	1			Terms such as "modest agreement", "reasonable agreement", "fairly well" etc, are used throughout Chapter 9 when comparing models with observations, and the performance of models. It is not clear how these terms rate sequentially - for example, is 'modest agreement' a higher level of agreement than 'reasonable agreement'? Elsewhere 'skill' is used to describe model performance. It would be good if more consistent, and quantitative terminology could be used where ever possible. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Terminology has been streamlined and made more quantitative where possible with specific attention to the guidance notes.
9-4	9	0	2			Consistency in assessment numbers: Because chapter assessments continue to be refined, please check carefully all values (and the uncertainty ranges) carefully between tables, figures, main text, and summary text within your chapter. If numbers are taken from other chapters, please also ensure the latest results are used. Specific examples will be highlighted in our chapter comments. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Have checked this carefully during revisions and editing. Each LA has cross-checked their section with other chapters to the best extent possible.
9-5	9	0	3			Treatment of Uncertainty: please follow the IPCC guidance note carefully; use italics to highlight formal uncertainty assessments; use likelihood in conjunction with high/very high confidence only (except in exceptional cases); if likelihood is given for situations where confidence is less than 'high', we recommend to put confidence in brackets at the end of the sentence rather than combining both confidence and likelihood in text. Please note - usage of the formal terms from the uncertainty guidance note, (egg. "likely", "confidence" etc) should be restricted to the use within statements which report assessment findings. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. We have checked this carefully during revisions and editing.
9-6	9	0	4			Format of Executive Summary (ES): As agreed at the third lead author meeting, we would ask that all chapters follow a consistent style for the ES. 1) The first sentence (or two) of each paragraph should be bolded to highlight the key message, with the subsequent sentences providing the detailed quantitative assessment. 2) Statements should incorporate the IPCC Uncertainty Language 3) Each paragraph must include a traceability to the underlying sections/subsections where the key message was drawn from (to the second level section heading), indicated using square brackets at the end of each paragraph. 3) Paragraphs should be grouped together under subtitles. The use of bullets should be avoided. 4) Finally, because the ES should be short and concise, lengthy textbook or chapeau type introductory text should be avoided. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. The ES was revised accordingly and submitted to TSU adhering to these guidelines.
9-7	9	0	5			Cross-chapter references AR5: suggest to update cross-chapter references to not just refer to Chapter number but to refer to specific section if appropriate. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Cross references have been more specific where suitable throughout the chapter.
9-8	9	0	6			References to AR4 and earlier IPCC assessments: be as specific as possible. Writing just AR4 without any reference is not useful to the reader. Please refer to specific chapter where possible. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Have done so where a specific reference is intended. In some cases the reference is more 'generic'.
9-9	9	0	7			Use of acronyms: In order to improve overall readability of the report, we would like to suggest that you please avoid acronyms that are not needed and/or are not used in more than one section of your chapter. [Thomas Stocker/ WGI TSU, Switzerland]	noted.
9-10	9	0	8			Personal pronouns: our strong preference is to minimize the usage of personal pronouns, e.g., we/us/our to the extent possible. Exceptions to this would be when the Chapter's assessments conclusions are presented as clear summary statements. [Thomas Stocker/ WGI TSU, Switzerland]	noted.
9-11	9	0	9			Please make sure to provide updates of relevant data from your chapter that will be collected in Annex II - Climate System Scenario Tables, to the Annex II Chair. Also, please take the time to critically check all the entries in Annex II that are based on your Chapter assessment or that you are using in your chapter assessment. [Thomas Stocker/ WGI TSU, Switzerland]	Noted. Ch09 has limited input to Annex II
9-12	9	0				Some consistency needs to be applied across Ch 2, 9,10,11,12,14 to the index names used for the extremes indices. For instance, annual maximum 5-day rainfall is referred to as R5dmax in Ch 12, RX5day in Ch 9, and	Taken into account. Have striven for consistency as chapter was revised and edited.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						R5d in Ch 14, and the warmest 10% of nights as TN90 in Ch10 and TN90p in Ch 2. This should be coordinated amongst all relevant chapters. [Lisa Alexander, Australia]	
9-13	9	0				Figures: please show the model spread (e.g., the standard deviation) for Figs. 9.2-9.4 [Massimo Bollasina, Italy]	Rejected. Space constraints prevent us from expanding these figures; model spread is amply documented in other figures.
9-14	9	0				The results presented in this chapter, and I presume in other chapters as well, partially suffer from the limited time available to carry out the analysis of the CMIP5 runs. This is unfortunate and might lead to misinterpretations. [Massimo Bollasina, Italy]	Noted. Where feasible, the chapter has been updated with information from new publications and new input to the CMIP5 archive.
9-15	9	0				Chapter 9 : The chapter includes many nice things. However its scope is very broad, and probably too broad. In particular, I think that the section on Climate Sensitivity is not well articulated with the other chapters. The main drivers of climate sensitivity, namely the water vapor, lapse rate and cloud feedbacks, but also forcings and fast adjustments to some extent, are assessed and evaluated in Chapter 7 and 8. Chapter 9 should not repeat these evaluations (aspects of the Chapter 9 discussion which are not already discussed in Chapters 7 and 8 could move to these chapters). Chapter 9 could then focus on : the evaluation of processes and feedbacks not covered by the other Chapters (e.g. the surface albedo feedback), the observational constraints on climate sensitivity (something which is also addressed in Chapter 12 : paleo, volcanic eruptions, historical period, etc.), and the relationship between model performance and climate sensitivity. [Sandrine BONY, France]	Accepted -- Duplication of material relative to chapters 7 and 8 has been removed and complementary aspects have been highlighted.
9-16	9	0				I feel that the overall discussion and organization of this chapter are very inadequate. I pose to the authors the following question: Why do we evaluate climate models? Two reasonable answers might be: (1) To learn where their flaws are so that we can be guided in our development of new parameterizations, resolution choices, etc.; (2) To decide which model(s) are most reliable for predicting the climate changes that are the reason such models have been created in the first place. Both are worthwhile goals - but only the first has been achieved to any degree. The community has been obsessed since AR4 with the idea of metrics to judge climate models and using them to weight climate model predictions of the future. But what the community has actually done since AR4, for the most part, is akin to the guy who loses his keys and looks for them under the lamppost because that's where the light is best, rather than looking for them in the places he is most likely to have lost them. Our metrics are based on what is available and easy to calculate rather than what is needed to isolate the processes that control climate change. We became enamored of validating model monthly mean geographic distributions of parameters using Taylor diagrams and various other "skill" metrics. But it has become obvious over the past few years (in papers that this chapter cites!) that the vast majority of these exercises have not actually been useful in differentiating the climate change predictions of "good" vs. "bad" models. Presumably this is telling us that (1) monthly mean geographic fields are usually not good indicators of climate change, (2) there are many different things we and the public want to know about climate change (climate sensitivity, subtropical drought, Arctic sea ice decline, etc.) and perhaps different aspects of the climate and the processes we use to represent how it changes in response to a change in environmental conditions are relevant to each thing, and there is no such thing yet as a "good" or "bad" model, (3) we have not yet done the hard work needed to identify which aspect of climate change is best predicted by which aspect of current climate behavior. Yet we remain addicted to the simple metrics even in the face of the evidence that suggests that they do not tell us about the things we wish to predict. Consequently, a short summary of the chapter as written might be "Since AR4 we have come up with many different ways to evaluate climate models against observations. These comparisons indicate that certain models are much better at simulating certain features of the current climate than others, but that we do not yet have an effective way to judge model predictions of climate change itself. However, we are going to make you read through 3/4 of the chapter before we show you that we do not actually know how to judge the accuracy of most climate model predictions of the future based on their ability to simulate the present climate." I strongly suggest that you eliminate charged words such as "skill" which do not actually apply to the predictions of climate change and re-orient the discussion toward (1) a philosophy of using observations of the current climate primarily to provide insights into future fruitful directions for model improvement, and (2) the need for the community to develop more useful process-based metrics than the simplistic ones currently used if we wish to actually differentiate one model's prediction from another's, an idea that I think receives one sentence of discussion in	Rejected. As an 'assessment' of published science, we can only provide an assessment of what is available. It is the state of the science that a variety of metrics have been used to evaluate climate models, but there has been limited progress on linking these performance metrics to 'reliability' of future climate projections. The approach has therefore necessarily been to provide as comprehensive a synthesis of model evaluation as possible, give page limits, realizing that this provides only a measure of the ability of models to reproduce the past and current climate, and the processes that shape it, and that this provides a physical basis for making use of such models in future projections. Note that we do have two subsections devoted entirely to "lessons learned from evaluation", in the sense the reviewer implies.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						the chapter. [Anthony Del Genio, United States of America]	
9-17	9	0				In ROFOD, I suggested to rerun the models with a reevaluated anthropogenic CO2 contribution to radiative forcing of ~ 0.0025 W/m2 per year for the reasons detailed in ROFOD and confirmed throughout this reviewer report. It has been ignored in SOD. Due to the enormous impact of such a difference in radiative forcing numbers on the next chapters and on the conclusions of AR5, it does not appear useful and even worst continuing to comment the subsequent chapters devoted to modeling because the inputs in terms of anthropogenic forcing are strongly exaggerated. It is not worth repeating the same recommendations several times. [François Gervais, France]	Rejected. The IPCC Assessment is based on peer-reviewed published science. The IPCC does not 'run' or 're-run' models -- this is up to the modelling centres and recommendations for specific climate model experiments should be directed to such centres.
9-18	9	0				This chapter is huge. It should give conclusions for improvements of the models. The representation of the hydrological cycle in climate models is short. For precipitation, the mean of the models ensemble is given. The meaning of averaging across models is questionable. To show the variance among models would be a way to recognize that considerable work is still needed in this area. The variation across models has indeed been done for sea level (fig 9.16) and ocean heat transport (fig 9.16). These figures show that the difference between models are big. The adequate resolution of western boundary currents is a big problem for the ocean. Unless a convergence against resolution is shown, the AOGCMs remain informative but limited prediction tools This chapter should be reduced, better synthesized, and the conclusions on model performance should be more modest than they are in the present document. The conclusions of the chapter should list the fundamental research issues that need to be tackled to improve the climate models (clouds, oceanic turbulence to quote but two). [Government of France]	Rejected. Chapter length has been addressed in final revisions and edits. Variance across the multi-model ensembles is illustrated in many figures in which results from all individual models are shown. The IPCC Assessment is aimed at synthesizing published research. It is not a forum for suggesting new research programs.
9-19	9	0				Please print uncertainty qualifiers always in italic letter format. [Government of Germany]	Taken into account; calibrated language statements has been indicated by italics
9-20	9	0				It would be good to mention the IAM some where in this chapter and to explain why they are not evaluated here. [Government of Germany]	Rejected. Integrated Assessment Models are not evaluated in this chapter as their use later in the report is limited to emulating more comprehensive models which are evaluated.
9-21	9	0				Seems repetitive. A lot of references in multiple sentences hide the clarity of the message/findings. Hard to understand what is new. How about outlining the major processes and make a table of what was AR4's weaknesses that have been improved by the new formulation and why? This table may be long, but it will capture the performance metric and establish the benchmark (if AR4 it is). [Government of India]	Taken into account. Have attempted to improve clarity through final edits. Model improvements since AR4 are discussed throughout and highlighted in the FAQ. In addition comparisons for a range of climate variables are shown in the summary figure which is being substantially revised.
9-22	9	0				There is no mention of Bay of Bengal. A major question remains: What will happen to the strength/weakness of the seasonally reversing winds in AR4 or in CMIP3 or in CMIP5? [Government of India]	Noted. Specific regional aspects are covered in Chapter 14. Summarizing model performance for all interesting regions is prohibited by space considerations.
9-23	9	0				There is ONE time mention of the Arabian Sea – Also, is there any impact of climate change on the OMZ? [Government of India]	Noted. Specific regional aspects are covered in Chapter 14. Summarizing model performance for all interesting regions is prohibited by space considerations.
9-24	9	0				No references of the works of Gadgil on Indian Monsoon. No references of the works of Sengupta, Vinaychandran or Shankar. What is the impact of temperature change on the bay of Bengal SST and/or salinity? Mitra et al. 2009 (Current Science) showed an incredible amount of change in the BoB temp and salt over last thirty years. [Government of India]	Noted. Specific regional aspects are covered in Chapter 14. Summarizing model performance for all interesting regions is prohibited by space considerations.
9-25	9	0				Please add new relevant referente that has been accepted by the international journal Climate Dynamics : Article title: "PHYSICALLY BASED EVALUATION OF CLIMATE MODELS OVER THE IBERIAN PENINSULA "; DOI: 10.1007/s00382-012-1619-2 ; Authors : M.C. Sánchez de Cos Escuin, J. Mª Sánchez-Laulhé, C. Jiménez Alonso, J.M. Sancho Ávila, E. Rodríguez-Camino [Government of Spain]	Rejected. Specific regional aspects are covered in Chapter 14. Summarizing model performance for all interesting regions is prohibited by space considerations.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-26	9	0				Many of the references are still in a phase of revision, please be sure that, in case of rejection for some of them, final conclusions are not affected. Many references cited in the text correspond to scientific papers submitted but not accepted at the time of this revision. All of them should be checked before the final draft will be released. [Government of Spain]	Noted. This has been carefully verified.
9-27	9	0				The lengths of subsections are very variable. Please, try to homogenize them. [Government of Spain]	Rejected. Varying length is inevitable. Some attempt to reduce this has been done in editing final version.
9-28	9	0				The number of CMIP3 or CMIP5 models used is only indicated in figures and parts of text. Generally, it is not explicitly pointed out. [Government of Spain]	Noted. Comment not clear. Models included in the chapter are listed in Table 9.1.
9-29	9	0				With respect to the Figures, it would be an advantage for the reader that CMIP3 and CMIP5 appear with the same colour. [Government of Spain]	Taken into account. During final revisions, figures have been re-drafted to use consistent colors.
9-30	9	0				Within Chapter 9, a number of acronyms used in this chapter are not explained when they first appear in the text. It is suggested to revise them. [Government of Spain]	Taken into account. Have checked this carefully during revisions and editing. In some cases, definitions are in Glossary.
9-31	9	0				The Likelihood Table (Table 1.1) and Confidence figure (1.12) should be repeated in the SPM, TS and each Chapter and the terminology should be applied consistently. As an alternative to repeating the complete table/figure the material should be restated briefly in the SPM, TS, and each chapter. [Government of United States of America]	Rejected. Table is in Chapter 1 for reference by readers of all chapters.
9-32	9	0				The comparison of climate models to observations in general lacks any method to define "how accurate is good enough". This is true of mean state comparisons, seasonal comparisons, ENSO comparisons, etc. A large part of the problem is that these comparisons are only crudely related to the desired decadal prediction accuracy. There is not a good physical reason to believe that capturing the mean state or seasonal cycle will be closely tied to the accuracy of decadal change prediction. Instead it can be considered circumstantial evidence: related but only in a very crude sense. A good summary of this issue can be found in the recent paper by Klocke, Pincus, and Quaas, J. Climate, 2011. Something on this fundamental issue needs to be mentioned as a current limitation. The most direct tests of decadal prediction accuracy will remain tests of model prediction of past decades of climate change. The challenge with these direct tests remains a) short climate records for many climate variables (natural variability confuses anthropogenic signals over a few decades), and b) poor observational accuracy from sampling and/or accuracy of current and past observations. In essence we are currently trying to gain confidence in climate model projections through a preponderance of circumstantial evidence: related but not direct tests of climate prediction accuracy. A similar case would be to test weather prediction accuracy by how well weather models produce mean climate statistics: its only very crudely related to the desired prediction. [Government of United States of America]	Noted. Decadal prediction skill is covered in Chapter 11.
9-33	9	0				Inconsistencies or errors in citation format throughout the chapter. For example, Page 9-8, lines 46, 53, 55; Page 9-10, lines 13, 14, 17, 25, 26 32. [Government of United States of America]	Editorial.
9-34	9	0				There is no discussion of model's ability to capture low level jets, which are critical features contributing to regional moisture fluxes and precipitation. If the authors would like to include such a discussion, they may wish to consult the paper by Hu and Fend ("Low-level Jets and Precipitation Variations in the U.S. Great Plains Simulated and Predicted in the CMIP5 Models"), which details an evaluation of current global models and indicates some skill at producing the Great Plains low-level Jet. [Government of United States of America]	Noted. Space limitations prevent a comprehensive review, but this has been considered.
9-35	9	0				The paper "North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology" by Sheffield et al. informs a number of sections in this chapter. [Government of United States of America]	Noted. Space limitations prevent a comprehensive review, but this has been considered.
9-36	9	0				Could the authors of this chapter cross check with the authors of Chapter 5 regarding the intermodel comparison of paleoclimate models. [Government of United States of America]	noted. There is close coordination with the model evaluation aspect of Ch05.
9-37	9	0				Table numbering seems off between main text and Tables 9.4-9.6. Can the authors recheck these? [Government of United States of America]	Taken into account. Corrected.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-38	9	0				in several places the chapter discusses climate sensitivity, which is fair because this is a major uncertainty in future projections. But by definition, this is the sensitivity to a given CO2 concentration and so misses any consideration of processes which affect the concentration (such as the carbon cycle feedback from climate). Hence it is true to say that clouds are the main uncertainty in ECS, but this is only half the full story. If you want to quantify uncertainty in the response to EMISSIONS, then carbon cycle processes are important too. Gregory et al (2009, J. Clim) actually show that carbon cycle feedbacks have the same magnitude and size of uncertainty as clouds, and hence should be given equal consideration. So I think the chapter needs some mention beyond sensitivity to CO2 CONCENTRATION and also consider sensitivity OF the concentration. TCRE is becoming a metric to quantify this and should be discussed here with equal weight as ECS and TCR. [Chris Jones, United Kingdom of Great Britain & Northern Ireland]	Noted. Whilst we recognise that there are Earth System feedbacks beyond the physical feedbacks that we include here, these are either covered elsewhere in the report, or do not contribute specifically to the climate sensitivity to doubling CO2. In the case of carbon cycle feedbacks both of these conditions hold - they are covered extensively in Chapter 6, and by definition do not contribute to the climate sensitivity to doubling CO2 (because they act through changing the CO2 concentration). We therefore exclude carbon cycle (and other biogeochemical) feedbacks here, but have included text to justify this and to point to the more detailed analysis in Chapter 6. The TCRE metric is much less widely-used than ECS, so we consider it inappropriate to give TCRE equal weight at this stage, especially given page-count restrictions.
9-39	9	0				Miss an evaluation of ocean chemistry as relevant for ocean acidification (pH, saturation state with respect to calcium carbonate) [Fortunat Joos, Switzerland]	Noted. Ocean acidification is dealt with in Ch06?
9-40	9	0				References that are supposed to be embedded in the text are often given in parantheses. That should be corrected. [Farahnaz Khosrawi, Sweden]	Editorial.
9-41	9	0				It would be desirable (though probably impossible) to use similar fonds and fond sizes in all figures. [Farahnaz Khosrawi, Sweden]	Noted. Figures have been homogenised to the extent possible and sensible.
9-42	9	0				Figure quality varies a lot, some are very nice (e.g. 9.37) while others are poor. A variety of different plotting tools are used, fonts differ, some labels are unreadably small, map projections do not conform to the standards agreed (e.g. 28 and 35), symbols are shown with shadings (e.g. 44) and can't be separated. Some figures use grids, others don't, some use titles others don't, some use a,b,c) for labels (in CAPS, others not), some don't use labels. It's a lot of work but I thing this book deserves attention to figure quality. Thousands of people will read it and use it for teaching. We should set standards not just in the assessment but also in the way the information is shown. [Reto Knutti, Switzerland]	Noted. Figures have been homogenised to the extent possible and sensible.
9-43	9	0				I congratulate the authors for this chapter. Given the difficulties in CMIP5 this was a difficult task, and clearly there is much to be updated for the final draft. The chapter however is long, probably too long, and while there is a lot of scientifically interesting material, partly of it is rather descriptive. Often it's just a discussion of what the models do well and what not, but without discussion of the implications on projections. This link needs to become stronger, in collaboration with the projections chapters. Calibrated language is still quite rare in this chapter and should be used not only in the summary statements. While many figures show individual models, the chapter is still very much focussing on the ensemble as a whole in the discussion. The conclusion that some clearly to much better on many variables than other models is not highlighted in the summary. I think it should be, even if we don't know how to relate that to projections. But not all models are equal, and I think it's fair to say that, also to point out that continued investments in model development as done at the large centers does lead to models that are more credible in reproducing what we observe. [Reto Knutti, Switzerland]	Noted. Length has been reduced for final version, and material have been updated with CMIP5 output not yet available for SOD. Use of calibrated language has been improved in final version. The literature we assess does not yet support strong statements regarding connections between model performance as evaluated against historical observations (the topic of this chapter) and the quality of future projections. In the final version, we have made an attempt to better draw out conclusions derived from identifying and evaluating individual models.
9-44	9	0				The evaluation of simple models and EMICs is essentially absent in this chapter, even though the chapter is clearly supposed to talk about all types of models. There is half a page on the discussion of climate sensitivity and a table, and results are shown in one single figure out of 45. I don't think this is a balanced presentation. Of course CMIP5 is the cornerstone of this assessment, but chapter 12 has quite a bit of projections on long term changes, cumulative carbon, commitment, etc. that is based on EMICs and simple models, including also discussions of projections uncertainties that are derived from probabilistic methods e.g. using MAGICC or 2D models. A lot of the climate sensitivity/TCR discussion is also based on such models. I don't understand why this is not part of the chapter. [Reto Knutti, Switzerland]	Taken into account. The EMIC evaluation was not complete at SOD due to the fact that papers providing the evaluation were being submitted and not available until late. Regarding the evaluation of simple models, there is very little literature on evaluating them individually given that they are rather straightforward to implement and straightforward to interpret. The one issue with simple models is their use in estimating PDFs of climate sensitivity and in this case the EMICs are specifically tuned to emulate the full ESMs we

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							evaluate in this chapter. The final draft includes additional EMIC evaluation material covering land and ocean carbon-cycle fluxes during the past 20-years, and ocean heat content changes since 1970. We are including EMIC results in additional figures where they can be evaluated alongside the CMIP5 results. We include a summary discussion of the EMICs capabilities for millennial time-scale projections and their ability to handle the climate-carbon feedbacks and the ocean-climate-carbon feedbacks. A difficult issue is the EMICs are a very broad category of models and cannot really be assessed in the same way as the fully-coupled ESMs that constitute the core of this chapter. Many EMICs are tuned in ways that ESMs are not, and the models like MAGIC that are used in Ch012 are specifically tuned to emulate the ESMs, and so therefore cannot be evaluated in the way the ESMs are. An attempt has been made to better explain this in the final version.
9-45	9	0				This chapter provides an assessment of the capabilities of the climate models considered in Chapters 10–14 of this report as well as in the Atlas, focusing particularly on developments since the IPCC Fourth Assessment Report (AR4). Only the sections are commented based on the reviewer’s expertise on the SSTs, cloud, radiation and surface temperatures related in historical simulations. [Jui-Lin (Frank) Li, United States of America]	noted
9-46	9	0				Sometimes "the AR4" and sometimes "AR4" are used. Check the consistence. [Hai Lin, Canada]	Taken into account.
9-47	9	0				Check the consistence of verb tense used. Somebody "finds" and somebody "found" were used. [Hai Lin, Canada]	Taken into account.
9-48	9	0				This is a comment on the glossary, related to the term "Earth sytem model", hence I put in ch9. My recomndation is to include in the glossary the definition of "Earth System". As it is now, it seems that "Earth system model" is a model of the "climate system", so why call it "Earth system model" and not "climate system model"? And what is the "Earth System" in this context? I recommend the glossary be updated with the current vocabulary. [Elisa Manzini, Germany]	noted. An issue for the Glossary. However, we define "Earth system model" in our chapter, based on the definition of AOGCM. This is logically coherent and defensible. By contrast, it is not required, though of course possible, to define "Earth system model" through a prior definition of "Earth system". We have not chosen this route.
9-49	9	0				I urge to include a precise definition of what a model bias is, either based on a physical concept or statistical concept (difference between expected value of some estimate and the true value). This might be clear, but often, I was surprised to learn, researchers and end users mix up several concepts here. [Douglas Maraun, Germany]	Noted. The term bias in this chapter is used in the way typically understood in model evaluation: namely the difference between a time-averaged model quantity and the corresponding time-averaged observational quantity.
9-50	9	0				In the regional downscaling section I would discuss the role of GCM errors in a more prominent way. Currently it is confined to the sentence „...skill varies with ... the AOGCM used as boundary conditions“ (p 65, l 55). This does not really point out the „garbage in, garbage out“ argument. I would, either in section 9.6.3, 9.6.4 or 9.6.5 add a section on the role of GCM errors for downscaling uncertainties. Maybe some examples could be given, e.g., the position of the storm tracks and their influence on local precipitation; or the Tropical Atlantic SST bias and its potential influence on the simulation of the African Monsoon. A general discussion can also be found in Maraun et al, (2010b), already in the reference list. See also the next comment, which is directly related. [Douglas Maraun, Germany]	Taken into account. Indeed, the quality of boundary conditions is very important for downscaling performance. Within space constraints, an attempt has been made to further highlight this. More exhaustive listing of specific examples is provided in the underlying literature
9-51	9	0				Section 9.7.2: I read this section with interest. In general I would recommend synchronizing the text better with Chapters 7 and 8, in particular concerning the definition of forcing and feedback. The below are some	Accepted -- definitions of forcings and feedbacks have been cross-checked and reconciled with chapters 7

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						suggestions. [Thorsten Mauritsen, Germany]	and 8.
9-52	9	0				The notions of confidence, level of agreement and evidence need to be better defined from the beginning of the chapter. It is not clear how a given degree of confidence is established. It is said that confidence is a function of evidence and agreement but what makes the agreement high or the evidence low is not always well defined (at least not until the end of the chapter). The only section where level of agreement and evidence are clearly defined is in the summary p.60 l.27-44. It would be useful to adopt the same clarity everywhere in the chapter. However even p.60 l.27-44 where the use of agreement and evidence is clear, the notion of confidence is not used in this paragraph so it does not help understanding how a level of confidence is established precisely. Does a high agreement with low evidence lead to a medium confidence? [RYM MSADEK, United States of America]	Taken into account. A footnote has been included at the first occurrence of the calibrated language.
9-53	9	0				Fig 9-45 (p 215) is the first figure mentioned in the chapter (p.3 l.42) and it is associated with key notions like "evidence", "confidence" and "agreement" that are used all along the chapter but are only defined at the end of the chapter. I would suggest moving this figure as Fig 9.1 and moving l.6-10 of p.78 to the earlier part of the chapter. It would also be useful to refer to Chap 1, p.2, l.35 when these terms are first used. Further, in Fig 9-45 that shows the level of confidence in x-axis and the model quality as y-axis, there are some variables like clouds or MJO for which there is high confidence despite a low-quality models. Even if confidence is defined only as a function of evidence and model agreement, it seems not logical to be able to reach high confidence for a phenomenon that is recognized as poorly represented in the models and for which observations are limited. Or does it mean that there is high confidence that the models behave poorly? I don't believe this is the point that is made but in all cases it should be clarified. [RYM MSADEK, United States of America]	Taken into account. As this is a summary figure it remains at the end of the chapter. It is referred to in the executive summary precisely because it is a summary figure. The text has been improved to clarify the manner in which the confidence statements are derived, and the specific values provided in the figure have been carefully reviewed for the final version.
9-54	9	0				This chapter evaluates the performance of AOGCM and ESMs models to be used for the future projections for the global as well as regional climate in the upcoming IPCC AR5 report. This chapter is well written and much improved as compared to FOD. [HASIBUR RAHAMAN, India]	noted. Thank you.
9-55	9	0				The table numbers are not consistent between the text and in the table (e.g., Table 9.6 (EMICs) in P76, L8 is labeled as Table 9.5 at the end of the file) [Koichi Sakaguchi, the United States]	Taken into account. Figure and table numbering have been carefully checked for the final version.
9-56	9	0				There are many duplicates in the reference list (e.g., Arora, V. K., et al., 2011a in P85 L38 and L41) [Koichi Sakaguchi, the United States]	noted. The system for including citations has been improved and such errors have been corrected in the final version.
9-57	9	0				Numbering of sub-sections is often not consistent. I point them out below. [Koichi Sakaguchi, the United States]	Taken into account. Some changes to sub-section numbering have been made in the final editing of the chapter
9-58	9	0				in general a very complete presentation [CLAUDIA STUBENRAUCH, France]	noted. Thank you.
9-59	9	0				Of the three chapters I read and reviewed this was by a long shot the best draft. There was real inter-section consistency in style and depth and the structure was clearly laid out. [Peter Thorne, United States of America]	noted. Thank you.
9-60	9	1	0	0	0	It is dishonest to talk about multi-model mean results in this chapter when modelling teams exchange software routines or modify them to bring them more into line with models that receive favourable comment. You are, in effect, implying an independence between models that is simply untrue. (I note that this is finally mentioned on page 80, at almost the end of the chapter. It should be mentioned at the outset so that readers are not deceived into believing an independence that does not exist.) [John McLean, Australia]	Rejected. There is no implication of model independence in the evaluation of model performance. The evaluation is conducted on the collection of models that is available, and the quantitative performance metrics are applied to models as individuals (as illustrated in most of the figures). We do try to provide summary statements regarding the multi-model ensemble (and its mean) by way of synthesizing results and because the ensemble is used in later chapters, where the issue of model independence is more prominent.
9-61	9	1	1	1	3	The chapter is called "Evaluation of Climate Models". But the methodology of evaluation seems focused on looking for signs of agreement between models and data. There is no discussion of hypothesis testing. Models embed hypotheses that may, in principle, be wrong. While there are discussions of errors and deficiencies in	Rejected. The methodology, as explained in the chapter, is to objectively confront model results with corresponding observations and to evaluate

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						small-scale local processes, there seems to be a working assumption that anything on the large scale is basically correct in the models, and model evaluation is only for the purpose of looking for signs of fidelity between models and data sets. To the extent discrepancies are observed, no matter how numerous, this does not seem to affect the prior judgment that the models are basically accurate. So, for the purpose of model evaluation, what would constitute a testable hypothesis? If you are going to recommend we have high confidence in the models as forecast tools, presumably they must be falsifiable. [Ross McKittrick, Canada]	performance using quantitative metrics where possible. Such metrics are applied and small and large scales. Where discrepancies are large, they are clearly noted and illustrated. There is no prior judgment that models are accurate, but it is noted that they are based on physical principles. The 'test' underlying this chapter is that physically-based models are able to reproduce observed climate (its mean state, its variability, and its long-term forced trend).
9-62	9	1	1	137	56	This Chapter is written very nicely, however small errors especially reference quoting and duplication of references are present. I have added one more aspect of evaporation minus precipitation on model SST bias in page 32 of chapter 9 from recently published article Pokhrel et al (2012) [SAMIR POKHREL, INDIA]	Taken into account. Citation errors have been corrected in final draft.
9-63	9	1	1	218	1	I think it is worthwhile noting somewhere in this chapter that improvements in the CMIP5 data/metadata definitions, quality control of data, publishing of data and revision, and model descriptions (via the multi model questionnaire) has improved dramatically compared to the approach taken in CMIP3. This is likely to assist with more efficient (and comprehensive, although some of this will occur post AR5) analysis and through more accurate evaluation (due to less errors and analysts and evaluators being better aware of what the data products represent). [Mark Collier, Australia]	Noted. Intro paragraph already refers to increasing quality in documentation.
9-64	9	1	1			General comment: I would like to thank the authors for collating a comprehensive review of current climate models. They have reviewed a tremendous amount of papers and evidence, providing some feeling for the performance of models versus observations, a prerequisite for discussing climate change integrations using these models. Thank you! However, I cannot resist the general comment that in particular the latter half of the chapter is 'over structured' and I will explain this in my comments below. There is some lovely detail, but sometimes the larger picture is getting lost in between the details. Sometimes less is more, even though I understand the temptation to highlight some details that are close to the heart of one of the authors, but at least it should be made clear how these details relate to the bigger picture of capturing climate and climate change. [Peter Braesicke, United Kingdom]	Taken into account. Thank you. Some of the detail have been reduced as edits are made to reduce length. Section beginnings now have a more systematic account of why we evaluate a particular aspect of the model solutions.
9-65	9	1	15			Add initial "J." Before "David Neelin" in contributing author list (i.e. to read "J. David Neelin") [J. David Neelin, United States of America]	Taken into account.
9-66	9	1	35			Not sure about the "tuning" in the title, it is part of the development, we confront the model with observations and change/optimize parameter settings [Peter Braesicke, United Kingdom]	noted. The box describes the model development process and the particular definition of 'tuning' intended here.
9-67	9	1	43			Not sure about the title, is it more "Sub-models and how they simulated elements of climate" [Peter Braesicke, United Kingdom]	noted. Section headings were chosen to separate evaluation of mean state and historical trend (9.4) from variability (9.5). Organization within 9.4 is by model component.
9-68	9	1	43			Where is the discussion of simulating trends in OH and the lifetimes of the GHG? There are some papers published/submitted on this topic. At least there should be a disclaimer that this is not evaluated here - in a revised 9.4.6 on composition, aerosols, chemistry, ozone. [Michael Prather, United States of America]	Rejected. We have made a choice on the climate phenomena and species we evaluate in the chapter and have given reasons for this choice in the individual sections. Stating what we have not evaluated would unnecessarily lengthen the chapter. We are also not aware of a study that evaluated OH and lifetimes of GHGs in the CMIP5 models, the set of models we are focusing on in Chapter 9. We have not introduced a new subsection on composition, see response below.
9-69	9	1	44	1	49	Is the order right? Atmosphere/Aerosol/Ocean/Sea ice/Land surface/Carbon cycle [Peter Braesicke, United Kingdom]	noted. order chosen to represent major physical components (atmosphere, ocean ...), followed by

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							earth system attributes (carbon, aerosols).
9-70	9	1	44			Section 9.4.1. is misleading in that it combines both physical atmospheric (precip) and composition and chemistry (ozone) - in the same way that aerosols are singled out, it would be best to have a clear pointer for where composition is treated. Add a new subsection if possible or merge with aerosols. Better to include composition in 9.4.6. [Michael Prather, United States of America]	Rejected. In the outline of the chapter, ozone changes are discussed together with associated climate impacts like trends in SH surface climate and lower stratospheric temperatures. By not separating composition from the atmosphere we can provide a more direct link to how composition changes matter for climate.
9-71	9	1	55	2	4	Too many sub-section, "value added"/"skill"/"errors and uncertainties" are not independent enough [Peter Braesicke, United Kingdom]	Taken into account. Material under the subsection headings has been revised and reorganized, and discussed in the order of skill, value added, and errors and uncertainties, providing a better flow.
9-72	9	1		200		1. Legal Disclaimer. IPCC and/or any of its representatives/associates/affiliates/divisions/governing bodies/subsidiaries will not use my name in the IPCC documents and publications, unless they make the entire text of the dissenting minority opinions expressed in my review available to general public, the text of this Disclaimer included. In no case will they mention me as a person who had endorsed or otherwise approved the presently reviewed Draft, unless fraudulent content is removed in the final version of the IPCC Fifth Assessment Report. [Igor Khmelinskii, Portugal]	noted. As clearly laid out in the review process, all review comments and responses to them will be made public after publication of the report.
9-73	9	1		200		2. This and the following Paragraphs, up to and including Paragraph 8, refer to the entire Chapter 9. Chapter 9 is the key part of the entire Report, as it is supposed to discuss the climate models, which allegedly provide (the only existing) material evidence that the warming observed in the second half of the 20th century is caused by greenhouse gases generated by humanity, which I will henceforth refer to as the "Anthropogenic Global Warming" (AGW) hypothesis. In fact, apart from models, there is no other way to establish cause-and-effect relation between greenhouse gases (most importantly, carbon dioxide) and climate, as we are unable to perform well-controlled experiments on our climate system. Interpretation of historic climate data does not provide any alternative demonstration of such relation, as any such interpretation is based on the same (wrong, as I shall demonstrate) climate models, and thus amounts to circular reasoning. [Igor Khmelinskii, Portugal]	noted. This chapter concerns evaluation of climate models by quantitatively comparing model results to corresponding observations. The use of such models to attribute observed climate change to anthropogenic forcing is covered in Chapter 10.
9-74	9	1		200		3. The Scientific Method had been defined, for example, by Richard Feynman (Feynman, Richard (1965), The Character of Physical Law, Cambridge: M.I.T. Press, ISBN 0-262-56003-8.; p. 156) as follows: "In general we look for a new law by the following process. First we guess it. Then we compute the consequences of the guess to see what would be implied if this law that we guessed is right. Then we compare the result of the computation to nature, with experiment or experience; compare it directly with observation, to see if it works. If it disagrees with experiment it is wrong." As a consequence of this definition, a single piece of contradictory evidence is sufficient to reject a hypothesis, whereas no amount of corroborating evidence may prove or confirm a hypothesis - by stating otherwise one would commit a logical fallacy called "affirming the consequent/denying the antecedent" (see http://en.wikipedia.org/wiki/Fallacy). The logical fallacy of this Chapter is in making the (implicit and ever present in the Report) statement that it is the anthropogenically produced carbon dioxide that is causing the global warming, based on the knowledge that carbon dioxide and other greenhouse gases affect climate to some extent. In other words, they accept the AGW hypothesis as final truth, without even trying to use the Scientific Method and test the hypothesis. By doing that, this key Chapter and the entire Report assume a non-scientific dogmatic approach, as all of the previous Reports uniformly do, which necessarily and inevitably produces non-scientific (and, as I shall demonstrate, fraudulent) conclusions. [Igor Khmelinskii, Portugal]	noted. This chapter concerns evaluation of climate models by quantitatively comparing model results to corresponding observations. The use of such models to attribute observed climate change to anthropogenic forcing is covered in Chapter 10.
9-75	9	1		200		4. In effect, Chapter 9 failed to address two key questions, which must be addressed before one tries using the models for understanding present and future climate: (1) Do the models represent the physics of our terrestrial climate system correctly? (with the emphasis on "correctly") (2) Have the predictions made by these same (or slightly modified) models five, ten or fifteen years ago become true? These should be the questions to ask before one tries making any climate predictions and/or policy recommendations based on the models. I shall address these questions in paragraphs 5, 6, and 7, and 8, respectively. [Igor Khmelinskii, Portugal]	noted. This chapter concerns evaluation of climate models by quantitatively comparing model results to corresponding observations. Results presented in this chapter directly address the question of how well physically-based models of the climate system are able to reproduce observations. Evaluation of short-

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							term climate predictions is covered in Chapter 11.
9-76	9	1		200		5. I will discuss only two of the publications that allow us to reject the AGW hypothesis, although, according to Paragraph 3, a single piece of contradictory evidence is sufficient to reject a hypothesis. The first one is the paper by Lindzen and Choi ("On the determination of climate feedbacks from ERBE data", R. Lindzen, Y.-S. Choi, Geophys. Res. Lett. 36, 2009, pp. L16705. doi:10.1029/2009GL039628.). These authors demonstrate in their Fig. 2 and in the rest of the paper that the IPCC climate models distort the essential physics of the terrestrial climate system. Indeed, all of the climate models produce a reduction in the outgoing infra-red radiation upon an increase in surface temperature on Earth, whereas the experimental results, extracted from satellite data, evidence an increase in the outgoing infra-red radiation, in the same conditions. Therefore, our terrestrial climate system behaves as if it were in a stable state of equilibrium: indeed, its behaviour corresponds to the Le Chatelier's principle (the system always reacts to any change in such a way that the externally imposed change is partially compensated; see, for example, http://en.wikipedia.org/wiki/Le_Chatelier%27s_principle), whereas the climate models behave like an unstable system, amplifying all and every temperature change. We therefore see that the models that reproduce the conditions and conclusions of the AGW hypothesis do so at the cost of distorting essential physics of the terrestrial climate system. Therefore, the AGW hypothesis is wrong, as it has not been reproduced in models describing the climate physics correctly. Moreover, the AGW hypothesis can not be reproduced in models describing climate physics correctly, as it is impossible to attribute the warming of the 20th century to carbon dioxide based on correct models. The important conclusion that necessarily and inevitably follows from this paper is that the AGW hypothesis is wrong, as it is only viable in the virtual reality of the IPCC climate models, fundamentally different from the physical reality of the terrestrial climate system. [Igor Khmelinskii, Portugal]	noted. Assessment of models regarding radiative transfer is covered in Chapter 8.
9-77	9	1		200		6. The second publication I will discuss is the one by Khmelinskii and Stallinga ("Climate Change in the XXIst Century: Mechanisms and Predictions", I. Khmelinskii and P. Stallinga, in Proceedings of the 6th IASME / WSEAS International Conference on ENERGY & ENVIRONMENT (EE '11), Cambridge, UK, February 20-25, 2011, eds. Z. Bojkovic et al., RECENT RESEARCHES in ENERGY & ENVIRONMENT, WSEAS Press, 2011, ISSN: 1792-8230; ISBN: 978-960-474-274-5, pp. 26-31. Available: http://www.wseas.us/e-library/conferences/2011/Cambridge/EE/EE-02.pdf). These authors in their Fig. 1 and in the text of the paper analyze the recent history of the global average Sea Surface Temperature (SST) and compare it to the recent history of atmospheric carbon dioxide concentrations, in an attempt to find the signature of the AGW in the SST data. In fact, they could find no such signature, due to the fact that human carbon dioxide emissions started growing exponentially in the second half of the 20th century, whereas SST had two (virtually identical) growth periods, one of which in the first half of the 20th century, when little or no excess carbon dioxide had been liberated into the atmosphere by humanity. These authors therefore conclude that the AGW hypothesis has to be rejected, based on the recent SST history. Note that SST is a better indicator of the climate evolution than the global average temperature, being unaffected by interfering factors such as the Urban Heat Island effect. The latter distorts climate data gathered on the continents, with additional uncertainty introduced by the corrections made to compensate for it. [Igor Khmelinskii, Portugal]	noted. Relationships between greenhouse gas forcing and climate response are discussed in Chapter 10. In this chapter, the ability of models to reproduce observed changes in SST and ocean heat content is objectively and quantitatively evaluated.
9-78	9	1		200		7. In Paragraphs 5 and 6 I discussed two papers, each of the two providing sufficient grounds to reject the AGW hypothesis. I shall not discuss any further evidence against the AGW hypothesis, considering it rejected, according to Feynman's definition presented in Paragraph 3. Additionally, I conclude that the IPCC climate models are wrong, as they obviously distort the essential climate physics, and therefore any and all of their results and conclusions should be expressly and unconditionally rejected and disregarded in their entirety. [Igor Khmelinskii, Portugal]	noted. This chapter concerns evaluation of climate models by quantitatively comparing model results to corresponding observations. Both the models and observations contain errors and uncertainties and these are discussed throughout the report. Disagreement between models and observations indicates model shortcoming and areas for improvement, but does not imply that the essential physics is incorrect. The extent to which such uncertainties affect attribution of climate change to anthropogenic forcing is discussed in detail in Chapter 10.
9-79	9	1		200		8. It is well known that there has been no global warming for the last 15 years, contrary to the IPCC predictions produced by IPCC climate models for the same period of time. Moreover, we have reasons to	noted. Solar forcing of the climate system is discussed in Chapter 8. A new box is being included in Chapter

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						believe (see, for example, H. Abdussamatov, 2008, The Sun defines the Climate, http://www.gao.spb.ru/english/astrometr/abduss_nkj_2009.pdf) that instead of the "global warming" we are in for a new Little Ice Age, already in progress, which will be similar to the Maunder minimum of solar activity as regards temperatures and other climate-related consequences. Therefore, the IPCC models have not (because they distort climate physics) and will not (for the same reason, and also because they neglect solar change) predict future climate, and thus should be rejected and disregarded, as I have obtained negative answers for the two key questions of Paragraph 4. [Igor Khmelinskii, Portugal]	9, based on various lines of evidence, discussing the reduced rate of warming in the recent decade and the possible causes.
9-80	9	1				Thank you for the improvements. Although I still don't like some of the literature, the representation of it is much better. [James Annan, Japan]	noted.
9-81	9	1				As explained in the comment above for chapter 7, I feel that the report is missing the point of cloud representation in current climate model : how it performs, how it has improved from AR4, and how it compares with the new available cloud observations (in particular Cloudsat/Calipso observations). It can be done either in chapter 7 or 9 but I would be a little bit upset if it was not discussed seriously somewhere in the AR5. [Frédéric HOURDIN, France]	noted. This is intended to be covered in Chapter 7 and so only limited discussion of this topic appears in Chapter 9.
9-82	9	1				The chapter is lacking overall references to major international climate observing bodies and activities, like e.g. GCOS or the ESA Climate Change Initiative. I recommend to revised the chapter accordingly [Alexander Loew, Germany]	Rejected. This chapter provides an assessment of climate models, not the bodies/activities involved in climate observations. Observations per se are covered in chapters 2-4.
9-83	9	1				The supplement to Chapter 8 in AR4 provided a highly valuable means of assessing the errors in temperature, precipitation, and the like in individual models as a function of location, altitude, latitude, etc. I would urge a similar supplement to the present chapter. [Stephen E Schwartz, United States of America]	Rejected. After extensive discussion, we have decided not to provide extensive supplementary material, favouring identification of individual models in the figures presented in the main body of the chapter.
9-84	9	1				Unfortunately, I do not have time to review the remainder of this chapter either. [Gunilla Svensson, Sweden]	noted
9-85	9	2	5	2	11	Many thoughts in this section should be used earlier to frame the chapter better [Peter Braesicke, United Kingdom]	Taken into account. Some attempt to better lay out organization of the chapter early on has been made during final edits.
9-86	9	2	12			Drop "and how would we know" [Peter Braesicke, United Kingdom]	rejected. Title of FAQ was extensively discussed and decision was that this phrase was important.
9-87	9	2			83	This chapter is well written. Minor issues previously noted have been corrected. [Larry Thomason, United States of America]	noted. Thank you.
9-88	9	3	1	3	36	This section is essentially introductory material and not an executive summary. Suggest to shorten this significantly. [Reto Knutti, Switzerland]	Accepted. Executive Summary has been thoroughly revised.
9-89	9	3	1	5	44	This Section can be improved. As an Executive Summary, it should stay focussed on concisely delivering the key messages from model evaluation studies and not stray into discussion. [Ian Smith, Australia]	Accepted. Executive Summary has been thoroughly revised.
9-90	9	3	1	5	44	It should stick to model evaluation and not stray into model projections. [Ian Smith, Australia]	noted. However, other reviewers have asked for more connection to projections.
9-91	9	3	1	5	44	It should not be referring to "confidence" with respect to model evaluations. [Ian Smith, Australia]	rejected. Throughout the report the confidence of various assessment statements is conveyed through careful use of calibrated language.
9-92	9	3	1	5	44	As with other Chapters (e.g. Chapter 12) the Executive Summary does not need to refer back to individual Chapter Sections and Boxes. [Ian Smith, Australia]	rejected. Traceability to individual sections is a requirement imposed on the report.
9-93	9	3	1	5	44	Please be consistent and refer to CMIP3 model results rather than AR4 results (or vice-versa). [Ian Smith, Australia]	noted. This has been carefully checked, but typically we refer to CMIP3 models, but also refer to assessment results as described in AR4.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-94	9	3	1	5	44	There is nothing in the ES about the upper tropical troposphere even though this is included in ES 12-3 line43. [Peter Stott, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Executive Summary has been thoroughly revised and this material is included in paragraph three.
9-95	9	3	1	5	46	There should be something in ES about aerosols - (9.4.6)? [Michael Prather, United States of America]	Taken into account. Executive Summary has been thoroughly revised. Aerosols are now discussed in the ES with specific attention to interactive aerosol components, the representation of the sulfur cycle, and the comparison with satellite based observations of aerosol optical depth.
9-96	9	3	1			Executive Summary: It would be good to include a comment on the contribution of increasingly complex models (e.g. ESMs; high top models) to the CMIP5 ensemble, particularly in relation to the advances made in representing basic model processes. Is there evidence that adding complexity is improving confidence in model projections, or are systematic biases still dominant? A statement on this would be helpful in informing plans for future model development. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Executive Summary has been thoroughly revised.
9-97	9	3	3	3	4	Suggest deleting these lines as they represent Introductory material, not a summary [Ian Smith, Australia]	Taken into account. Executive Summary has been thoroughly revised.
9-98	9	3	5	3	5	Suggested text: "...has increased since AR4." [Ian Smith, Australia]	Taken into account. Executive Summary has been thoroughly revised.
9-99	9	3	7	3	7	Suggested text:"In addition to coupledAtmosphere-Ocean..." [Ian Smith, Australia]	Taken into account. Executive Summary has been thoroughly revised.
9-100	9	3	8	3	10	Suggested text:"...predictions, this assessment has considered AOGCMs with biogeochemical cycles, referred to as Earth System Models (ESMs). [Ian Smith, Australia]	Executive Summary has been thoroughly revised and this description of ESMs is more complete.
9-101	9	3	10			The word "composition" might be good to use [Peter Braesicke, United Kingdom]	Taken into account. Executive Summary has been thoroughly revised.
9-102	9	3	12	3	12	Probably it is worth to point out that this time many RCM modelling groups managed to downscale historical simulations and scenarios from global CMIP5 simulations. Thus the new RCP scenarios can be studied on the global domain as well as in regions. In the past the regional scenarios were often based on older global scenarios. [European Union]	Taken into account. Executive Summary has been thoroughly revised.
9-103	9	3	12	3	15	Suggested text: "...(RCMs), which are used extensively in downscaling global scale climate results to regional scales, statistical downscaling (SD) methods, and Earth System Models of ..." [Ian Smith, Australia]	Taken into account. Executive Summary has been thoroughly revised.
9-104	9	3	13			And also later in section 9.6: Is there a policy of continuity of acronyms across chapters? SD is used in Chapter 4 for snow depth and here for "statistical downscaling". [Adrian Simmons, United Kingdom]	Taken into account. Acronym usage is being minimized, and made consistent via Glossary.
9-105	9	3	15	3	17	A group of models that is not being included in Ch9's assessment are the Simple Climate Models. However, the simple climate model MAGICC is used quite a bit in (i) calculating CO2 concentrations from IAM emissions which then serve as input for the AOGCMs/ESMs, and (ii) in projections of climate change in the 21st century and beyond in Ch12. It seems that this would warrant some mention here in Ch9. [Thomas Stocker/ WGI TSU, Switzerland]	Rejected. The discussion/evaluation of EMICs is being enhanced in the final version with attention to adequacy of using them for both long-term projections and probabilistic assessments based on calibrated ensembles. But simple models like MAGICC are not covered in the model evaluation chapter. Their use in computing CO2 concentrations is not really part of the AR5 WG1 -- the CMIP5 protocol takes the CO2 concentration as input. The comparison of EMICs and ESM simulated aspects of the carbon cycle is provided in Chapter 6. The use of MAGICC in Chapter 12 is limited to 'emulating' more comprehensive climate models, and so it is specifically tuned to reproduce the results of the models that are evaluated, and its simplicity precludes an assessment

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							of feedback analysis; therefore the metrics used to evaluate climate models are not relevant in the case of model emulation using simple climate models.
9-106	9	3	18	3	18	Another novelty is the coordination of idealized experiments, which support the analysis of processes and feedbacks. Thus the concept of model inter comparison can be extended to simulations designed to measure particular characteristics of climate models. In the past such simulations were done less systematically, i.e. mostly by individual groups. [European Union]	noted
9-107	9	3	19	3	19	"The evaluation of climate models requires the availability of high quality observational data" This is true, but not only. We also need the availability of a process-based analysis of data, since models are not able to reproduce all processes of nature, and we also need theoretical understanding. [Government of France]	noted. Aspects of process understanding are covered in Chapters 6 through 8.
9-108	9	3	19	3	20	I would suggest to split the first sentence into two and combine the second part of the first sentence with the third sentence. Thus, I would rewrite the second sentences as follows (first sentence ends at the comma): The uncertainty of these data also needs to be understood as described in chapter 2.5. [Farahnaz Khosrawi, Sweden]	Taken into account. Executive Summary has been thoroughly revised.
9-109	9	3	19	3	25	A more balanced approach to performance metrics would bet better; they are important for traceability and characterisation of models; but they do not answer all question. They depend on data and error descriptions. Choices of metrics are subjective and should be clearly linked to a purpose. I doubt that a best model exists. Please phrase mre carefully. [Peter Braesicke, United Kingdom]	Taken into account. Executive Summary has been thoroughly revised. The discussion of multiple lines of evidence for evaluation is used throughout the chapter, and an assessment of the models is presented accordingly. In the final paragraph of the ES, the discussion focuses on specific concerns about using metrics as objective weights for selecting a perfect model, which is rejected as premature.
9-110	9	3	19	3	25	This paragraph is a good example of the overall bias of the chapter. Nowhere in this paragraph does the reader get even a hint of the fact that these "advances" in model evaluation though the use of "performance metrics" have shown virtually no skill in differentiating different model predictions of climate change. We religiously cling to the idea that simple performance metrics will tell us which model is better for its intended purpose while ignoring the scientific evidence that the metrics currently developed simply do not work for what we want them to do (e.g., Pincus et al. 2008, Klocke et al. 2011, and Collins et al. 2012 for climate sensitivity, but a similar lack of evidence for telling good from bad models in other aspects of climate change also exists). [Anthony Del Genio, United States of America]	Taken into account. Executive Summary has been thoroughly revised. The remit of this chapter is clear from the title, "evaluation of climate models".
9-111	9	3	19	3	25	In AR1 was a Chapter entitled "Validation of climate models" In the First Draft of AR2 was a similar Chapter. I commented at the time that since no attempt at genuine validation was being attempted, the Title was inappropriate. To my surprise, you agreed with me, and in the next Draft you not only altered the title to "Evaluation of climate models" but you also changed the word ":validation": to ":evaluation" no less than fifty times throughout the Chapter. Since then, the word "validation" has been forbidden. Not only that, you also banned the use of the word "prediction: and replaced it by the word ":projection": All this is an admission that none of your midels are capable of "climate prediction"; all you get are "projections" where yoi have to believe the initial assumptions before you take any notice of them. "Validation" would require an extensive series of tests to discover the predictive capabilities of the models in all circumstances for which they are to be used, to a satisfactory level of accuracy. Not only has this never been done, there has, up to now, been no discussion on how it may be done. Because of this it is possible to assert that the models shoukd not be used for forecasts until it has been done. You have chosen the lesser procedure of :Evakuation: which essentially relies on educated guesses mae by the modellists themselves, who have a conflict of interest in the matter, and therefore should not be believed. ,"Simulation" is merely another word .for "correlation" which never proves cause and effect, Your entire exercie depends on the opinions of people who have a conflict of interest in judging the effectiveness of midels, when their true eggettiveness has never bee tested.. [Vincent Gray, New Zealand]	noted. Model evaluation is defined in section 9.1 as involving quantitative comparison of model results with corresponding observations. This word has come to be used in the scientific literature to clearly distinguish this process of confronting models with observations to assess their ability to represent the climate system, from the 'validation' of predictions made with such a model (by comparing a prediction with an observed value after the fact). Predictions and predictability are described in detail in Chapter 11 (where the term 'prediction' is clearly defined). Predictions are distinct from projections, the subject of Chapter 12. Simulation is not a synonym for correlation. A simulation is the result of a physically-based model, and this is clearly distinct from 'correlation' which is a statistical concept.
9-112	9	3	19	3	36	Suggest omitting all this text as it does not represent a summary. It is Introductory material that belongs in another part of the Report. [Ian Smith, Australia]	Taken into account. Executive Summary has been thoroughly revised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-113	9	3	20	3	24	Despite of the emphasis on the 'performance metrics' in Executive Summary (P3 I20-I24), there is no paragraph which summarizes the recent development of cloud metrics. [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	noted. The Executive Summary is intended as a synthesis of the main results; it cannot provide details on all types of metrics. Clouds and their representation in climate models is dealt with in detail in Chapter 7.
9-114	9	3	20	3	25	An important advance in model evaluation since the AR4 is the quantitative evaluation of model performance in simulating the vertical profiles of clouds, water vapor and aerosols, which have been enabled by the A-Train constellation. A number of studies have evaluated CMIP5 model performance in simulating the clouds and water vapor vertical profiles (e.g., Jiang et al., 2012; Li et al., 2012; Su et al., 2012; Tian et al., 2012). [Government of United States of America]	noted. The Executive Summary is intended as a synthesis of the main results; it cannot provide details on all types of metrics. Clouds and their representation in climate models is dealt with in detail in Chapter 7.
9-115	9	3	20			"Observational data are described in Chapters 2-5." Comment: The observational data described Chapter 2-5 are incomplete. For example, the most advanced NASA satellite constellation, the "A-Train" satellite observations (L'Ecuyer and Jiang; 2010) has not been introduced or mentioned in Chapter 2. However, a large number of recent papers have evaluated performance of CMIP5 models using the A-Train satellite datasets (e.g. Jiang et al. 2012, Li et al. 2012, Su et al. 2012, Stephens et al. etc.). References: L'Ecuyer, T.S., and J.H. Jiang, "Touring the atmosphere aboard the A-Train," Physics Today 63, 7, 36-41, 2010. Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. Su, H., J.H. Jiang, C. Zhai, V.S. Perun, J.T. Shen, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, C. Morcrette, J. Petch, M. Ringer, J. Cole, M. Mesquita, T. Iversen, J.E. Kristjansson, A. Gettelman, L. Rotstayn, S. Jeffrey, J.L. Dufresne, M. Watanabe, H. Kawai, T. Koshiro, T. Wu, E.M. Volodin, T. L'Ecuyer, J. Teixeira, and G.L. Stephens, "Diagnosis of Regime-dependent Cloud Simulation Errors in CMIP5 Models Using A-Train Satellite Observations," J. Geophys. Res., in press. [Government of United States of America]	noted. This comments should be directed to Chapter 2.
9-116	9	3	20			"Observational data are described" -- please change to "Observational data and information from paleoclimate archives are assessed" [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Executive Summary has been thoroughly revised.
9-117	9	3	21	3	24	It is important to also emphasize that another advance in model evaluation since the AR4 has been the use of vertically resolved global observational metrics, e.g. for cloud and water vapor, (e.g. Jiang et al. 2012, Figure 11). These new metrics are also multivariate metrics, i.e. use collocated multi-variable observations to evaluate more than one variables at same time, and also examining the relationships of the observed and simulated variables to large-scale environment (e.g. Su et al. 2012). References: Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. Su, H., J.H. Jiang, C. Zhai, V.S. Perun, J.T. Shen, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, C. Morcrette, J. Petch, M. Ringer, J. Cole, M. Mesquita, T. Iversen, J.E. Kristjansson, A. Gettelman, L. Rotstayn, S. Jeffrey, J.L. Dufresne, M. Watanabe, H. Kawai, T. Koshiro, T. Wu, E.M. Volodin, T. L'Ecuyer, J. Teixeira, and G.L. Stephens, "Diagnosis of Regime-dependent Cloud Simulation Errors in CMIP5 Models Using A-Train Satellite Observations," J. Geophys. Res., in press. [Government of United States of America]	noted. The Executive Summary is intended as a synthesis of the main results; it cannot provide details on all types of metrics. Clouds and their representation in climate models is dealt with in detail in Chapter 7.
9-118	9	3	27	3	36	ES : « designed specifically for this purpose » is not necessary. « similiar » -> similar. Lines 33-34 are misleading, as the inter-model spread does not assess the same uncertainty as model-observation	Taken into account. Executive Summary has been thoroughly revised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						comparisons. Lines 34-36 : these lines should make it clear how difficult it is generally (and often impossible) to relate model errors at the climate scale to errors in model formulation or processes. Documenting the models is very important, it may actually help to make progress in that direction, but probably not as much as suggested by these few lines. Physically-based approaches should be developed for this purpose. I would make the lines 35-36 shorter (e.g. « ..by the more extensive documentation of the models since the AR4. »). [Sandrine BONY, France]	
9-119	9	3	27	3	36	A reference to simulations of the past climates (glacial, holocene) is missing in this paragraph as well as a bullet on past climates. Even if uncertainties are associated with paleodata and conditions of experiments, simulations prior to the historical period are the only way to evaluate models on a large climate change. Moreover for the first time simulations have been done with the same models within CMIP5. [SYLVIE JOUSSAUME, France]	Taken into account. Executive Summary has been thoroughly revised.
9-120	9	3	27	3	36	The availability of coordinate multi-model experiments is not a novelty - indeed CMIP is at its phase 5. This paragraph could be made more informative, if it reports what is special and new in CMIP5 with respect to CMIP3. Also "however comprehensive analysis is just getting started" does not seem to be relevant here (it sounds like text for a project report)" [Elisa Manzini, Germany]	Taken into account. Executive Summary has been thoroughly revised.
9-121	9	3	27			There seems to be an over emphasis on the multi-model mean. As the authors well know, the models can't all be right. Somehow the impression is that the multi-model mean is the best representation of the output of the models, but this can hardly be correct; it is just that the best model is not known. This point should be underscored in the text where the multi-model mean is introduced and justified [Stephen E Schwartz, United States of America]	Taken into account. Executive Summary has been thoroughly revised. This paragraph does not mention multi-model mean, but rather the availability of a multi-model ensemble. In the body of the chapter, the individual members of the ensemble are shown and metrics computed for each individually.
9-122	9	3	28	3	28	CMIP3 and CMIP5: Some readers may wonder about CMIP4. It may be worth to add a footnote explaining that the name "CMIP4" was skipped to avoid confusion with "C4MIP". Additionally CMIP"5" has the advantage to be in synch with AR"5". [European Union]	rejected. This is a nomenclature detail not suitable for the Executive Summary.
9-123	9	3	28	3	28	...Projects (CMIP3 and CMIP5), allows for ... [Hai Lin, Canada]	Taken into account. Executive Summary has been thoroughly revised.
9-124	9	3	29	3	32	Why say "historical" observations - are there any other kind? Rather than "historical period" say "recent historical period" or else it implies 5B yrs. "...available in large numbers" - what does that mean - data size, number of exp, variables ?? - could be better. [duane waliser, United States of America]	noted. Executive Summary has been thoroughly revised. General usage is that 'historical' refers to the period over which human records have been kept; as opposed to pre-historical or paleo indicators.
9-125	9	3	30	3	30	Large number of what? Do you mean a large number of Publications? [Farahnaz Khosrawi, Sweden]	noted. Executive Summary has been thoroughly revised. Large number refers to the number of models and their output.
9-126	9	3	30	3	30	analyses of what? Be more precise. [Farahnaz Khosrawi, Sweden]	rejected. Analysis of the model results is implied.
9-127	9	3	31	3	31	"smimiliar"à"similar" [Jui-Lin (Frank) Li, United States of America]	Editorial.
9-128	9	3	31	3	31	similiar --> similar [François Massonnet, Belgium]	Editorial.
9-129	9	3	31	3	31	similiar -> similar [Bart Van den Hurk, Netherlands]	Editorial.
9-130	9	3	31	3	32	But note that direct comparison of models with historical observations needs to be conducted with care since "good" models could have "poor" forcings imposed. Note also that even though they have been given common inputs the forcings could be very different - eg similar sulfate emissions could leave to very different forcings from indirect effects. [Peter Stott, United Kingdom of Great Britain & Northern Ireland]	noted. This is covered to some extent in the body of the chapter. Executive Summary is intended to be concise.
9-131	9	3	31			delete "the"; why only similar? [Peter Braesicke, United Kingdom]	Taken into account. Executive Summary has been thoroughly revised. 'Similar' because the CMIP5 protocol does not constrain all details of forcing.
9-132	9	3	32	3	34	Please rephrase, an intermodal spread is useful but cannot replace observations. [Peter Braesicke, United Kingdom]	noted. Executive Summary has been thoroughly

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Kingdom]	revised. Text does not state it is a replacement, merely that it provides some information.
9-133	9	3	32	3	36	Please include a discussion of the uncertainties in observations [Government of United States of America]	noted. This is mentioned earlier in the draft ES. Text being significantly revised.
9-134	9	3	32	51		For me the external forcing for the climate system is solar; CO2 released becomes part of the system [Peter Braesicke, United Kingdom]	noted. This is a definitional issue. Here 'external' refers to 'external to the model'; i.e. something specified as input to the model. Solar and CO2 both satisfy this in the experiments in which CO2 is specified.
9-135	9	3	34	3	34	I would suggest to put the text in the parantheses " by examining inter-model spread" in an own sentence. [Farahnaz Khosrawi, Sweden]	noted. Executive Summary has been thoroughly revised.
9-136	9	3	36			Last phrase in sentence is unclear. Reword. [Ronald Stouffer, United States of America]	Taken into account. Has been clarified.
9-137	9	3	40	3	40	The list introduced in this line refers to "climate models". Previously (line 5), four classes of models have been introduced. It is not clear, if the summary starting at line 40 refers to all the 4 classes or just one. I recommed to solve this ambiguity. [Elisa Manzini, Germany]	noted. It is a generic statement. Some attempt at clarification has been made during revisions.
9-138	9	3	40	3	40	Please specify: does the historical climate period equal the instrumental period? [Thomas Stocker/ WGI TSU, Switzerland]	noted. Yes this is the intent as per common usage of the term 'historical'.
9-139	9	3	41	3	41	respects': do you mean 'aspects'? [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	rejected. 'respects' is the correct word.
9-140	9	3	41			It is suggested to substitute "respects" by "aspects". [Klaus Radunsky, Austria]	rejected. 'respects' is the correct word.
9-141	9	3	42	3	42	Write "in many models" or "many of them" instead of just "many" [Farahnaz Khosrawi, Sweden]	rejected. 'many' refers to 'respects', not models.
9-142	9	3	42	3	42	It feels very odd for the first figure referenced to be 9.45. As much as anything how do you expect the reader to easily find Figure 9.45 which presumably will be 100+ pages later in the report? [Peter Thorne, United States of America]	noted. 9.45 is the summary figure which attempts to bring together assessment results spread throughout the chapter. It is therefore the figure most relevant to the Executive Summary.
9-143	9	3	42			Figure 9.45 and the examples/discussions hereafter, why is the confidence of cloud simulation high but the model quality (for simulating clouds) low. This should be explained clearly, because it's difficult to understand. While more than half of the climate models show improvements from CMIP3 to CMIP5 in simulating column-integrated cloud amount, the quality of model simulated clouds at different pressure levels is poor. For example, for the 19 CMIP5 models studied, the modeled mean cloud water contents (CWCs) over tropical oceans range from 3% to 15 times of the observations in the upper troposphere (p<300 hPa) and 40% to 2 times of the observations in the middle and lower troposphere (Jiang et al. 2012). Reference: Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Briant, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	Taken into account. Assessment statements in the figure have been carefully checked and were subject to revision pending results of more recent model output and published papers.
9-144	9	3	42			Figure 9.45 is full of acronyms and hard to decipher, especially at the beginning of the chapter. Suggest omit mentioning it here. It is sufficient to have it explained later in the Chapter, Page 9-78. [Government of United States of America]	Rejected. 9.45 is the summary figure which attempts to bring together assessment results spread throughout the chapter. It is therefore the figure most relevant to the Executive Summary.
9-145	9	3	45	3	45	The reference surface air temperatures (2meter) are from ERA-Interim which is from assimilation that is not an "observation" which should be mentioned in the text. [Jui-Lin (Frank) Li, United States of America]	Rejected. Details regarding the observations or observationally-based information is contained in the body of the chapter. The Executive Summary is intended to communicate high-level results.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-146	9	3	45	3	45	replace "models" by "large multi-model ensembles" [Michel Petit, France]	rejected. This is a generic statement.
9-147	9	3	45	3	48	Suggested text: The AOGCMs realistically simulate the surface temperature on continental and larger scales (Section 9.4.11, Figure 9.2) and the global-scale surface temperature increase over the historical period, especially the last Fifty year (Section 9.4.1.3.1, Figure 9.8). [Ian Smith, Australia]	Taken into account. Executive Summary has been thoroughly revised.
9-148	9	3	45	3	51	This comment can be demonstrated to be misleading by a simple comparison of model output vs. observations based on the past 34 years of data. To do this, simply reference the model and observations to 1979-1983 (i.e. give them all the same starting point as in a race) and see how they spread apart through the following years (see my congressional testimony, House subcommittee on Energy and Power, 20 Sep 2012.) What I presented in the testimony is the type of information that communicates the problem with models quite well - and it seems to me this is the duty of the IPCC. Model sfc trends average 0.23 to 0.24 C/decade for 90 runs of the CMIP5 RCP4.5 1979-2012 trends. Obs. surface average 0.16 C/decade (includes some conatimation from surface development and geographic misrepresentativeness), and tropospheric trends reduced to the surface average 0.12 C/decade. This is obvious from the available data and is an important period because this is when the alleged impact of GHG should be most readily measureable. [John Christy, United States of America]	Taken into account. Comparisons of short records (34 years is 'short' in the climate context) are subject to large uncertainties related to natural variability. Box 9.2 has been added to Chapter 9 to specifically address the reduced rate of warming over the recent 15 years and the comparisons of models to observations over this short period.
9-149	9	3	45	3	51	I disagree strongly with the conclusions of this paragraph. If Chapter 7 is to be believed, then we have a long way to go in (a) knowing the climate sensitivity, and (b) knowing the climate forcing due to aerosols and cloud-aerosol interactions. It is interesting that the CMIP5 models do not show the negative correlation between forcing and sensitivity that Kiehl (2007) showed for the CMIP3 models, as stated later in the chapter, but if in fact this is the case, then doesn't it imply that the uncertainties in forcing and feedback and not adequately tested by the 20th Century temperature record? How then can one make the statement of "very high confidence that models generally respond correctly to external forcing"? It is one thing to make the statement that no model has yet been able to reproduce the upward trend in global temperature during the last half of the 20th Century without including greenhouse gas forcing (something that AR4 said but which I don't think is ever said in this chapter). That is a fair statement to make, one which puts the onus on those who disagree to come up with a model that reproduces the trend without GHG increases. But that is not the same as what you say. Perhaps you need to better define what you mean by "generally respond correctly" - if by that you mean simply that in models the temperature warms when we add GHGs, as expected from basic physical principles, that is OK. But if you want to conclude that agreement with the rate of temperature increase is a metric of model fidelity, then you must also conclude either that (a) aerosol effects on climate are known accurately enough or are small enough not to be much of a research concern going forward, and/or that (b) cloud feedback is less uncertain than we imagine or has not yet occurred to a sufficiently great extent to matter (in which case historical evidence of cloud trends that you discuss later are irrelevant). [Anthony Del Genio, United States of America]	Taken into account. Executive Summary has been thoroughly revised. Specific confidence statements has been updated as new information became available. The statement attempts to convey that fact that when models are provided with available estimates of external forcing, they respond in important aspects of their solutions in a way that is consistent with observed climate change. We have clarified the wording as Executive Summary has been revised, to reflect better the quantities for which tis is true. The decomposition of that response into individual forcings is discussed in more detail in Chapter 10.
9-150	9	3	45	3	51	There's a discrepancy between Figure 1.4 and this paragraph. You say there is "very high confidence" that models simulate realistically the global surface temperature trend especially over the last 50 years. But Figure 1.4 shows the models predict a uniform upward trend in the global average after 1990 (except the dip at the 1992 Pinatubo volcano, which was programmed in later), while the observed global average bounces around then runs flat for a long time, dropping out by 2010 below the low end of the model range. The discrepancy is obvious in that diagram. Your claim to have "very high confidence" in the model forecasts therefore needs explanation. You cite Figure 9.8 as support for the claim of a match with observations, but that Figure doesn't support it at all, in fact it shows that during the past few decades the match has gotten worse. It shows that during the pre-1992 interval, when you got to peek at the answer, the models matched (or were tuned to match) the observed trends. But after 1990 when the models could be said to have begun generating forecasts, the model temperatures clearly trend upwards while the observations do not, and the discrepancy opens up so much that by 2011 the observed mean temperature is at or below the entire range of model projections. This appears to be the longest interval in which the model mean and the observed mean do not cross, and the discrepancy is widening over time. You can't ask readers to accept at face value a claim that the models are doing better and better when the discrepancy is getting bigger and bigger. The language of "very high confidence" in this case appears out of place. [Ross McKittrick, Canada]	Taken into account. Executive summary has been thoroughly revised. A new Box 9.2 has been added to Chapter 9 to explicitly address the issue of reduced rate of warming over the recent 15 years and the apparent mismatch between models and observations over this period. Specific confidence statements have been updated as new information becomes available. Figure 9.8 shows results of historical simulations -- there are no 'forecasts'.
9-151	9	3	45	3	51	"Realistic" needs to be defined. Is this referring to the physical realism of process representations or to the	Taken into account. Executive summary has been

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						closeness to historical observations? In either case, what are the bounds of "realism"? Same for "generally responds correctly". [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	thoroughly revised. "Correct" and "realistic" are no longer used.
9-152	9	3	45	3	51	Given the importance of testing models using independent evidence (see Chapter 9, page 9, lines 1 to 17 and my comments on chapter TS, page 33, lines 51 to 54) and the likelihood that surface temperatures have been used to tune the models, the fact that surface temperatures and their recent changes are represented well in present climate models cannot be used to give "very high confidence" that models will respond correctly to changing greenhouse gases. Confidence statements about predictions need to be made on the basis of independent evidence. [David Webb, United Kingdom]	Taken into account. Exact wording in the Executive summary has been revised. The statement attempts to convey that fact that when models are provided with available estimates of external forcing, they respond in important aspects of their solutions in a way that is consistent with observed climate change. We have clarified the wording as Executive Summary has been revised, to reflect better the quantities for which tis is true. Detailed disagreement over recent 15 years is the subject of a new box in the Chapter.
9-153	9	3	45	4	26	Phrases such as "simulate realistically" or "correctly simulate" are used repeatedly but it is entirely unclear what these phrases mean. Either they need to be clearly (and consistently) defined (which at the very least requires multiple metrics) or they should be avoided. The likelihood that some metrics may show agreement with observations for spurious reasons (e.g. error cancellation) should be explicitly acknowledged. [Rowan Sutton, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Executive summary has been thoroughly revised. "Correct" and "realistic" are no longer used in the ES.
9-154	9	3	45		51	The range of equilibrium response to 2xCO2 in recent CMIP5 simulations of the global mean surface temperature (2.1 to 4.7°C, page 5, line 4) is still a large uncertainty for this basic global-mean climate sensitivity factor. This is not remarkable progress over the interval of confidence (1.5 to 4.5°C) proposed by the Charney workshop some 35 years ago. The realism of climate projections could be toned down a bit considering the remaining error bar. [Government of France]	Taken into account. Exact wording in the Executive summary is being revised. The statement attempts to convey that fact that when models are provided with available estimates of external forcing, they respond in important aspects of their solutions in a way that is consistent with observed climate change. We have clarified the wording as Executive Summary has been revised, to reflect better the quantities for which tis is true. Detailed disagreement over recent decade is the subject of the new Box 9.2 in the Chapter.
9-155	9	3	45			"There continues to be very high confidence that models simulate realistically the surface temperature on continental and larger scales". The statement begs the question of the accuracy and the use of the results. On a scale of 288 degrees, a 3 degree error is remarkable accuracy. 1 %, but as measured on a scale of important changes to climate that would result from a change in global temperature of 1 or 2 degrees, a 3 degree error assumes a much greater importance. There are further questions regarding accuracy at continental scales; regarding accuracy by season; an error of 3 degrees in the growing season might have different implications from an error in winter. Are there compensations in the annual mean from seasonal errors. . . I would ask the authors to comment on the statement of high confidence in realistic simulation of the surface temperature in the models in the context of Figure 9.14, which shows some models low in sea surface temperature by as much as 4 degrees at important high latitudes of the NH. Issues of sea ice; poleward heat flux, for example. The discussion at page 9-32 does not really deal with the consequences of these errors. [Stephen E Schwartz, United States of America]	Taken into account. Exact wording in the Executive summary is being revised. The statement attempts to convey that fact that when models are provided with available estimates of external forcing, they respond in important aspects of their solutions in a way that is consistent with observed climate change. We have clarified the wording as Executive Summary has been revised, to reflect better the quantities for which tis is true. Detailed disagreement over recent 15 years is the subject of the new Box 9.2 in the Chapter.
9-156	9	3	45			"simulate realistically" -- suggest to clarify how you determine whether something is realistic or not the formulation is used rather often in the chapter when describing the quality of a model simulation. [comment applies to many more instances throughout the ES and the Chapter Sections] [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Executive summary has been thoroughly revised. "Correct" and "realistic" are no longer used in the ES.
9-157	9	3	46	3	46	Consider rephrase "...on continental and larger scales" , "larger scales" of what? [Jui-Lin (Frank) Li, United States of America]	Taken into account. Exact wording in the Executive summary has been revised.
9-158	9	3	46	3	48	Can we really have "very high confidence that models simulate realistically the global-scale surface temperature increase especially the last 50 years" while we have such limited knowledge of the actual size	Taken into account. Exact wording in the Executive summary is being revised. The statement attempts to

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						of the aerosol forcing over that period? Some models may, e.g., be getting trends that look reasonably correct but for a somewhat incorrect balance of forcings. I'd scale back to "high confidence". [Anthony Hirst, Australia]	convey that fact that when models are provided with available estimates of external forcing, they respond in important aspects of their solutions in a way that is consistent with observed climate change. We have clarified the wording as Executive Summary has been revised, to reflect better the quantities for which tis is true. Detailed disagreement over recent 15 years is the subject of the new Box 9.2 in the Chapter. Box 9.2 confirms that over the 1951-2012 period, both forcing and GMST trends are in excellent agreement between CMIP5 mean and Ch02/Ch08 best estimates. The "very high confidence" is hence justified.
9-159	9	3	46			I am worried by this general statement about the « confidence that models simulate realistically the surface temperature on continental scales ». I think the comment should be more precise on for instance the amplitude and phase of seasonal cycle to better than X% or large scale patterns, or the ocean/continent contrasts or so. Indeed, I am convinced that the agreement in the mean level of temperature is essentially a question of tuning. An error by 1W/m2 on the radiative fluxes changes typically the mean temperature by 1 K. And the simulation of clouds in global models is not robust enough to claim that we are able to simulate a priori the global Cloud Radiative Effect to better than probably 10 W/m2. We experienced changes of up to 8 W/m2 when just increasing the vertical resolution of our model (Hourdin et al., 2012, Clim. Dyn., doi: 10.1007/s00382-012-1411-3). It means for me that the agreement on the global mean temperature comes for a large part from the tuning of some important parameters. This tuning can be done either explicitly (as we do in our team) or implicitly through the long historical process of model version selection in each team. Hopefully, it seems that once this tuning of radiative forcing is done properly, the models simulate the global surface temperature reasonably well (to less than a few K probably). [Frédéric HOURDIN, France]	Taken into account. Exact wording in the Executive summary is being revised. The statement attempts to convey that fact that when models are provided with available estimates of external forcing, they respond in important aspects of their solutions in a way that is consistent with observed climate change. We have clarified the wording as Executive Summary has been revised, to reflect better the quantities for which tis is true. Detailed disagreement over recent 15 years is the subject of the new Box 9.2 in the Chapter.
9-160	9	3	47	3	47	The discrepancies in global mean temperature in recent years (Fig. 9.8) needs to be clearly explained, otherwise some readers will find it difficult to accept that there is very high confidence that models simulate the last fifty years. [Government of Australia]	Taken into account. Exact wording in the Executive summary is being revised. The statement attempts to convey that fact that when models are provided with available estimates of external forcing, they respond in important aspects of their solutions in a way that is consistent with observed climate change. We have clarified the wording as Executive Summary has been revised, to reflect better the quantities for which tis is true. Detailed disagreement over recent 15 years is the subject of the new Box 9.2 in the Chapter.
9-161	9	3	47	3	47	replace "models" by "large multi-model ensembles" [Michel Petit, France]	rejected. This is a generic statement.
9-162	9	3	48	3	48	Suggest omitting the sentence beginning "Together with the fact...". This statement is not relevant to a summary of model evaluation. [Ian Smith, Australia]	rejected. We feel it is important to remind readers that models are based on physical principles.
9-163	9	3	48	3	49	This needs more nuance - GCMs incorporate many physical principles but also many paramterisations. It is not sufficient to claim that predictions must therefore be good. Predictions must be assessed independently. [Gavin Schmidt, United States of America]	Taken into account. Exact wording in the Executive summary has been revised. This chapter does not deal with climate predictions; rather it deals with simulations of historical climate. While true that models include many parameterizations, these are generally based on some physical principles or understanding. The claim is not that physical basis is sufficient, but rather that it contributes to confidence. Many aspects of model performance are assessed in this chapter.
9-164	9	3	48			It would be useful to know the magnitude of aerosol forcing (direct and indirect) that results from the prescribed (or modeled) aerosol conc. [Henning Rodhe, Sweden]	noted. Aerosol processes and radiative effects are covered in Chapter 7.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-165	9	3	49	3	51	This seems too confident. What about regional details of responses? What about responses to aerosol forcing which have large uncertainty? Maybe qualify that you are referring to global means? Or tone down the language. [Government of United States of America]	Taken into account. Exact wording in the Executive summary has been revised. Different spatial scales are more explicitly assessed separately. Aerosol aspects are dealt with primarily in Chapter 7, and responses in Ch10..
9-166	9	3	49	3	51	Delete those 3 lines. This optimistic oversimplification is not consistent with the principal steps described in Box 9-1. [Michel Petit, France]	Taken into account. Exact wording in the Executive summary has been revised.
9-167	9	3	50	3	50	Given the substantial differences between model sensitivities (Table 9.4), it is unclear how this statement of 'very high confidence' that models respond to greenhouse gases 'correctly', can be made. Box 9.1 uses the word 'plausible' and 'significant warming'. Such qualifications should be included, especially if 'high' confidence is retained. [Government of Australia]	Taken into account. Executive summary has been thoroughly revised. "Correct" and "realistic" are no longer used in the ES.
9-168	9	3	50	3	50	Given the substantial differences between model sensitivities (Table 9.4), it is unclear how this statement of 'very high confidence' that models respond to greenhouse gases 'correctly', can be made. Box 9.1 uses the word 'plausible' and 'significant warming'. Such qualifications should be included, especially if 'high' confidence is retained. [Ian Watterson, Australia]	Taken into account. Executive summary has been thoroughly revised. "Correct" and "realistic" are no longer used in the ES.
9-169	9	3	50	3	51	"very high confidence" that the models "generally respond correctly" seems to mix high and low precision statements. [Peter Stott, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Exact wording in the Executive summary has been revised. The statement attempts to convey that fact that when models are provided with available estimates of external forcing, they respond in important aspects of their solutions in a way that is consistent with observed climate change. We have clarified the wording as Executive Summary has been revised, to reflect better the quantities for which it is true. Detailed disagreement over recent 15 years is the subject of a new box in the Chapter.
9-170	9	3	50	3	55	It is not possible for models to provide the 'correct' response. These claims need to be more qualitative. [Gavin Schmidt, United States of America]	Taken into account. Executive summary has been thoroughly revised. "Correct" and "realistic" are no longer used in the ES.
9-171	9	3	50			What does the statement « robust evidence that models perform less well on precipitation than on surface temperature » mean ? What metrics are you using to compare those two completely different variables ? By the way , I was impressed by the significant and very robust biases that persists in the surface temperatures in CMIP5 models, over continents during summer (systematic warm bias by a few K on average) or over the upwelling regions in the tropics (by 2-3 K on average on all the upwelling regions. Moreover, it is the temperature bias which explain a very large part of the rainfall biases on some regions (for instance for the West African monsoon, see eg Roehrig et al., in revision in climate dynamics, ftp://cnrm-ftp.meteo.fr/pub-moana/roehrig/CMIP5_WAM/rev1/Roehrig_etol2012_JClim_rev1.pdf) On large areas. [Frédéric HOURDIN, France]	Taken into account. Exact wording in the Executive summary has been revised. Statement is intended to convey the fact that agreement with observations is poorer for precipitation than for temperature when using the same sort of metric (e.g. Fig 9.6)
9-172	9	3	50			"correctly": I struggle with that word. Strictly when a response is correct it means it is 100% correct, true, accurate. But a model is always imperfect, so fundamentally the model response can never be fully correct. It may be correct in terms of the sign (i.e. warming as a response to increased CO2) but I suggest to either define "correct in terms of X, Y..." or reword this. One could go further and say the statement requires some observed or other evidence to test the model with, but since the correct response to future CO2 unknown and predicted by the model, that is impossible. I personally say the evidence of what models can do, the observations they reproduce, and the physical principles imply that they are adequate/useful for predicting climate change. Like in weather forecast, a model is never perfect, but it can be very useful for predictin certain things, but certainly not everything. [Reto Knutti, Switzerland]	Taken into account. Executive summary has been thoroughly revised. The adjective 'correct' has been replaced.
9-173	9	3	50			"generally respond correctly" meaninglessly vague. [Steven Sherwood, Australia]	Taken into account. Executive summary has been thoroughly revised. The adjective 'correct' has been replaced.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-174	9	3	51			It is suggested to substitute "greenhouse gases" by "greenhouse gas concentrations". [Klaus Radunsky, Austria]	Taken into account.Sentence has been removed during extensive ES revision.
9-175	9	3	53	3	56	It is not clear whether in this discussion of large scale patterns pertains to spatial patterns alone or spatial patterns and precipitation intensities too. [Government of United States of America]	Taken into account. Executive summary has been thoroughly revised. Change in precipitation intensities is no longer elevated to ES.
9-176	9	3	53	3	56	The current text strays from model evaluation into model projections. Suggested replacement text:"The simulation of large-scale patterns of temperature and precipitation has improved since the AR4. but precip[itation is less well simulated than is surface temperature." [Ian Smith, Australia]	Taken into account. Executive summary has been thoroughly revised. Change in precipitation intensities is no longer elevated to ES.
9-177	9	3	55	4	2	It is still debatable (IMHO) to render the 'medium confidence' assertion to the models ability to correctly simulate precipitation increases and decreases, if there is only limited support in observations.by observed trends. [Hans Visser, The Netherlands]	Taken into account. Executive summary has been thoroughly revised. Change in precipitation intensities is no longer elevated to ES.
9-178	9	3	55			"correctly": I struggle with that word. Strictly when a response is correct it means it is 100% correct, true, accurate. But a model is always imperfect, so fundamentally the model response can never be fully correct. It may be correct in terms of the sign (i.e. warming as a response to increased CO2) but I suggest to either define "correct in terms of X, Y..." or reword this. One could go further and say the statement requires some observed or other evidence to test the model with, but since the correct response to future CO2 unknown and predicted by the model, that is impossible. I personally say the evidence of what models can do, the observations they reproduce, and the physical principles imply that they are adequate/useful for predicting climate change. Like in weather forecast, a model is never perfect, but it can be very useful for predictin certain things, but certainly not everything. [Reto Knutti, Switzerland]	Taken into account. Executive summary has been thoroughly revised. The adjective 'correct' has been replaced.
9-179	9	3		4		I found the use of e.g. "evidence...agreement" statements far too liberal here. For example, you don't need to say that there is "strong evidence" that ENSO is simulated better by CMIP5 than CMIP3, it either is or isn't, you just have to look. Where hedging comes in is in making statements about the real world that cannot be proven. [Steven Sherwood, Australia]	Rejected. Here in the ES is the place where the calibrated language is warranted, except where we can make statements of fact. On ENSO, for example, we cannot; nor can we in many other places.
9-180	9	3		5		ES : As Chapter 9 assesses the models used for projections and predictions, a general comment is « How good is good enough ? ». It is a difficult question to answer, but a critical one. This issue is also what makes the evaluation of climate models fundamentally different from the evaluation of NWP models. It should be discussed somewhere in the chapter (even if it is to say that we don't know how to answer the question..). [Sandrine BONY, France]	Noted. We discuss this in sections 9.8.2 and 9.8.3
9-181	9	3		5		For the casual user it is not trivial to understand the policy relevance of many points listed in the Executive Summary. I get as a main message "Model approaches have greatly improved in time and are rock-solid now." I had expected some critical reflection in the Executive Summary on potential effects of fundamental gaps and uncertainties in mainstream model approaches. The last sentence of the executive summary reads like a formal disclaimer. In my view it should be expanded to reflect and assess the remaining uncertainties, gaps and deficiencies of state of the art models. It should be expanded into one or several paragraphs based on the discussions in paragraph 9.7 of the underlying chapter. [Jochen Harnisch, Germany]	Taken into account. Executive summary has been thoroughly revised. Model shortcomings are spelt out more explicitly.
9-182	9	3				Figure 9.8:. This figure is very important and it or some version of it is likely to become iconic of AR5. However it is misleading in that the quantity plotted is anomaly; that is the quantities being compared have different mean values subtracted from them, as indicated in the device at the top of the figure. To my thinking it is essential to plot temperature, not just temperature anomaly, and compare to measured GMST. The device at the top of the figure is a step in the right direction, but I respectfully suggest that the actual temperatures be plotted and compared with observations. It is a much stronger statement of the level of current modeling skill. See Tredger E (2009) On the evaluation of uncertainty in climate models. PhD thesis, London School of Economics, London http://cats.lse.ac.uk/homepages/edward/TREDGER_Thesis.pdf ; Figure 3.1 p. 71. Also Stevens B. and Schwartz S. E.: Observing and Modeling Earth's Energy Flows. Surveys Geophys. 33	Rejected. We show GMST for 1961-1990 reference period, to provide the information. For the global mean, the primary observational variable is the temporal anomaly and not the absolute value. Models' ability to simulate change is more imporant than simulating the absolute; moreover, as discussed in box 9.1, some models tune GMST, some others do not. Overall, Figure 9.8 does show the quantity most relevant more climate model evaluation. In addition, the absolute bias is shown in the inset at the top of the figure so as to clearly convey the biases.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						<p>779-816 (2012). DOI 10.1007/s10712-012-9184-0 Figure 11.</p> <p>Also Mauritsen, T., et al. (2012), Tuning the climate of a global model, J. Adv. Model. Earth Syst., 4, M00A01, doi:10.1029/2012MS000154, Figure 1.</p> <p>These figures show that the spread in GMST of AR4 models greatly exceeds the change in GMST over the twentieth century and indeed over expected temperature change in the 21st century, about 3 K This would have major effects on ice lines, vegetation, etc, and ultimately in climate response to forcing. A difference of 3 K is half way to an ice age. So it is misleading to present only temperature anomaly and not temperature itself. The departures of modeled temperature from observations and its implications must be shown and discussed.</p> <p>Note, for comparison, in Figure 9-24 the value in seeing actual sea ice extent, not anomaly. [Stephen E Schwartz, United States of America]</p>	
9-183	9	4	1	4	1	"this is been detected" should be either "this has been detected" or "this is being detected" [Sonya Legg, United States of America]	Editorial.
9-184	9	4	1	4	1	Change the last word "is" to "has". [Hai Lin, Canada]	Editorial.
9-185	9	4	1	4	2	Split the sentence into two sentences [Farahnaz Khosrawi, Sweden]	Editorial.
9-186	9	4	1	4	2	It is not clear what you want to say here. Does this mean that the models agree with each other, but not with measurements? [Farahnaz Khosrawi, Sweden]	Editorial.
9-187	9	4	2	4	2	Suggest omitting unnecessary references to Chapters 10, 12 and 14. [Ian Smith, Australia]	Accepted. Executive Summary has been thoroughly revised.
9-188	9	4	2	4	2	It feels odd to make the statement regarding observational confidence and then not cross-reference the observed chapter. This also applies generically to several of these bullets. The forward looking modelling chapters are referenced but the observational chapters are not in several cases. Yet many times the confidence in the observations, which must stem from the observational chapters (and should be checked for consistency) is alluded to. [Peter Thorne, United States of America]	Accepted. Executive Summary has been thoroughly revised.
9-189	9	4	4	4	8	ES : « to simulate surface temperature and precipitation » : it is a bit vague. The statement should be made more specific (patterns ?, variability ?, trends ? extremes?). « not to be simulated equally well, and... » : an alternative would be : « has not improved as much, but.. » [Sandrine BONY, France]	Taken into account. ES has been thoroughly revised.
9-190	9	4	4	4	8	The current text refers to "confidence" that model skill has improved. This is meaningless, since the aim of model evaluation is to demonstrate whether improvement has occurred or not. Suggest omitting this text as it repeats the previous point [Ian Smith, Australia]	Rejected. This paragraph speaks about regional scales.
9-191	9	4	4			"confidence is less than for larger scales.... high confidence simulated ... with some improvement." this is a classic example of a statement which would be much more informative and useful if clear guidance on the spatial scales where CMIP5 models are not realistic in simulating the past, and likely (or virtually certain) to be misinformative in projecting the future. As currently phrased, it is unhelpful to know models are less able to simulate at smaller scales than larger, and potentially misleading to say they have improved from CMIP3 if both are believed to be unrealistic. This is discussed in LA Smith (2012) Predictability and Insight, Int J Forecasting (in review). [Leonard Smith, United Kingdom]	Taken into account. ES has been thoroughly revised. We cannot cite a paper "in review".
9-192	9	4	5	4	7	"Nevertheless since AR4' --> I consider this sentence as being rather strange and uninformative: why is it worthwhile to mention that there is now SOME improvement since AR4. The adjective 'some' suggests only minor improvement. Why adhere the judgement 'there is high confidence' to such a statement, instead of just saying that there are some improvements since AR4. [Hans Visser, The Netherlands]	Taken into account. ES has been thoroughly revised.
9-193	9	4	6	4	6	High confidence of improvement is not very useful; what is the confidence that it is simulated well (for some definition of "well")? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. ES has been thoroughly revised. On this specific point, it is one of the primary questions to this chapter whether models have improved (see FAQ), and one of the primary

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							quandaries that it is hard to define a standard for goodness for purpose. Hence, statements of improvement are important. But it has been sharpened considerably.
9-194	9	4	6			What is the related metric? [Peter Braesicke, United Kingdom]	Taken into account. ES has been thoroughly revised.
9-195	9	4	6			"with some improvement" -- this remains rather vague if not supported by details about what has been improved (and what not) [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. ES has been thoroughly revised.
9-196	9	4	7	4	8	This sentence is not clear. What does "...not to be simulated equally well.." mean? All regions are not simulated with the same skill? [Government of Australia]	Taken into account. ES has been thoroughly revised.
9-197	9	4	7	4	8	Unclear sentence: 'not to be simulated equally well' --> equally well as what?; 'the assessment remains difficult' --> assessment of improvement in the simulation of precipitation since AR4?? [Hans Visser, The Netherlands]	Taken into account. ES has been thoroughly revised.
9-198	9	4	9	4	11	It is not convincingly shown in the chapter that the CMIP5 models realistically simulate the annual cycle of Arctic sea-ice extent. [Gunilla Svensson, Sweden]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-199	9	4	9	4	12	It is shown in the document that many models can replicate the observed annual cycle in the extent and the observed trend over the last decades. How this is possible with almost all problems identified in the CMIP3 models regarding storm tracks, Arctic ocean circulation, Arctic clouds remain as described in the chapter? Almost no sea-ice model updates have been made since CMIP3 (except for some models providing more advanced albedo schemes)? Did the increased horizontal resolution magically solve everything? I would argue very strongly that it is very likely that the models get the right answer for the wrong reason and the present result is because of focus on this while tuning the model. This strong statement might give the impression to society and policy makers that the problem is solved, it is not! [Gunilla Svensson, Sweden]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-200	9	4	10	4	10	Whether or not this is impressive depends entirely on whether the annual cycle is tuned for. [Gavin Schmidt, United States of America]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-201	9	4	10	4	12	These statements on the CMIP5 model performance for the Arctic sea ice extent sounds very positive, and seem to suggest that no essential model improvements would be needed. How does this relate to the statements on page 9-40, line 1-10 which state that various processes and feedbacks are still poorly described in the models. [Hans Visser, The Netherlands]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-202	9	4	10	4	14	I also do not understand how you can have high confidence in models' ability to simulate Arctic sea ice trends. If one looks at the ensemble of one model's simulations of recent and future decades one finds a pretty wide spread in Arctic sea ice decline rates. If the observed decline happens not to lie outside the full range of ensemble members is that enough to give us confidence that we have it right? Have the sea ice energy budgets of the models been validated against observations, and do we know whether the models agree that the decline is largely due to local ocean warming vs. latitudinal heat ocean transport vs. latitudinal atmospheric heat transport vs. cloud effects on downwelling longwave radiation vs. changes in sea ice radiative or thermodynamic properties? Or do they get a similar answer for different reasons? Do we simulate mixed-phase Arctic clouds and their response to changing concentrations of ice nuclei well enough to be confident of the clouds' effect on sea ice decline? Does the less impressive simulation of the Antarctic trend make you more circumspect? [Anthony Del Genio, United States of America]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-203	9	4	10	4	14	Thickness of the sea ice is an important parameter that can affect heat budget at the polar regions. Therefore, model capability should also be evaluated for the thickness of the sea-ice and the related changes. [Government of Canada]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-204	9	4	10	4	14	Suggested text: "CMIP5 models realistically simulate the annual cycle of Arctic sea-ice extent, including the trend in Arctic sea-ice extent over the past several decades. There is a clear improvement over CMIP3 models." [Ian Smith, Australia]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-205	9	4	10	4	14	Please indicate the skill at simulating the annual cycle of Antarctic sea-ice. [Ian Smith, Australia]	Accepted.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-206	9	4	10	4	14	Suggested text:" Most models simulate a decrease in Antarctic sea-ice extent over the past several decades which is not seen in the observations." [Ian Smith, Australia]	Taken into account. Sentence rephrased.
9-207	9	4	10	4	14	I would prefer to break up this statement into two: one (line 10-12) which comments on the performance of the CMIP5 models for the Arctic, and another (line 12-14) which comments on their performance for the Antarctic, which in fact is considered as not satisfactorily since the trends in model results and observations are different for this region. [Hans Visser, The Netherlands]	Accepted. Statement revised.
9-208	9	4	11	4	11	I think this is overstating the consensus about the trends and whether models capture them properly. I think there is still an ongoing debate whether the discrepancy between the multi-model ensemble and the observations points to a lack of sea ice sensitivity of the models or can be explained by natural variability (e.g Winton 2011). While some models may produce realistic trends, others clearly don't. I don't think the constraint business isn't sufficiently settled to make that statement. [Axel Schweiger, United States of America]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-209	9	4	11	4	12	This statement is too strong. The magnitude of the CMIP5 multi-model mean trend in summer Arctic sea ice extent is close to the observed one up to 2005, but is significantly underestimated thereafter. [Thierry Fichefet, Belgium]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-210	9	4	11	4	12	I don't see a justification for "high confidence that they realistically simulate the trend in Arctic sea ice extent over the past decades". Instead I see in Fig. 9.24 a strong indication that CMIP5 models are still under-predicting the recent trend. Even if some models are simulating a trend as large as observed are they doing it for the right reason? For example are they warming too much globally in recent decades and that's why they are simulating a large enough trend in sea ice? In the text where this is discussed (Sect 9.4.3) the authors state that there is high confidence that the CMIP5 models capture the first-order behavior of the Arctic sea ice...particularly the seasonality and the trend... This means to me that the models have the general right idea about the trend (i.e., the sign), but not necessarily that they have the magnitude right. [Thomas Knutson, United States of America]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-211	9	4	11	4	12	It is not convincingly shown in the chapter that the CMIP5 models realistically simulate the trend in Arctic sea-ice extent [Gunilla Svensson, Sweden]	Taken into account. Assessment on annual cycle and trend of Arctic sea ice has been re-phrased.
9-212	9	4	12	4	14	Decreasing trend in Antarctic sea ice but increasing trend in the observations. Isn't that contradicting? [Farahnaz Khosrawi, Sweden]	Taken into account. Sentence rephrased.
9-213	9	4	13	4	13	larger inter-model spread [European Union]	Editorial.
9-214	9	4	13	4	13	It is unclear what is meant by 'intermodal spread'. [Hans Visser, The Netherlands]	Editorial.
9-215	9	4	13	4	13	"...larger intermodel spread" - larger than what? [duane waliser, United States of America]	Editorial.
9-216	9	4	13			Be more specific regarding the period [Peter Braesicke, United Kingdom]	Accepted. Period now given.
9-217	9	4	13			This sentence would be much more useful if it indicated whether the observed trend is within the model spread or not. [Steven Sherwood, Australia]	Taken into account. Within space constraints, reformulation makes contrast clearer.
9-218	9	4	16	4	19	Suggested text: " Many models realistically simulate the observed recent trend in ocean heat content. This adds confidence to their representations of the global energy budget." [Ian Smith, Australia]	Accepted, thank you
9-219	9	4	16	4	19	this seems a rather weak statement given the conclusion "many models simulate realistically" is immediately questioned by "some models show substantial deviations from the observations". And then it's not clear how these substantial deviations between observations and models could "give confidence in using climate models to assess the global energy budget"? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account, statement modified.
9-220	9	4	17	4	17	"substantial deviations from the observations" - please be specific - do you mean in the "mean" the "variability" ? [duane waliser, United States of America]	Taken into account, statement modified.
9-221	9	4	21	4	23	Suggested text: "Simulations of the El Nino Southern Oscillation (ENSO) has improved from CMIP3 to CMIP5. Several models now realistically simulate the frequency spectrum and the amplitude of sea surface temperatures." [Ian Smith, Australia]	Taken into account, statement modified.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-222	9	4	21	4	30	Chapter 9: Executive Summary. The document gives many details concerning the assessment of the simulation of interannual to centennial variability (Section 9.5.3). However, the executive summary of chapter 9 only devotes a few lines to it. It would be desirable that both parts will be more balanced. [Government of Spain]	Taken into account. The ES has been extensively re-written.
9-223	9	4	21	4	30	Please include a discussion about the veracity of teleconnections when the modes of variability are discussed. [Government of United States of America]	Taken into account. The ES has been extensively re-written, including assessment of modes of variability. "Teleconnections" on its own is, however, too unspecific.
9-224	9	4	21			We recommend deletion of the word "robust". [Government of United States of America]	Taken into account. Assessment of improvement of ENSO simulation is more explicitly backed up in ES.
9-225	9	4	24	4	24	"However, this improvement is not universal across the CMIP5 ensemble". Isn't that just a repetition of what you said in the sentences before and thus could be skipped? [Farahnaz Khosrawi, Sweden]	Taken into account. The ES has been extensively re-written.
9-226	9	4	24	4	27	... and the Quasi-Biennial Oscillation...: This is true for a relatively small number of models. Most models still do not simulate a QBO at all because they are not configured to simulate the QBO (in terms of vertical resolution), but those models which do simulate the QBO obtain it because they simulate the essential mechanisms (wave-meanflow interaction due to resolved and unresolved waves). The NAO is however hard to miss and represented in probably all AOGCMs. The sentence, however, suggests that both modes of variability are about equally well represented. [European Union]	Taken into account, statement modified.
9-227	9	4	24	4	27	It is strange that - on the one hand - it is stated that other important modes are simulated realistically, although the confidence in this assessment (?do you mean this statement, namely 'other important modes are simulated realistically') is only medium. [Hans Visser, The Netherlands]	Taken into account. The ES has been extensively re-written.
9-228	9	4	25	4	25	The QBO is not a common, and certainly not universal, feature of the CMIP5 models. [Gavin Schmidt, United States of America]	Taken into account, statement modified.
9-229	9	4	26	4	26	Omit the text : "...although the confidence in this assessment is only medium". As mentioned above, such a statement is meaningless. [Ian Smith, Australia]	Taken into account, statement modified.
9-230	9	4	27	4	27	Consider move "CMIP5 models are able to realistically simulate northern-hemisphere surface temperature variability on timescales of decades to centuries (Figure 9.33)." at line 21. [Jui-Lin (Frank) Li, United States of America]	Taken into account. ES has been extensively re-written.
9-231	9	4	27	4	28	Please also comment on the Southern Hemisphere. [Government of Germany]	Rejected. No data exist to allow for reconstructions in the Southern Hemisphere.
9-232	9	4	27	4	28	We feel this statement is too confident. There are only approximations available from paleo proxies to compare with. And the observed record is quite short for making such a claim. Also models differ from one another a lot on things like Atlantic multidecadal variability (i.e., the AMO...or perhaps the aerosol forced version of "AMO"). So leading off with "CMIP5 models are able to realistically simulate..." seems too confident and in need of tone down. [Government of United States of America]	Taken into account. Paragraph has been re-written to clarify which element of variability is meant here (spectrum of NH surface temperature).
9-233	9	4	27	4	28	This statement is not consistent with the fact that models show a huge range in the simulated amplitude of surface temperature variability (e.g. Hawkins & Sutton, GRL, 2012 doi: 10.1029/2011GL050087). [Rowan Sutton, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Paragraph has been re-written to clarify which element of variability is meant here (spectrum of NH surface temperature).
9-234	9	4	28	4	30	The sentence could be omitted since it is not clear what signal and noise you are referring to as well what is meant with attribution and detection studies (though this becomes clearer with continuing reading the chapter). Alternatively, to omitting you could add an extra item and be more precise. [Farahnaz Khosrawi, Sweden]	Rejected. Signal and noise are well understood in their meanings here.
9-235	9	4	28	4	30	Omit the last sentence- it is meaningless in an Executive Summary on model evaluation. [Ian Smith, Australia]	Rejected. Implications for D&A are important.
9-236	9	4	28			"realistically simulation variability on time scales from decades to centuries" - From the century time scale, the statement is overdone. Observations are the issue. [Ronald Stouffer, United States of America]	Rejected. This statement is based on published analysis of paleo-climate reconstructions and simulations..

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-237	9	4	32	4	32	Do you really mean progress in the 'assessment' of model simulations, or rather progress in the 'performance' of model simulations? [Hans Visser, The Netherlands]	Taken into account. Executive Summary text has been thoroughly revised.
9-238	9	4	32	4	34	suggest to add fig.9.37 in page 206 and show the number of models in the figure. [Lei Huang, China]	noted. This comment refers to the figure, not the Executive Summary. Comment is unclear, but the number of models is evident from the axis labels on the upper panel of the figure.
9-239	9	4	32	4	40	The ability of models to simulate tropical cyclones is a problem of great interest to the modeling community and AR readers alike. There should be more discussion in the Executive Summary of modeling capability WRT Tropical Cyclones than the limited statement that "Some high-resolution atmospheric models have realistically simulated tracks and counts of tropical cyclones." Examinations of CMIP5 models, such as Suzana Camargo's paper titled "Global and regional aspects of tropical cyclone activity in the CMIP5 models" concur with the existing Ch. 9 ES statement, but also indicates that significant biases remain when comparing global models with observations. The conclusion is that models cannot currently provide skill in future projections of tropical cyclone activity, but that projections of environmental conditions influencing tropical cyclone activity may be more fruitful. [Government of United States of America]	Taken into account. Executive Summary text has been thoroughly revised. Statement regarding Tropical Cyclones has been modified based on updated information.
9-240	9	4	33	4	33	I would suggest to write here "extreme weather events" instead of just "extreme events" [Farahnaz Khosrawi, Sweden]	Editorial.
9-241	9	4	33	4	33	Suggested text: "...extreme events. The global distribution..." [Ian Smith, Australia]	Editorial.
9-242	9	4	33	4	34	medium evidence' and 'high agreement' appear as strange twins here. Is this referring to the situations where the model-community strongly believes that the global distribution of temperature represented by the models is ok, but that there is not much evidence available to test this? Or is there also a considerable amount of evidence that does not confirm the model results. In the latter case I would definitely not agree that the models represent the global distribution of the temperature extremes well. In the former case I would rather prefer to reformulate the statement, e.g. as 'There is high agreement that well by models. This is confirmed by the currently available evidence.' [Hans Visser, The Netherlands]	Taken into account. Executive Summary text has been thoroughly revised.
9-243	9	4	33			"medium evidence, high agreement" -- could this be expressed as a confidence statement? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Executive Summary text has been thoroughly revised.
9-244	9	4	34	4	34	Please specify what sort of "temperature extremes" (e.g. monthly, daily, etc) [Rowan Sutton, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Executive Summary text has been thoroughly revised. Statement for Executive Summary was intended to be generic, but has been phrased more precisely.
9-245	9	4	35	4	39	Suggested text: "...20th century is also well captured but models tend to overestimate the observed increase of warm temperature extremes and underestimate the increase of cold temperature extremes. CMIP5 models tend to simulate...extremes compared to CMIP3 models." [Ian Smith, Australia]	Editorial. Thank you.
9-246	9	4	36	4	37	Is it not better to replace 'warming of warm temperature extremes' by 'increase of warm temperature extremes', and 'warming of cold temperature extremes' by 'decrease of cold temperature extremes'? [Hans Visser, The Netherlands]	Editorial. Wording chosen to avoid misunderstanding as to the 'sign' of the change.
9-247	9	4	37	4	38	The process by which the real atmosphere produces extreme precipitation is usually through organized mesoscale convective systems that produce rain over a widespread area for a long time, aided by synoptic weather patterns that sometimes stall rather than propagating. Mesoscale organization of convection is not simulated at all by any climate model except GFDL CM3. Does this not matter, and are Clausius-Clapeyron and the moist adiabatic lapse rate all we need to know to get extreme precip right? Has anyone examined the synoptic conditions in which extreme precip occurs in the real world and in models and concluded that the models are getting the right answer for the right reason? These types of conclusions are symptomatic of the chapter's emphasis on results as opposed to the process by which the models get those results. [Anthony Del Genio, United States of America]	Taken into account. The Executive Summary provides high-level conclusions. Where possible, based on available literature, model results and the processes that underly them are assessed, but this is of course limited by available space.
9-248	9	4	38	4	38	for me it is not obvious that more intense precipitation extremes are THUS more realistic. Remove the word	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						"thus" [Bart Van den Hurk, Netherlands]	
9-249	9	4	39	4	30	Some ... cyclones' -->Can you give more information on this subset of models? What makes that their simulation shows more realism as compared to other models? I.e.what can the model-community learn fromthis? [Hans Visser, The Netherlands]	Taken into account. Executive Summary is necessarily brief. Important aspect is noted at the beginning of the sentence: high-resolution.
9-250	9	4	39	4	40	The sentence is here somewhat lost. The connection to the rest of this item becomes not clear. [Farahnaz Khosrawi, Sweden]	Taken into account. Executive Summary text has been thoroughly revised.
9-251	9	4	40	4	40	Write "storm tracks" instead of just "tracks" [Farahnaz Khosrawi, Sweden]	rejected. 'Tracks' refers to Tropical Cyclones -- the subject of this sentence.
9-252	9	4	42	4	43	If the trend is prescribed, is not it natural to expect that it is well represented then? Check the content and phrasing. [Government of France]	Taken into account. The point is that time-evolving stratospheric ozone is now included. Text revised slightly to make this more clear.
9-253	9	4	42	4	43	"the trends inis well represented" - should be "trend" or "is" should be "are". [Sonya Legg, United States of America]	Editorial.
9-254	9	4	42	4	43	what does "well represented" mean, and what does it mean for a prescribed trend to the well represented? [Rowan Sutton, United Kingdom of Great Britain & Northern Ireland]	Taken into account. The point is that time-evolving stratospheric ozone is now included. Text revised slightly to make this more clear.
9-255	9	4	42	4	43	whether prescribed or calculated interactively' --> this information is not very informative on this summary level [Hans Visser, The Netherlands]	Taken into account. The point is that time-evolving stratospheric ozone is now included. Text revised slightly to make this more clear.
9-256	9	4	42	4	45	I don't fully agree; there is high confidence that ozone matters, but figure 9-10 does not shout 'well represented' at me, please rephrase [Peter Braesicke, United Kingdom]	Agreed. Has been changed to medium agreement and robust evidence for improvements.
9-257	9	4	42	4	45	The assertion here and elsewhere (p. 9-31 lines 59-60+; Section 9.4.1.3.5) that stratospheric O3 trends are well represented in CMIP5 models may contradict Santer et al. (2012). A preprint of that paper says models including O3 chemistry "have errors in their simulations of historical ozone changes" (citing Eyring et al. 2012). Santer et al. go on to select a "more reliable" subset of CMIP5 models that prescribe rather than compute O3. Figure 9.10 shows total column O3 whereas Santer et al. may be talking about the vertical distribution of O3, so perhaps there's no real contradiction. In any case, checking the final published version of Santer et al. seems warranted. [Curt Covey, United States of America]	Agreed. Has been changed to medium agreement and robust evidence for improvements.
9-258	9	4	42	4	45	suggest to show relevant section or figure. [Lei Huang, China]	Editorial
9-259	9	4	42	4	45	Suggested text:"Trends in stratospheric ozone, whether prescribed or calculated interactively, are well represented in CMIP5 models and constitutes a significant improvement over CMIP3 models.The representation of associated impacts on high latitude climate has also improved." [Ian Smith, Australia]	Thanks for the suggestion. The statement has been reworded, also in responses to other comments, see above.
9-260	9	4	42	4	49	The executive summary is a bit short on ESM aspects. [Olivier Boucher, France]	Taken into account. Executive Summary text has been thoroughly revised and more on ESM aspects included.
9-261	9	4	43			trends are, not trends is [Julia Hargreaves, Japan]	Editorial
9-262	9	4	43			The following wording is suggested: ... , are well represented. [Klaus Radunsky, Austria]	Editorial
9-263	9	4	44	4	45	Please give an example for these impacts or refer to the respective section. Are there any known impacts in other areas? [Government of Germany]	Taken into account. Executive Summary text has been revised and word 'impacts' replaced to avoid confusion.
9-264	9	4	44	4	45	Split the sentence into two sentences. End the first sentence after AR4 and continue then with "There is high confidence....". [Farahnaz Khosrawi, Sweden]	Editorial
9-265	9	4	47	4	47	It is not a good idea to equate the term "Earth System Model" to the inclusion of a carbon cycle. That would be	rejected. This sentence does 'equate' carbon cycle to

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						a carbon cycle model. The Earth System is much wider. [Gavin Schmidt, United States of America]	ESM. Indeed ESM is defined (as per general usage in the climate modelling community) in section 9.1.2
9-266	9	4	47	4	49	you might consider referring here to Ch6, where ESMs and their representation of the C-Cycle are assessed [Thomas Stocker/ WGI TSU, Switzerland]	Accepted.
9-267	9	4	51	4	51	delete 'of course' [Gavin Schmidt, United States of America]	editorial
9-268	9	4	51	4	51	Use of qualifier ', of course,' here seems colloquial and also assumes knowledge of the reader that I doubt can be assumed. It would be better to rewrite explicitly as to why this is the case I think. [Peter Thorne, United States of America]	Taken into account
9-269	9	4	51	4	57	Some mention should be made of the difficulty models have in simulating changes in baroclinic instability as this affects their ability to produce the changes in the growth and preferred location of the storm tracks. Suggest adding text in quotes to this paragraph: There are, of course, many areas of model performance that remain to be improved. There is large inter-model spread, and systematic biases are evident in a number of important aspects. For example, models have problems simulating the mean temperature structure of the Tropical Atlantic Ocean (Section 9.4.2.5.2), the diurnal cycle of precipitation (Section 9.5.2.2), the Madden-Julian Oscillation (Section 9.5.2.2), and clouds and cloud radiative effects (Section 9.4.1.1.2). "Models also have difficulty in simulating the changes and trends in atmospheric baroclinic instability and hence the changes in the growth rate and preferred location of mid-latitude storm formation." In some cases, model results are in general agreement with observations, but the observational uncertainty precludes definitive statements about model quality. [Jorgen Frederiksen, Australia]	Taken into account. Executive Summary text has been thoroughly revised; however, some judgement has to be used to select items that are important, relevant and comprehensible to readers of the Executive summary and so it cannot be an exhaustive list.
9-270	9	4	51	4	57	Some mention should be made of the difficulty models have in simulating changes in baroclinic instability as this affects their ability to produce the changes in the growth and preferred location of the storm tracks. Suggest adding: Models also have difficulty in simulating the changes and trends in atmospheric baroclinic instability and hence the changes in the growth rate and preferred location of mid-latitude storm formation. In some cases, model results are in general agreement with observations, but the observational uncertainty precludes definitive statements about model quality. [Government of Australia]	Taken into account. Executive Summary text has been thoroughly revised; however, some judgement has to be used to select items that are important, relevant and comprehensible to readers of the Executive summary and so it cannot be an exhaustive list.
9-271	9	4	51	4	57	Please consider mentioning or adding a bullet about the largest model spread in simulating the ice cloud and water vapor in the upper troposphere. This information is not available in AR4 since the CMIP3 output did not provide vertically resolved cloud water contents, and A-Train data have not been available yet. Thus this is a new and important findings for AR5 (Jiang et al. 2012). Additional note: This new finding, resulted from a combined water vapor and ice cloud observations from A-Train satellite instruments, is very important to future improvement of climate models. In the upper troposphere, cloud and water vapor are strongly coupled (e.g., Su et al., 2006a; Jiang et al. 2010), where water vapor exerts the strongest greenhouse gas effect (Soden et al., 2005). Importance of the upper tropospheric water vapor radiative feedback has been emphasized by many studies (e.g. Held and Soden, 2000; Allan et al., 2003; Colman, 2003; Soden and Held, 2006; Bony et al., 2006; Soden et al., 2008; Gettelman and Fu, 2008; Dessler and Sherwood, 2009; Sherwood et al., 2010). In addition, the water vapor concentration at the tropopause is directly linked to the stratospheric water vapor amount, which can affect the surface energy budget and stratospheric ozone chemistry (Solomon et al., 2010). In light of these findings, there is a need to understand the processes that cause the large spread in the modeled upper tropospheric clouds and water vapor in order to reduce the uncertainties for climate projections. References: (in addition to some references already in the Chapter 9 reference list) Soden, B. J., D. L. Jackson, V. Ramaswamy, M. D. Schwarzkopf, and X. Huang, The radiative signature of upper tropospheric moistening, Science, 310(5749), 841844, doi:10.1126/science.1115602, 2005. Held, I. H., and B. J. Soden, Water vapor feedback and global warming, Annu. Rev. Energy Environ., 25, 441475, doi:10.1146/annurev.energy.25.1.441, 2000. Solomon, S., K. Rosenlof, R. Portmann, J. Daniel, S. Davis, T. Sanford, and G. Plattner, Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, Science, 303, 2010. Su, H., W.G. Read, J.H. Jiang, J.W. Waters, D.L. Wu, and E.J. Fetzer,	Taken into account. Executive Summary text has been thoroughly revised, although it cannot provide an exhaustive list of all findings discussed in the body of the chapter.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						"Enhanced positive water vapor feedback associated with tropical deep convection: New evidence from Aura MLS," Geophys. Res. Lett. 33, L05709, doi:10.1029/2005GL025505, 2006a. Jiang, J.H. et al. "Five-year (2004-2009) Observations of Upper Tropospheric Water Vapor and Cloud Ice from MLS and Comparisons with GEOS-5 analyses," J. Geophys. Res. 115, D15103, doi:10.1029/2009JD013256, 2010. [Government of United States of America]	
9-272	9	4	51	4	57	Stratospheric variability is another aspect that needs to be improved. Given the tendency to raise the top of climate models, I suggest to point out that we need to advance on this aspect of modelling (especially on the factors controlling stratospheric variability). [Elisa Manzini, Germany]	rejected. The list is intended to be 'examples', not exhaustive. The IPCC Assessment does not make recommendations about research directions.
9-273	9	4	51	4	57	Suggested text:"Many areas of model performance remain to be improved. Models have problems simulating the mean temperature of the Tropical Atlantic Ocean....model quality." [Ian Smith, Australia]	Editorial
9-274	9	4	51	4	57	"There are of course ...". This is terrible! Climate sensitivities still differ by a factor of two, but they are not mentioned, and after twenty years (see my comments on chapter TS, Page 38) clouds are still a problem. And this is "of course" slipped in right at the end. [David Webb, United Kingdom]	Taken into account. Executive Summary text has been thoroughly revised.
9-275	9	4	51		57	Some mention should be made of the difficulty models have in simulating changes in baroclinic instability as this affects their ability to produce the changes in the growth and preferred location of the storm tracks. Suggest rewriting as follows: "There are, of course, many areas of model performance that remain to be improved. There is large inter-model spread, and systematic biases are evident in a number of important aspects. For example, models have problems simulating the mean temperature structure of the Tropical Atlantic Ocean (Section 9.4.2.5.2), the diurnal cycle of precipitation (Section 9.5.2.2), the Madden-Julian Oscillation (Section 9.5.2.2), and clouds and cloud radiative effects (Section 9.4.1.1.2). Models also have difficulty in simulating the changes and trends in atmospheric baroclinic instability and hence the changes in the growth rate and preferred location of mid-latitude storm formation. In some cases, model results are in general agreement with observations, but the observational uncertainty precludes definitive statements about model quality." [Carsten Frederiksen, Australia]	Taken into account. Executive Summary text has been thoroughly revised; however, some judgement has to be used to select items that are important, relevant and comprehensible to readers of the Executive summary and so it cannot be an exhaustive list.
9-276	9	4	53			"temperature structure" could be clarified - I assumed you meant the 3D structure, but referring to section 9.4.2.5.2 it seems to be referring to the 2D spatial pattern [Julia Hargreaves, Japan]	Taken into account. Executive Summary text has been thoroughly revised.
9-277	9	4	53			For me, the bias of the tropical SSTs is not restricted to the tropical Atlantic ocean. A very similar, and robust as well, bias is present on the upwelling regions in general and in particular on the Pacific oceans. The bias is probably more important for the Atlantic because the basin is smaller. The consistency of this bias among the CMIP5 models and between CMIP3 and CMIP5 is a very impressive feature (Roehrig et al., 2012, ftp://cnrm-ftp.meteo.fr/pub-moana/roehrig/CMIP5_WAM/rev1/Roehrig_et_al2012_JClim_rev1.pdf), which could be emphasized as a major failure of our community (no improvement at all between AR4 and AR5). [Frédéric HOURDIN, France]	Taken into account. Executive Summary text has been thoroughly revised.
9-278	9	4	55			We suggest that the authors consider adding that upper tropospheric clouds and water vapor amounts have large inter-model spreads (e.g., Jiang et al. 2012). [Government of United States of America]	Taken into account. Executive Summary text has been thoroughly revised; however, some judgement has to be used to select items that are important, relevant and comprehensible to readers of the Executive summary and so it cannot be an exhaustive list.
9-279	9	4				Consider adding a bullet about improvements in ice cloud simulation from AR4 to AR5. This is one of the important areas of progress from AR4 to AR5 (e.g. Jiang et al. 2012, Figure 5), also see following references: 1. Balcerak, E., "Comparison with observations shows cloud simulations improving", RESEARCH SPOTLIGHT, Eos, Volume 93, Number 38, 18 September 2012. (http://www.agu.org/journals/eo/v093/i038/2012EO38_tabloid.pdf) 2. Benka, S.G., "Cloud simulations improving in climate models", Physics Today, Volume 65, Issue 10, October 2012.	Taken into account. Executive Summary text has been thoroughly revised; however, some judgement has to be used to select items that are important, relevant and comprehensible to readers of the Executive summary and so it cannot be an exhaustive list.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						(http://www.physicstoday.org/daily_edition/physics_update/cloud_simulations_improving_in_climate_models) Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. (Jiang et al. Figure 5; also see the AGU highlight and Physics Today article about improvement in cloud simulations). [Government of United States of America]	
9-280	9	5	1	5	1	I think there is a need for a bullet describing the main outcomes of the evaluation using past climates such as the last glacial maximum and the mid Holocene. The description in part 9.4.1.4 shows an overall capacity of models to reproduce large scale features (e.g. the 1.5 land/ocean ratio) but also some underestimations which are important to be quoted, such as an understimation of the polar amplification also quoted in chapter 5. [SYLVIE JOUSSAUME, France]	Taken into account. Executive Summary text has been thoroughly revised. Palaeo-modelling has been included in the ES where it strengthens the assessment of model capability.
9-281	9	5	4	5	4	What do you mean with equilibrium climate sensitivity? [Farahnaz Khosrawi, Sweden]	noted. 'Equilibrium climate sensitivity' is defined in the glossary.
9-282	9	5	4	5	5	It seems to me it is time to start saying this a bit more strongly to decision makers and the public, namely that, over the last 30 years, with climate models going from quite simple to much more sophisticated, the climate sensitivity has really not changed very much at all. While we understand that exactly how clouds change, which is somewhat uncertain, can affect the climate sensitivity a bit, that the climate sensitivity has stayed about the same as we have added a deep ocean, land surface processes, gone to a finer grid, etc., suggests, at least to me, that the climate sensitivity must be determined by very basic aspects of the Earth system and the governing laws, so, namely, the seasonal cycle of solar energy, the Stefan-Boltzmann equation, the Clasuis-Clapeyron equation, the general structure of the atmosphere (so the perfect gas law), land-sea geography, etc. and other relationships that we have had in the models from the start. That the climate sensitivity is in the range we have is also suggested by the paleoclimatic record also adds to confidence that we have the range about right. Thus, we need to say that, at least for processes and the system on time scales of centuries, there is less and less likelihood that the sensitivity we have to deal with is going to go down or up much at all--and that there is just no indication that the skeptics with their claim of a very low sensitivity will be proven correct. [Michael MacCracken, United States of America]	Taken into account. This is, however, primarily an issue for the synthesis of ECS discussion in Chapter 12.
9-283	9	5	4	5	7	And the unstated conclusion is that the pages and pages of model evaluation that you are about to show us actually don't tell us which models are better or worse for their intended purpose. [Anthony Del Genio, United States of America]	noted. This is an assessment of published results and reflects the state of the science at the time of writing.
9-284	9	5	4	5	13	The continuing wide spread of climate sensitivities is 'disappointing'. This needs to be stated here and elsewhere in the document in order to help direct funding into needed areas of research. [David Webb, United Kingdom]	noted. While it may be 'disappointing', it is the state of the science assessed at this time. The IPCC report is not a forum to lobby for funding.
9-285	9	5	5	5	7	ES : The second part of this sentence is weak and potentially misleading. Why focusing on the influence of time-mean errors for assessing our confidence in climate sensitivity ? Either other aspects of model evaluation (more relevant for evaluating our confidence in climate sensitivity) should be mentioned, or this statement should be removed. [Sandrine BONY, France]	Taken into account. Text revised. The statement was intended to address concerns, raised by other reviewers of previous drafts, that biases in the mean temperature are critical.
9-286	9	5	6	5	6	which enhances confidence' --> I do not see immediately why this would enhance the confidence in model simulations (haven't checked the argumentation in section 9.7.2.1) [Hans Visser, The Netherlands]	Taken into account. Text revised. The statement was intended to address concerns, raised by other reviewers of previous drafts, that biases in the mean temperature are critical.
9-287	9	5	6	5	7	Suggested text: ".climate sensitivity." (omit reminder of sentence) [Ian Smith, Australia]	Editorial. Text revised. The statement was intended to address concerns, raised by other reviewers of previous drafts, that biases in the mean temperature are critical.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-288	9	5	7	5	7	Please clarify what is meant by time-mean. Suggestion: "simulated absolute values of the global-mean surface temperature". [Government of Germany]	Taken into account. Text revised.
9-289	9	5	9	5	10	Isn't there equally large uncertainty associated with carbon-cycle feedbacks? Please qualify this if it arises from just AOGCMs and not ESMs. [duane waliser, United States of America]	noted. Carbon cycle feedbacks do not enter into equilibrium climate sensitivity (as defined in glossary).
9-290	9	5	9	5	11	Suggested text:"The primary factor contributing to the spread of climate sensitivity values continues to be cloud feedback effects. This applies to simulations of both the modern climate and that during the last glacial maximum. There is a strong correlation between SST.....in the models." [Ian Smith, Australia]	Accepted, thank you!
9-291	9	5	9	5	13	ES : A gap of this chapter (and of chapter 7) is the lack of interpretation of the spread of cloud feedbacks in multi-model ensembles (it is only discussed a bit in chapter 9 for multi-physics ensembles)...while it was one focus of the AR4. What can we say about it in the AR5 ? [Sandrine BONY, France]	Taken into account. This is primarily an issue for Chapter 7.
9-292	9	5	9	5	13	Although cloud feedbacks remain the primary factor contributing to the spread in equilibrium climate sensitivity, there have been improvements in climate model simulations of clouds, including ice and liquid water path, and cloud fraction of highly reflective clouds (Jiang et al., 2012; Klein et al., 2012). [Government of United States of America]	noted. Some areas of improvement are discussed in the body of the chapter and in Chapter 7.
9-293	9	5	10			Need to refer to 9.7.2.3.2 [Julia Hargreaves, Japan]	Editorial
9-294	9	5	11	5	13	The confidence in water vapour feedback needs to be based on more than just correlations with SSTs. Indeed this correlation alone does not imply a global positive water vapour feedback. Suggest this be changed to indicate that a wide range of evidence indicates that models get this correct. Briefly enumerate representation of regional, seasonal, interannual variations in humidity, representation of humidity trends, etc. [Government of Australia]	Taken into account. Text revised.
9-295	9	5	11	5	13	How does the strong correlation between SST and water vapor imply a "realistic and positive water-vapor feedback"? More is needed here. [Ronald Stouffer, United States of America]	Taken into account. Text revised.
9-296	9	5	12	5	12	The abbreviation SST has not been introduced yet. [Farahnaz Khosrawi, Sweden]	Editorial.
9-297	9	5	12	5	13	Why does a strong correlation between water vapour and SST point at a strong feedback? It can also be a strong response from one variable to the other [Bart Van den Hurk, Netherlands]	Taken into account. Text revised.
9-298	9	5	15			I think this title is misleading; for me this section tries to answer the question if better model performance might mean a better forecast. It highlights that some model differences with respect to observations are mirrored in deviations in future projections of the same models. It does not define a particular application or defines credibility (are predictions better when the recent climate is captured better?). I would call it "Relating model performance for the recent past relates to confidence in prediction" [Peter Braesicke, United Kingdom]	Taken into account. Text revised.
9-299	9	5	17	5	22	ES : It should be added : « However, such correlations emerge only for a few quantities, such as... » [Sandrine BONY, France]	Taken into account. Text revised.
9-300	9	5	17	5	22	This is quite misleading since only a handful of papers have demonstrated that inter-model differences in modern day skill scores correlate to differences in projections. You are taking something that is very rare, and implying that it is commonplace. [Gavin Schmidt, United States of America]	Taken into account. Text revised. Original text made clear that evidence was 'emerging' -- this clearly does not imply 'commonplace'.
9-301	9	5	17	5	22	Suggest omitting since this refers to model projections and belongs elsewhere in the Report. [Ian Smith, Australia]	rejected. Many review comments have been critical of there not being enough in this chapter to relate model performance to future projections.
9-302	9	5	24	5	27	So how can you use the word "reliability" when you haven't found a way to use these assessments to differentiate predictions of climate change? Most of what we have to date is a beauty contest. If I have a beautiful car and an ugly car side-by-side, can I tell which one is going to start up when I turn the key in the ignition? The one that starts is the more "reliable," not the more beautiful one. A Jaguar is prettier than a Honda, but Jaguars tend to spend more time in the repair shop. [Anthony Del Genio, United States of America]	Taken into account. Text revised to avoid the word 'reliable'.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-303	9	5	29	5	37	Chapter 9: Executive Summary. It would be desirable that the order of appearance of the different contents in the executive summary will follow the structure of the chapter. In particular it is recommended to move the paragraph concerning Regional Downscaling to the top of page 5, just before that related to Climate Sensitivity in the CMIP5 Ensemble. [Government of Spain]	Accepted.
9-304	9	5	31	5	31	ES : This assessment is too general and should be made much more specific. Regional downscaling does provide added value. But what is the nature of this added value ? It is for simulations of the current climate or for climate projections ? For the latter, is it primarily in providing more geographical details, or in providing more confidence in the projections? In some cases, e.g. over highly variable topography of for some small-scale regional phenomena, the added value can come from both. However in many other cases, especially when the climate change signal at the regional scale is similar to (or largely constrained by) the climate change signal in the driving global model, the added value might be more in providing details (which can be useful for some impact studies) than in providing confidence. The ES should be much more specific regarding the nature of the added value. [Sandrine BONY, France]	Taken into account. Executive Summary text has been revised. It must be noted however, that there is as yet rather little published literature on this topic and this limits the extent to which very specific statements can be made.
9-305	9	5	31	5	31	"providing credible ...": We recommend that the authors qualify the statement by adding the word "some" between "providing" and "credible." [Government of United States of America]	Editorial
9-306	9	5	31	5	31	This is ambiguous. Does this refer to spatial patterns or trends? One does not imply the other. [Gavin Schmidt, United States of America]	Taken into account. Text revised.
9-307	9	5	31	5	31	Suggested text:"..downscaling methods offer a means..." [Ian Smith, Australia]	Editorial
9-308	9	5	31	5	33	I don't agree; this statements is not universally true and depends on region, the impact of topography, etc. - please rephrase [Peter Braesicke, United Kingdom]	Taken into account. Text revised.
9-309	9	5	31	5	37	I agree with the fact that « downscaling adds value both in regions with highly variable topography and for various small-scale phenomena » but downscaling (in particular in the frame of the Cordex with very large domains) adds also another source of dispersion in the models . Part of this additional dispersion may be for good reasons (additional processes, larger exploration of the possible model physics). But the fact that the coupling with the general circulation is not accounted for in the limited area model may introduce an other source of dispersion and error which may be linked purely to the methodology. For a very sensitive region such as the Sahel, most impact studies depend to first order on the annual cumulated rainfall, whose variations (either due to internal variability or external forcing) are very large scale. For that particular region, the mean rainfall patterns are not better represented in Cordex (see eg. Nikulin et al., 2012, already in the reference list) than in AMIP/CMIP5 simulations (see Roehrig et al., in revision in climate dynamics, ftp://cnrm-ftp.meteo.fr/pub-moana/roehrig/CMIP5_WAM/rev1/Roehrig_et_al2012_JClim_rev1.pdf). In this case, the response of global models, even with a rather coarse grid, may be more reliable than that of the regional models. It is important to emphasize that this question has not been investigated so far. So I would be much more restrictive at the beginning of the paragraph and I would emphasize that downscaling must be every time checked carefully in terms of consistency or added value with respect to the large scale models and that the additional information provided at intermediate scales (resolved both by both the global and regional model) is not more reliable, a priori, in regional than in global model. It is very important for me to advertize strongly that the Cordex exercise is a very interesting first coordinated exercise but that it is by no way the gateway of impact studies to climate projections. [Frédéric HOURDIN, France]	Taken into account. Text revised.
9-310	9	5	39	5	44	The conclusions read well, and appear well justified by the chapter material. Suggest add at end "... and for the attribution and understanding of past climate change.", which reflects the other principal usage of models in the report. [Government of Australia]	Taken into account. D&A added to the ES bullet on suitability of models.
9-311	9	5	39	5	44	This sounds like a summary and thus I would suggest to start this paragraph with "In summary,..... ". [Farahnaz Khosrawi, Sweden]	Editorial
9-312	9	5	39	5	44	This paragraph seems to allude that the overall performance of the CMIP5 models cannot be judged to be superior to that of the CMIP3 models. If so, I recommend to make the this point more explicitly and to explain (or speculate) why. [Elisa Manzini, Germany]	rejected. Text is explicit that some aspects of performance have improved. Executive Summary is no place for speculation.
9-313	9	5	39	5	44	Suggest omitting this text since it does not add materially to the key messages. [Ian Smith, Australia]	rejected. We feel it is important to conclude the

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							Executive Summary with a concluding paragraph.
9-314	9	5	39	5	44	this paragraph provides a nice summary. It might be better placed upfront in the ES before going onto the details of the many components; i.e., we suggest to consider moving it page 3, line 37, just before the "Simulations of Historical Climate subsection" [Thomas Stocker/ WGI TSU, Switzerland]	Noted. Executive Summary has been substantially revised. Our feeling was that it was important to end with some overall conclusion.
9-315	9	5	41	5	41	The lack of significant improvement between AR4 and the present report is 'disappointing'. This needs to be stated here and elsewhere in the document in order to help direct funding into needed areas of research. [David Webb, United Kingdom]	Rejected. While it may be 'disappointing', this chapter is intended to present the state of scientific evaluation of climate models, not to provide arguments for research funding.
9-316	9	5	42	5	42	"CMIP3 multi-model ensemble remains useful for many applications": Ambiguous. Which applications? [Elisa Manzini, Germany]	Taken into account. Text revised. Intent here is to provide a general statement. It is not possible to provide a detailed list.
9-317	9	5	42	5	44	Models still exhibit ... future climate.' --> This might sound to many people to good to be true: despite their errors and biases models are ok" : could you provide some extra argumentation which makes this statement more plausible? [Hans Visser, The Netherlands]	Taken into account. ES has been largely re-written and this paragraph removed.
9-318	9	5	43	5	44	If you believe that models provide a "physically-sound" basis for future projections, why have the metrics described in this chapter not been successful in separating good from bad model predictions of the future? At what level are they physically sound (the radiative effect of GHG increases to be sure, the water vapor feedback at least qualitatively), and at what level is it premature to consider them physically sound (cloud feedback, drought severity, hurricanes)? I understand the desire to make clear statements in the face of confusion in society about how to regard scientific uncertainty, but one must better differentiate the more certain and less certain aspects, even in summary paragraphs such as this. [Anthony Del Genio, United States of America]	Taken into account. Text revised. The statement reflects the overall assessment by the author team based on the totality of evidence presented in the chapter, all of which cannot be repeated in the Executive Summary.
9-319	9	5	43	5	44	A comment such as that in Chapter 12 (pg. 74, 36-39), would be beneficial here. "There are inevitable uncertainties around future external forcings, and the climate system's response to them, further complicated by internally-generated variability. These uncertainties make the use of multiple scenarios and models a standard choice if we are to assess and characterise them, describing a wide range of possible future evolutions of the Earth's climate. [Government of Australia]	rejected. Such a statement really does belong in Chapter 12. This chapter is concerned with evaluating models, not with their specific application to future projections which is the remit of Chapter 12.
9-320	9	5	44	5	44	Change "predictions and projections of future climate" to "generating possible projections of the future climate, consistent with our current physical understanding but subject to the limitations of model representation? " [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text revised. Intent here is to provide a clear concise summary statement.
9-321	9	5	46			There is nothing in the summary on simple or intermediate complexity models, which reflects essentially that there is very little in the chapter on those as well. Of course CMIP5 is the dominant source of information in AR5, but there are significant conclusions on allowed carbon emissions, climate sensitivity, commitment and uncertainty in global projections that are derived from simpler models. This chapter needs to assess all classes of models, not just CMIP5. [Reto Knutti, Switzerland]	Taken into account. Text revised to include material on EMICs.
9-322	9	5	46			There isn't any summary statement on individual model performance, e.g. some models clearly being better for worse overall wrt observations. This isn't an easy topic, and I agree there is no simple way to weight models, but one could easily defend a statement that a simple comparison to observations clearly reveals that some models are more advanced, realistic, and complete than others. And that there is no a priori argument to treat them equally for all cases. And that the choice of the models to trust will likely depend on the question at hand. Etc. [Reto Knutti, Switzerland]	Taken into account. Text revised to include a bit more on 'spread' in model performance and the need to assess suitability for particular purposes. Certainly the Executive Summary is not a place to list attributes of individual models. Further, there is no compelling scientific evidence for an overall performance metric that would allow one to rank or categorize models unambiguously. Therefore the summary statement must necessarily remain somewhat generic.
9-323	9	6	1			I am surprised that section 9.1 does not discuss the model equations and formulations for the dynamical core, and their improvements, focusing only on improvements in parameterizations. Examples of equation choices include Boussinesq/non-Boussinesq, hydrostatic/non-hydrostatic. Example of model formulations include the choices of vertical and horizontal coordinate, for example pressure coordinate/isopycnal coordinate for ocean	noted. This chapter cannot provide a comprehensive review of model development.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						models. An example of a CMIP5 ocean model with new vertical coordinate is given in Dunne, John P., Jasmin John, Alistair Adcroft, Stephen M Griffies, Robert W Hallberg, Elena Shevliakova, Ronald J Stouffer, W F Cooke, Krista A Dunne, Matthew J Harrison, John P Krasting, Sergey Malyshev, P C D Milly, Peter Philipps, Lori T Sentman, Bonita L Samuels, Michael J Spelman, Michael Winton, Andrew T Wittenberg, and N Zadeh, October 2012: GFDL's ESM2 global coupled climate-carbon Earth System Models Part I: Physical formulation and baseline simulation characteristics. Journal of Climate, 25(19), doi:10.1175/JCLI-D-11-00560.1. [Sonya Legg, United States of America]	
9-324	9	6	5			"Climate models constitute the primary tools available for investigating the response of the climate system to various forcings. Climate models constitute the primary tools for making climate predictions...etc." We recommend that the authors improve the statement by adding ", along with observations," between "Climate models" and "constitute." [Government of United States of America]	noted. The focus of this chapter is on models and they are the primary 'tools'. The combination of models and observations to decompose forcing and response is covered in detail in Chapter 10.
9-325	9	6	8	6	20	reference is made here only to Chapters 10 through 12, however this should be extended to also mention Chapters 13 and 14 which also deal with projections and Ch9 includes, e.g., a subsection on sea level model evaluation. In addition, the Annex I: Atlas should be mentioned already here in conjunction with these Chapters [Thomas Stocker/ WGI TSU, Switzerland]	accepted. Text revised, also in ES.
9-326	9	6	9	6	10	This seems an odd way to cross reference these chapters - chapter 10 is d&a and 11 and 12 projections. Would be clearer to say so explicitly here. [Peter Thorne, United States of America]	taken into account. Text revised.
9-327	9	6	10	6	10	What does "necessarily an incomplete evaluation " mean? [Jui-Lin (Frank) Li, United States of America]	noted. It means that all possible aspects of model evaluation are not included. The focus is on those aspects particularly relevant to the use of model results later in the report.
9-328	9	6	11	6	11	... Projects CMIP3 and... [European Union]	editorial
9-329	9	6	11	6	12	...(CMIP3 and CMIP5) (Meehl et al., 2007a; Taylor et al., 2012) ... [Hai Lin, Canada]	editorial
9-330	9	6	11	6	12	"(CMIP3 and CMIP5 (Meehl et al., 2007a); (Taylor et al., 2012)" can be written as "CMIP3 (Meehl et al., 2007a) and CMIP5 (Taylor et al., 2012)" [SAMIR POKHREL, INDIA]	editorial
9-331	9	6	12	6	12	Section 9.1.1 : instead of « a set of well-controlled and increasingly well-documented climate model experiments», I would rather say « coordinated and thus consistent set of climate model experiments ». [Sandrine BONY, France]	taken into account. Text revised.
9-332	9	6	12	6	12	... as these constitute... [European Union]	editorial
9-333	9	6	17	6	20	CMIP3 models were already evaluated in AR4. So, possibly, here analyses of CMIP3 model-output that have been published after AR4 are assessed as well? [Elisa Manzini, Germany]	noted. This is stated explicitly in the previous sentence.
9-334	9	6	18	6	19	"where only CMIP3 results are available": presumably should be "where only results based on CMIP3 are available" . [Martin Juckes, United Kingdom]	editorial
9-335	9	6	22	6	22	"direct approach"? Is there an "indirect" approach? [Elisa Manzini, Germany]	noted. This is stated in the last sentence of the paragraph.
9-336	9	6	24	6	24	The authors probably mean "Chapters 2 through 5" instead of "6". [Government of Germany]	noted. Observations related to the carbon cycle are included in Chapter 6.
9-337	9	6	24			The discussions in Chapter 2 about uncertainties of satellite observations are mainly focused on the uncertainty in the satellite observation derived trends. There is no mention of uncertainty in the observed climatological mean values. We suggest that this topic should be added to the discussion here. [Government of United States of America]	rejected. Observational uncertainties are to be discussed in the relevant observational chapter, in this case, Chapter 2.
9-338	9	6	25	6	26	Please also consider that sometimes only observed anomalies to the mean are available. [Government of Germany]	taken into account.
9-339	9	6	25			The authors should define what "short" means for the observational record (e.g., 10 years, etc.). [Government	rejected. We cannot provide a specific definition of

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						of United States of America]	'short' -- it depends on the context and on the variable in question, the timescales of its variability, etc.
9-340	9	6	26	6	28	Section 9.1.1 : As mentioned already in a comment about the ES, it should be explained that the inter-model spread does not assess the same uncertainty as model-observation comparisons. [Sandrine BONY, France]	taken into account. We do not claim that it is the same, only that it provides some information where model/obs comparisons are unavailable.
9-341	9	6	26			I suggest using model spread or range rather than uncertainty, or at least being explicit. Model spread may not be representative for uncertainty. In chapter 12 we try say uncertainty when a formal uncertainty assessment is possible, and model spread when it's simply an ensemble of opportunity from the models. [Reto Knutti, Switzerland]	taken into account. We intentionally put 'uncertainty' in inverted commas to indicate that model spread is not a full/formal measure of uncertainty; but in some cases it is all that is available, and it is better than nothing at all.
9-342	9	6	28	6	28	The meaning of the word 'uncertainty' is undoubtedly vague, but the glossary provides some guidance. Following that definition, the quotation involving the word in this line would be unnecessary. See also page 19, line 14. Chapter 12.2.2 also provides a lengthy discussion. [Ramon de Elia, Canada]	taken into account. We intentionally put 'uncertainty' in inverted commas to indicate that model spread is not a full/formal measure of uncertainty; but in some cases it is all that is available, and it is better than nothing at all.
9-343	9	6	37	6	48	Section 9.1.2 : It could be mentioned somewhere that models using the same physics but in much simpler configurations (e.g. aqua-planet or single-column) are also used to evaluate some particular processes (SCM evaluations against observations and/or process models, should there be a subsection in 9.1.3.1 synthesizing and assessing GASS evaluations of models at the process level ?), or to help understand how the climate system work and help interpret inter-model differences in model results (an important part of CMIP5). [Sandrine BONY, France]	Accepted. The chapter focuses on evaluating the ESMs that are used in subsequent chapters on projections and detection and attribution. While not directly evaluated in the chapter, component models or simplified models (like Aqua-planet) are used in individual sections to strengthen the assessment of the more complex models, and they are indeed an important part of CMIP5. One example is the Lu et al. (2007) study discussing the tropical circulation.
9-344	9	6	37	7	40	This overview section concentrates on models which are relevant for the assessment report, which is very good (compared to the first draft of this chapter). Nevertheless, the chapter still contains some references which are not relevant, e.g. does refer to examples of exercises with other types of atmospheric models which do not belong to this assessment report. See specific comments below. [Martin Dameris, Germany]	Rejected. While the chapter focuses on evaluating ESMs, other types of atmospheric models are used in addition where they strengthen the assessment.
9-345	9	6	40	6	40	Use "require" rather than "using" [duane waliser, United States of America]	Editorial
9-346	9	6	41	6	44	Add sensitivity studies (especially for attribution studies) and process studies. [Martin Jukes, United Kingdom]	Noted and added in the revision.
9-347	9	6	42	6	42	it should be "paleo" [Annalisa Cherchi, Italy]	Editorial
9-348	9	6	42	6	42	... paleo... [European Union]	Editorial
9-349	9	6	46	6	46	Could you explain here how you treat models providing ensembles and models providing just one simulation? How do you average them? [Chiara Cagnazzo, Italy]	Thank you for this clarification request. The treatment is different depending on the use of the model simulations. When taking the MME mean, all available models from the CMIP archives are included with equal weight. For certain figures (e.g., Figure 9.8), only a single ensemble member is included where a comparison of an ensemble mean (with unequal ensemble sizes) would unfairly compare results. This issue is clarified in specific context for a given figure.
9-350	9	6	46			"Examples of the latter include". Suggest to remove "of the latter. We suggest it would be more clear that the examples are for the applications of simplified models. [Government of United States of America]	Accepted and changed in the revision.
9-351	9	6	54	6	56	Why are these models used particularly for seasonal to decadal applications? Please explain. [Martin Jukes, United Kingdom]	Noted and clarified in the revision to be distinct from simulations requiring biogeochemical feedbacks.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-352	9	6	55	6	55	higher than what? [Bart Van den Hurk, Netherlands]	Noted. This is typically clear in the context of models used for seasonal and decadal predictions as distinct from ESMs.
9-353	9	6	55	6	56	Seasonal and especially decadal predictions are currently in research state, so references to the respective projects are necessary. [Government of Germany]	Given the general nature of this chapter, references for seasonal and decadal predictions are provided in Chapter 11 and omitted here.
9-354	9	6		7		I expect that I don't need to mention that this could be modified in light of Fasullo and Trenberth – which considers the correlation between biases and climate sensitivity. [Julia Hargreaves, Japan]	Unfortunately, it is not clear where this comment is referencing.
9-355	9	7	2			Instead of writing "used on its own" a more specific description should be given such as "decoupled from the ocean part" [Government of Germany]	Noted and considered in revision. We note that an AGCM can also be coupled to a slab-ocean model and thus may not be completely decoupled.
9-356	9	7	4			Sec 9.1.2.2: I think it's worth saying explicitly that ESMs ARE GCMs, with additional functionality, just so there's no ambiguity that we've moved away from GCMs as climate models [Chris Jones, United Kingdom of Great Britain & Northern Ireland]	Noted and incorporated in revision.
9-357	9	7	10	7	10	Strike "potentially" - it is important" [duane waliser, United States of America]	Accepted.
9-358	9	7	13	7	13	Table 9.1 is very useful [Government of United Kingdom of Great Britain & Northern Ireland]	Agreed. Note: Table 9.1 has been moved to the appendix as Table 9.A1 and a new Table 9.1 is used within the main text.
9-359	9	7	14	7	21	Which models are ESMs? Is there a clearly defined distinction, or are "AOGCM" and "ESM" these vague terms with considerable overlap? [Martin Juckes, United Kingdom]	Thank you for clarifying this. AOGCMs and ESMs are defined in 9.1.2.1 and 9.1.2.2 respectively. These are further clarified in Table 9.1 and Table 9.A1.
9-360	9	7	14	7	21	The column "Aerosols" in this table leads to confusion. Most models have "Not implemented" in this column, giving the impression that aerosols are not taken into account. Of course they are, but the aerosol concentration is prescribed, while this column is about an active module calculating aerosol concentrations. To avoid confusion, the column head should be changed into "Interactive Aerosols", or, alternatively, the table entries for models with prescribed aerosol contributions should be "Prescribed" instead of "Not implemented". [Andreas Sterl, Netherlands]	Taken into account. Table clarified and simpler Table added
9-361	9	7	16	7	16	"... and flux correction implementation (not yet described);..." does not really belong into this column. Information on the usage of a flux correction, which is an important aspect of an AOGCM or ESM, should rather be added to the "Ocean" column [European Union]	Note: Table 9.1 has been moved to the appendix as Table 9.A1 and a new Table 9.1 is used within the main text. The flux correction is included for CMIP3 models as a separate column but not in Table 9.A1 describing CMIP5 models. The description is no longer included in the caption for Table 9.A1.
9-362	9	7	17	7	17	Please clarify what is meant by "code independence". This does not necessarily need to be part of the caption, but could be mentioned in the text. [Government of Germany]	Note: Table 9.1 has been moved to the appendix as Table 9.A1 and a new Table 9.1 is used within the main text. The discussion of model independence is provided in Section 9.2.2.
9-363	9	7	18	7	18	"(low or high top)" should be avoided. Table 9.1 simply provides information on the pressure or height of the uppermost level, but does not judge whether this is "low or high top" [European Union]	Table 9.1 has been moved to the appendix as Table 9.A1 and a new Table 9.1 is used within the main text. The term HT has been kept here to indicate a distinct model category.
9-364	9	7	20	7	21	Please make sure to update this list, if more information becomes available. [Government of Germany]	Agreed. Note: Table 9.1 has been moved to the appendix as Table 9.A1 and a new Table 9.1 is used within the main text.
9-365	9	7	21	7	21	I suspect the date given in the caption (12 Nov 2011) is wrong as the same date was given in the FOD but this Table has clearly been updated considerably! [Tim Johns, United Kingdom of Great Britain & Northern Ireland]	Accepted. Note: Table 9.1 has been moved to the appendix as Table 9.A1 and a new Table 9.1 is used

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							within the main text.
9-366	9	7	32	7	39	It is unclear in Section 9.1.2.4 that regional climate models are largely equivalent with dynamical downscaling methods (although semantically they are not equivalent). Consider clarifying in p9-7, L34 that RCMs are "physics-based limited area models". In addition, the last sentence in that subsection, p9-7, L38-39 should be clarified to contrast empirical and statistical methods. [Government of United States of America]	Noted and incorporated in the revision.
9-367	9	7	34	7	39	Should mention explicitly in this paragraph that RCMs require the specification of lateral boundary conditions. It is at best only implied here. [duane waliser, United States of America]	This was considered in the revisions but not included given that a limited area model requires a lateral boundary implicitly.
9-368	9	7	41			Model improvements are slightly out of place here, this should be mentioned as part of the development cycle description for each model family. [Peter Braesicke, United Kingdom]	Noted: Model improvements often apply across different model types. A change in the atmospheric model affects both the AOGCM as well as the ESM and if the model is used to tune an EMIC, than even that is affected. Given this and the importance of reporting model improvements it is appropriate to have a separate section here.
9-369	9	7	43	7	47	Unclear sentence. [Ramon de Elia, Canada]	Taken into account: The sentence has been simplified.
9-370	9	7	44	7	47	This sentence is not well constructed and overuses commas, making it difficult to read. Suggest breaking this sentence up. [Government of Australia]	Taken into account: The sentence has been simplified.
9-371	9	7	45	7	45	... components, entirely... [European Union]	Editorial
9-372	9	7	45	7	45	Remove extra comma [Stephen Griffies, United States of America]	Editorial
9-373	9	7	45	7	45	consider rephrase "...in the various model components entirely new model components" [Jui-Lin (Frank) Li, United States of America]	Taken into account: The sentence has been simplified.
9-374	9	7	45	7	45	remove one comma in "components,,entirely" [Hai Lin, Canada]	Editorial
9-375	9	7	45	7	45	Is there text missing in this line? [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Taken into account: The sentence has been simplified.
9-376	9	7	45	7	45	,, --> , [François Massonnet, Belgium]	Editorial
9-377	9	7	45	7	45	double comma is used between components and entirely [SAMIR POKHREL, INDIA]	Editorial
9-378	9	7	45	7	45	Repeated commas here. [Ian Simmonds, Australia]	Editorial
9-379	9	7	45			... components,, (typo) [Government of France]	Editorial
9-380	9	7	46			"spatial and temporal resolution of the models is improved". The mention of temporal resolution in this context is ambiguous. High spatial resolution usually requires shorter time stepping, but the time-step itself may not have climatological significance, except for geo-bio-chemical processes? Did the authors refer to such processes? [Government of United States of America]	Taken into account: The references to spatial and temporal have been removed as they are unnecessary in this introductory statement and are explained in more detail later in the report.
9-381	9	7	47	7	48	I would say that the assembly of all model components should also produce a realistic climate. [Ramon de Elia, Canada]	Noted: The role of tuning is extensively explained in Box 9.1. We feel that there is no need to provide extensive detail in what is merely an introductory statement.
9-382	9	7	48	7	48	... uncertain or not observable model parameters... [European Union]	Noted: The role of tuning is extensively explained in Box 9.1. We feel that there is no need to provide extensive detail in what is merely an introductory statement.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-383	9	7	53	8	25	Box 9.1: Considering the variety of ways in which the term "model" is used in different fields of activity, this is a necessary and important reminder, forcefully expressed. [Robert Kandel, France]	Noted
9-384	9	7	53	9	23	BOX 9.1: This is a very good and concise explanation of tuning. [Andreas Sterl, Netherlands]	Noted
9-385	9	7	53	9	25	Box 9.1: When discuss model turning, it may be needed to emphasis here that: (1) observational constrains on a tunable parameter cannot not be exceed; (2) also need to make sure that tuning of a parameter (e.g. cloud) would not result in significant errors of another parameter (e.g. water vapor) that is also constrained by observation; (3) Since this box is to provide the readers of some knowledge, it would be good to give a few examples of most often used tunable parameters, e.g. gravity wave drag parameters are not very constrained by data, and so are often tuned to improve the climatology of stratospheric zonal winds; or the threshold relative humidity for making clouds is also tuned often to get the most realistic cloud cover and global albedo. [Government of United States of America]	Noted: As mentioned in the box, tuning practices applied in various modelling centers are usually not reported (Page 9, Line 1). So while we agree with the comment, there is no hard evidence that it is true. When writing the box we rely only on the (very small amount of) literature available to us, which is cited throughout the box.
9-386	9	7	53			Need to emphasize here that (1) observational constrains on a tunable parameter cannot not be exceeded; (2) tuning of a parameter (e.g. cloud) should not result in significant errors of another parameter (e.g. water vapor); (3) Since this box is to provide the readers some general knowledge, it would be good to give a few examples of mostly used tunable parameters. For example, gravity wave drag parameters are not very constrained by data so that they are often tuned to improve the climatology of stratospheric zonal winds. Precipitation efficiency is an unknown parameter in climate models and thus often tuned to make precipitation and clouds close to the observed values. [Government of United States of America]	Noted: As mentioned in the box, tuning practices applied in various modelling centers are usually not reported (Page 9, Line 1). So while we agree with the comment, there is no hard evidence that it is true. When writing the box we rely only on the (very small amount of) literature available to us, which is cited throughout the box.
9-387	9	7	55	7	55	Replace "laws of nature" with "laws of physics". Laws of nature has a much wider meaning including chemistry, biology. [Josephine Brown, Australia]	Rejected: Treatements of biochemistry are now common in ESMs and those are based on laws beyond just physics.
9-388	9	7	55	8	39	I have difficulty with your statement that the models being evaluated are "based on" fundamental laws of nature. You go on to say that many of the most important processes (clouds, turbulence, biogeochemical processes, aerosols, precipitation, the carbon cycle, etc.) cannot be represented in terms of their fundamental physical laws so they must be approximated through empirical parameterizations. So I suppose the models are "based on" fundamental laws of nature the way this or that movie is "based on" a true story. It would be more accurate, I think, to say that the models are mainly built up using approximations to fundamental processes that are put together in a way that try to avoid, as much as possible, violations of fundamental laws of nature, especially the laws of conservation. [Ross McKittrick, Canada]	Noted: The function of this box is to provide information on the model development process, including the approximations that need to be made and why this need arises. Parametrizations, while conceptual simplifications of unresolved processes, are still based on the fundemental laws of physics, in that they are designed to conserve mass, energy, momentum etc. The first law of thermodynamics is the foundation of any parametrization of condensation, even though we cannot describe the formation of every droplet. Given this and the rather extensive discussion of approximations that follows, we see no need to change this statement.
9-389	9	8	1			I prefer preinciple over law andwould stress the fact that a system of equations can be written to express the model [Peter Braesicke, United Kingdom]	Noted: We prefer to remain with the word "laws" for consistency and to stress the strong difference of climate models from other models readers might be familiar with, such as economic of hoydrological models, many of which are statistical models. We feel that the use of "mathematical expressions" instead of "equations" makes the statement more readable to the lay reader.
9-390	9	8	3	8	3	Could you also mention here approximations (such as hydrostatic...) [Chiara Cagnazzo, Italy]	Rejected: The box is meant to provide general information on the model development process and the hydrostatic approximation is a rather specfic and very technical example. The general need for and implementation of approximations are discussed later in the box.
9-391	9	8	6			the solution/solving [Peter Braesicke, United Kingdom]	Accepted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-392	9	8	7	8	9	Mention here that there are several "Good Modelling Practices Guidelines" that are trying to uniformize the evaluation process in all type of models. They should be mentioned here as well as recommended for future developers of climate users: [1] Van Waveren, R. H., Groot, S., Scholten, H., Van Geer, F. C., Wösten, J. H. M., Koeze, R. D. and Noort J. J. (1999) Good Modelling Practice Handbook (STOWA Report 99-05), Utrecht, The Netherlands. [2] Council for Regulatory Environmental Modeling (CREM) (2009) Guidance on the Development, Evaluation, and Application of Environmental Models, US Environmental Protection Agency, Washington, DC. [3] Kimmins J.P., Blanco J.A., Seely B., Welham C., Scoullar K. 2010. Forecasting Forest Futures: A Hybrid Modelling Approach to the Assessment of Sustainability of Forest Ecosystems and their Values. Earthscan Ltd. London, UK. 281 pp. ISBN: 978-1-84407-922-3. [Juan Blanco, Spain]	Noted: These studies are not fundamental to the evaluation of climate models. As the purpose of this section is to describe the current state-of affairs, they were not included here.
9-393	9	8	11	8	11	should read biochemical and not biogeochemical (if only seen for plants and not for entire ecosystems) [European Union]	Accepted
9-394	9	8	11			Change "spatial scales" to "spatial and/or temporal scales". The time steps of some global models preclude resolving cloud and turbulence processes. [Government of United States of America]	Accepted
9-395	9	8	16	8	16	Change "(e.g., (Randall et al., 2007))" to "(e.g., Randall et al., 2007)" [SAMIR POKHREL, INDIA]	Editorial
9-396	9	8	18			Complex climate models, with reasonable configurations (3° resolution on the horizontal and 40 vertical layer for instance) can be integrated on several years on recent laptops for instance. So I would rephrase as something like : « Long term integrations with high resolution complex climate models require ... » [Frédéric HOURDIN, France]	Accepted: We replace "complex" with "state-of-the-art".
9-397	9	8	23			As written, this sentence is contradicted in the paragraph beginning on p9-64, L50. Change "generally leads to" to "can lead to". [Government of United States of America]	Rejected: There is no disagreement between the statement here and that on p9-64 once the qualification that follows is taken into account. With no other changes, higher resolution provides better mathematical accuracy but not necessarily a better model.
9-398	9	8	25	8	25	Turbulence is not expected to be resolved in a GCM. This is quite independent of the computing costs. It would be better to write: The finite resolution of climate models implies... [European Union]	Accepted. Text modified.
9-399	9	8	25	8	27	"Currently affordable climate model resolutions imply that the effects of certain processes must be represented through parameterisations. " In reality, certain effects will always have to be represented through parameterizations, no matter the level of supercomputing capability. For example, even with cloud resolving models, cloud droplet distribution will need to be parameterized. Even if cloud droplets themselves could be resolved, surface tension and phase change of each droplet would need to be parameterized, and so on. We recommend that authors consider the following change: "Regardless of the supercomputing power, model closure at the resolution limit requires that certain processes be represented through parameterisations." [Government of United States of America]	Noted: We agree with this statement. However, we feel it is of rather philosophical nature. This report assesses the current generation of climate models and we aim to point out that in those some key processes, such as those related to clouds and precipitation are not resolved. We have changed the sentence based on comment 9-398.
9-400	9	8	25			The text speaks of "affordable" simulations. Affordable for which purpose and timescales? The computational challenge should be mentioned explicitly here in relation to the necessary application. [Alexander Loew, Germany]	Noted: We mention the role of supercomputing limitations in making model choices above. We feel that discussing choices made in individual applications separately is too technical for the purpose of this box.
9-401	9	8	29	8	30	Please consider to give one or two examples as in (I). [Government of Germany]	Accepted: We added the carbon cycle as an example.
9-402	9	8	30			"Hence compromises to include or exclude certain processes or components in a model must be made, recognizing that an increase in complexity generally leads to an increase in computational cost. (Hurrell et al., 2009)" Rather than a compromise, it is a mathematical prioritization of the most important drivers of the solution, based on how much complexity is affordable. We recommend that the authors consider something along the lines of the following: "Once the time scale of interest to the modeler is determined, processes can be ranked using mathematically based analyses and then included if they meet a threshold of importance. This threshold would be based on how much complexity can be afforded given the available computing power to perform the simulation." [Government of United States of America]	Noted: This comment refers to scale analysis in deciding which terms to retain in an equation based on the scales of interest. We are not aware that decisions whether or not include components of say the carbon cycle in a climate model are made in this way. We also feel that the word compromise is appropriate in the larger context of this Box, as one might compromise a model component for resolution

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							in other components.
9-403	9	8	31	8	32	leads to an increase in computational cost (Hurrell et al., 2009) and sometimes to additional systematic biases if additional feedbacks are introduced. [European Union]	Noted: While this statement is correct, it is not relevant to the discussion in this part of the box, as we are trying to establish the reasons for the compromises made in model design, not discuss model errors at this point.
9-404	9	8	32			position of fullstop [Peter Braesicke, United Kingdom]	Editorial
9-405	9	8	34	8	38	I understand the point but I think that mixing the problem of initial conditions and model uncertainties is very confusing. Both points could be addressed in a unique iii) point as done here, but separating the two aspects. I would propose something like "Due to the sensitivity of the meteorology to the initial state, a unique simulation represents one of the possible pathways the climate system might follow, and ensemble simulations or long term equilibrium simulations are required to construct the statistics that define the "climate" of one particular model. Ensemble simulations are used also to evaluate the uncertainties that come from the model formulation itself, either by varying the tunable parameters for a given model or considering multi-model simulations » [Frédéric HOURDIN, France]	Taken into account: The sentence has been rephrased.
9-406	9	8	34			"Due to uncertainties ... initial model state" - There are other important reasons for running ensembles - variability and sampling issues for example. Delete phrase. [Ronald Stouffer, United States of America]	Taken into account: The sentence has been rephrased.
9-407	9	8	35			"one of the possible pathways": The combination of model formulation error and model initial state error here oversimplifies the issue. The bullets are about computational cost, but the issue is accuracy for a given level of complexity, when it comes to model formulation. Uncertainty in the model formulation does lead to uncertainties in the model accuracy, but "possible pathways" should refer to model ensembles with different but equally likely initial states. If a model formulation contains an error within the uncertainty, that does not mean the resulting simulation is one of the possible pathways a climate system might follow. It means that it is the best estimate our model is able to provide with a given initial state. Given the level of uncertainty in our model formulation, we should run a number of models with different formulations (such as cloud physics parameterizations of grid discretizations such as cubed sphere, lat-lon etc) with the same initial state and determine the spread of uncertainty. This does increase model cost. But initial state uncertainty is a different issue that should be addressed as a separate point. We recommend that the authors consider rephrasing along the following lines: "It is difficult to rank the relative importance of mathematical formulations within a global climate model, especially considering the range of needs these models are expected to provide. Due to uncertainties in the model formulation, such as cloud physics parameterization or grid discretization schemes, a single model simulation is one point in a spread of model error. To allow some evaluation of this uncertainty it is necessary to carry out a number of simulations with models with different formulations, which incurs significant computational cost." [Government of United States of America]	Taken into account: The sentence has been rephrased.
9-408	9	8	37			We recommend that the authors change the phrase "either of which increases" to "both of which increase". [Government of United States of America]	Accepted
9-409	9	8	39	8	39	Another limitation due to computation cost regards the length of the simulations: the longer, the more expensive. This is why we sometimes need to compromise on the length of the model spinup (e.g. ocean). This could eventually go as iv) after i) resolution, ii) complexity, and iii) ensemble. [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Taken into account: We have added length of simulation as one of the reasons for trade-off in the sentence on p9-8 line 40.
9-410	9	8	42	8	42	Would it be possible to add somewhere here or above that one of the key elements for running and analysing a climate model is that it has to conserve mass, energy and momentum? All models conserve dry mass, but wet mass is not exactly conserved, and the global water budget is not balanced in every CMIP model (in some models, the imbalance is larger than the interannual variability in global precipitation, and starts to diverge at the beginning of the simulation: Liepert and Previdi, 2012, ERL, 7, 014006). [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Noted: We feel that the issue of in-principle vs in-practice conservation is an important but rather technical aspect. Therefore we have not added it here.
9-411	9	8	42	8	42	Similarly, the global radiative budget at TOA is also not always in balance (Trenberth et al, 2009, BAMS, 90; Trenberth et al, 2011, J Clim, 24, 4907-4924) as written in the caption of Fig. 9.21 that could eventually come	Noted: The tuning of the TOA balance has been clarified to apply to the pre-industrial period only, as it

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						here, or in lines 51-52, instead of as a note in the caption, as radiative balance at TOA is a crucial aspect of climate modelling. [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	is the imbalance that drives recent changes in climate.
9-412	9	8	43	8	44	Although individual component testing is certainly an important part of the model development process, this sentence doesn't adequately convey the fact that climate modelling centres increasingly base next-generation model developments on a previous-generation coupled AOGCM and evaluate performance of changes against the previous model quite routinely in a coupled framework (e.g. in free-running control climate mode, initialized seasonal hindcast mode). I suggest the model development process is more balanced between 'unit' testing and 'integration/system' testing in practice. [Tim Johns, United Kingdom of Great Britain & Northern Ireland]	Noted: We do not feel that we elevate individual component testing over system testing. We merely mention that the former happens and then move on to the full system. We have added that the system undergoes a "systematic evaluation" to the next sentence to make this clearer.
9-413	9	8	43	9	23	We recommend that the authors emphasize that all models of a physical system across the computational science community are "tuned" using observational constraints of subgrid scale parameterizations. We recommend the following alternative language: "This final parameter adjustment procedure is usually referred to as model tuning and is a necessary procedure performed to reconcile subgrid parameterizations within any model of a physical system, using observed values as a basis. As model tuning aims to match observed climate system behaviour, it is connected to judgments as to what constitutes a skillful representation of the Earth's climate." [Government of United States of America]	Noted: Tuning is not limited to sub-grid scale parameters. "Implicitness" parameters in numerical schemes for instances are often "tuned" to keep the model solution stable. As the details of tuning are not documented by the modelling centres, we feel we cannot go beyond the broad statement we make. As the comment provides no evidence from the literature, we also see no way to include a statement that other fields commonly tune their physically-based models. Of course, statistical models are always "tuned", but they constitute a different class of models.
9-414	9	8	46	8	46	Change "(Barnier et al., 2006); (Griffies et al., 2009)" to "(Barnier et al., 2006; Griffies et al., 2009)" [SAMIR POKHREL, INDIA]	Editorial
9-415	9	8	46			Change "input" to "conditions" [duane waliser, United States of America]	Accepted
9-416	9	8	48		55	Tuning goes beyond general balance requirements or observational constraints. Tuning is much used in "parameterized process-level formulas" (in the form of adjustable coefficients that need to be tuned). Climate or Earth system models involve hundreds of such empirical coefficients in their "physics packages". Adjusting such empirical coefficients is a major source of uncertainty in climate simulations. In particular, tuning is required whenever a parameterized formula is imported from another model. The existence of such insufficiently-documented model adjustments is ignored or discounted in this write-up and should be acknowledged as reality. [Government of France]	Noted: The existence of many parameters in GCMs is acknowledged. However, most of those are determined when the parametrizations are designed. As stated in the text, very few of these parameters are adjusted at the time of the final model adjustment described here. We do state that the details of tuning are not documented by the modelling centres (SOD, p9-9, L1). We have added a sentence to state that the need for tuning likely adds to uncertainties in the models to the paragraph following this one..
9-417	9	8	49	8	55	Box 9.1 : The sentence « As model tuning aims to match observed climate system behaviour... » can be misleading, as it can be considered as the way climate models are evaluated and improved. I suggest to remove this sentence. I also suggest to make it clearer that the observed quantities considered as part of the tuning process of coupled climate models are most of the time global or zonal mean quantities, and that the tuning rarely affects at first-order the large-scale patterns simulated by models, including the long-standing model biases (e.g. double ITCZ, too extended cold tongue). Line 54-55 page 8 and Line 1 page 9 : Hourdin et al. (Clim. Dyn, 2012) describes the tuning process of the IPSL-CM5 model and may be added to the list. [Sandrine BONY, France]	Taken into account: We keep the sentence as it needs to be seen in the context of the whole box, which describes the model development process in its entirety. We have added that the large-scale constraints are often global averages, but since the tuning is not well documented in the literature, we cannot go beyond a tentative statement. We have added the Hourdin et al. (2012) reference
9-418	9	8	50	8	50	"skillful" -> "skillful" [Timothy Merlis, United States of America]	Accepted
9-419	9	8	51	8	51	TOA energy balance is only needed for the pre-industrial control - not the whole historical simulation [Gavin Schmidt, United States of America]	Accepted: Replaced "observed historical" with "pre-industrial"
9-420	9	8	52	8	53	"The models used in this report almost universally contain small adjustments to parameters in their treatment of clouds...". We disagree with the characterization of the adjustments as "small". Take for example the parameterization of autoconversion rate, Fig. 1 in Hsieh et al., JGR, 2009 shows that, depending on condensate amount the predicted rate can differ by 2-3 orders of magnitude. We would therefore recommend	Accepted: The word has been removed.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						removing the word “small”. A TOA energy balance can be (and often is) achieved by the compensating effects of substantial adjustments to commonly used parameters. Using the word “small” also seems to contradict with the statement in lines 1-2 of p. 9-9 that most modeling centres do not provide details about how they tune their models. [Government of United States of America]	
9-421	9	8	53	8	55	I would remove the word « small » before adjustment. I do not know how to decide whether the adjustment we do in our models are small or large. They change the radiative balance more than a doubling of CO2 for instance. For cloud tuning, you could add references to Hourdin et al., 2012 : LMDZB... (already in the bibliography of the chapter) for our model (the paper describe the new paradigm of model tuning under the constraint of single column evaluation by comparison with LES) and Golaz et al., 2012, J. of Climate, Sensitivity of the Aerosol Indirect Effect to Subgrid Variability in the Cloud Parameterization of the GFDL Atmosphere General Circulation Model AM3, J. Climate, 24, 3145–3160. doi: http://dx.doi.org/10.1175/2010JCLI3945.1 which discusses the role of tuning on the anthropogenic forcing. [Frédéric HOURDIN, France]	Accepted. "small" has been removed and the references have been added.
9-422	9	8	54	8	54	TOA energy balance is not an observed constraint. And certainly wasn't observed for the PI. It is an inferred constraint. [Gavin Schmidt, United States of America]	Taken into account: The word "observed" has been removed.
9-423	9	9	1	9	9	"Modelling centres do not routinely describe how they tune their models". This makes it difficult (impossible?) for the community to judge the models or to investigate the reasons for the large spread in model sensitivities. A stronger statement is needed. It also needs to be included in the Key Uncertainties section of the Technical Summary section. [David Webb, United Kingdom]	Noted: We feel this statement is sufficiently informative.
9-424	9	9	2	9	3	The expression "tune" is quite negative and invites criticism, although it is widely used in the community. Suggestion to use a more neutral wording, like "adjust to observations". Please change throughout the chapter. [Government of Germany]	Rejected: The word "tune" is in common use in the climate science community and does not by itself imply a negative action. Race cars are tuned to improve their performance. It is the purpose of this box to describe the process of model development including the final tuning and its consequences.
9-425	9	9	2	9	3	This statement together with the next paragraph seems to call for a documentation of the tuning procedure. Are there any efforts for future evaluations to prevent evaluating the models with the datasets that have been used for the tuning procedure? [Government of Germany]	Noted: We do not know if the details of the tuning procedure at various centres will be better documented in the future. If they are future reports can make use of that information in their evaluation efforts.
9-426	9	9	8	9	9	Glosses over the dimensionality issues of parameter optimisation in large models - with more than a few parameters a full optimisation becomes computationally impossible and even with optimistic projections of computational power, will be impossible for a very long time to come. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account: The sentence has been changed to state that the methods "can not be applied" rather than have not yet been applied.
9-427	9	9	9			Section 9.1.3.1. As elsewhere in this chapter and the chapter on clouds, there is a long list of papers and models, and a positive comment about each one, but in the end in most cases, I except sea ice, there is very little proof that they are anything more than a fashionable rearrangement of deck chairs. (OK - sorry for that but it is the obvious comment). What I think are needed are more hard statements saying which developments have resulted in significant improvements of the climate models and why. [David Webb, United Kingdom]	Noted: The fundamental improvement of parametrizations is an important part of climate model development. Due to model complexities and compensating errors, as single improvement in a parametrization does not automatically lead to better overall model results, and it is impossible from the existing literature which of the many changes between model generations leads to which improvement. Hence, this section simply aims to document changes at the process level.
9-428	9	9	11	9	17	Transparency and reproducibility is a base paradigm of science. Does the CMIP5 protocol not foresee an appropriate documentation of the model tuning exercises? I know it is more a philosophical question and can	Noted: There is no documentation of tuning across all models that we can use in this assessment. It is up

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						perhaps not be implemented in practice, as a lot of model tuning is based on expert knowledge rather than objective methods. However, one might want to think about how to deal with this fundamental problem for future CMIP exercises. [Alexander Loew, Germany]	to the CMIP community to alleviate this in their next intercomparison project.
9-429	9	9	11	9	17	This is a key statement and needs to be repeated strongly in the chapter introduction and elsewhere. It should also be referred to in the Key Uncertainties section of the Technical Summary section. [David Webb, United Kingdom]	Noted. This is the introductory section of the chapter and that is why the statement is made here. Owing to space limitations, repetition is generally avoided. Concerning elevation to the Executive Summary and Technical Summary, it must be noted that this statement is behind the assessments performed here and, thus, is implicit in the ES/TS text. Without the context of the box, however, this paragraph is incomprehensible, and owing to this and its technical nature, the suggestion is not followed.
9-430	9	9	15	9	17	"insight into model skill", "stringest test of model quality": If you really believe this, then it should be possible to discard all but a small subset of models, whose predictions of future climate will be pretty much the same. But it has not been possible to do this, since even models that are "skillful" or "high quality" as judged by some metric often disagree in their climate projections - as you demonstrate throughout the chapter and in the references you cite. In the absence of useful metrics for climate change, one should not claim anything about skill or model quality, because the supposed skill is meaningless for what the model is intended to be used for. [Anthony Del Genio, United States of America]	Rejected: We do believe this. It has been made clear throughout the chapter that performance measured against current climate is a necessary but not sufficient condition for a model's utility in future climate projections. This does not make the evaluation using current climate meaningless. Rather it makes it essential. As stated, the use of many quantities, measures and techniques does provide insight into model quality, even though this does not usually lead to model selection in an "overall performance" sense. Model rejection based on more detailed applications is possible and does occur. For instance, if one wishes to study ENSO in a future climate, naturally models that do not have an ENSO are excluded from being used.
9-431	9	9	15			Section 9.2.1: This needs to refer to the concerns expressed in page 9, lines 1 to 17, i.e. that (a) assessments should not include fields close to those used to tune the model and (b) the extra difficulties that arise because modelling centres do not routinely describe how they tune their models. [David Webb, United Kingdom]	WRONG SECTION
9-432	9	9	17	9	17	The term 'emergent behaviour' may deserve a place in the glossary. The term 'un-tuned' is not clear. [Ramon de Elia, Canada]	Taken into account:
9-433	9	9	19	9	19	typo: are "a" reliable for future [European Union]	Accepted
9-434	9	9	19	9	19	... are reliable... [European Union]	Accepted
9-435	9	9	19	9	19	...models are reliable... [Hai Lin, Canada]	Accepted
9-436	9	9	19	9	19	insert "tool" after "reliable" [Bart Van den Hurk, Netherlands]	Editorial
9-437	9	9	19	9	19	remove "a" before "reliable" [duane waliser, United States of America]	Accepted
9-438	9	9	19	9	23	One way to express this is that it is essentially a Bayesian method where the future is conditional on some past observations, even though there isn't a formal Bayesian process in which this is produced. [Reto Knutti, Switzerland]	Noted
9-439	9	9	20	9	20	It should be noted that the tuning not only targets stable climatologies (in long pre-industrial control simulations), but sometimes also trends observed in the past. The latter may be useful to constrain the climate sensitivity, if several parameter choices exist, but there is the risk that the tuning is specific to the model biases occurring for the past forcing history. If that is the case, the model will be to some degree detuned once the forcing changes in its characteristics. [European Union]	Noted: We changed this sentence slightly. As the details of tuning are not documented throughout the modelling community, we cannot tell if observed trends were used in the tuning of any models used here.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-440	9	9	20	9	21	"Models are not tuned to match a particular future; they are tuned to reproduce a small subset of global constraints observed in the current climate." We recommend adding the following sentence (or something like it) before the sentence above: "More recently, as more preindustrial data has become available, tuning to a relatively unforced state for a zero net model energy balance matching the preindustrial period is being adopted as a procedure (Gent et al, 2011- already cited in the text.)" [Government of United States of America]	Rejected: As there is only one study supporting this rather strong statement and as centers do not document their tuning procedure preventing us to know about the wide adoption of this technique or not, we cannot include the proposed statement.
9-441	9	9	21	9	22	This is where you need to add a statement to the effect that "and models that do not include greenhouse gas increases cannot reproduce the past." This is our best evidence to date of anthropogenically forced climate change. [Anthony Del Genio, United States of America]	Noted: This is the subject of Chapter 10 and would be out of place to comment on here.
9-442	9	9	21	9	23	Here I would argue that significant warming is consistent with our physical understanding of CO2 radiative forcing and how feedback mechanisms (water vapor and surface albedo) act to amplify the response, in addition to models agreeing among each other. [Thorsten Mauritsen, Germany]	Taken into account: We have added: "consistent with our physical understanding" to the sentence.
9-443	9	9	21			current climate is not we;ll defined ans most metrics evaluate against the recent past, a term I would prefer [Peter Braesicke, United Kingdom]	Taken into account: The sentence has been modified
9-444	9	9	22	9	22	reword "reproduce the past,universally produce significant" [Jui-Lin (Frank) Li, United States of America]	Editorial
9-445	9	9	22			... past,universally (typo) [Government of France]	Editorial
9-446	9	9	23	9	23	"... the results presented in this report" and the prevous ones. [Elisa Manzini, Germany]	Taken into account: The sentence has been removed.
9-447	9	9	23	9	23	This feels a little grandiose and a potential hostage to fortune. It is a weak and unneeded ending to what is a very well put together box. For a start much of the report has nothing to do with models or projections - viz. Chapters 2 through 5. This alone is sufficient grounds to delete the sentence. [Peter Thorne, United States of America]	Taken into account: The sentence has been removed.
9-448	9	9	23			"It is this fact that underlies the broad consensus behind the results presented in this report" -- we are wondering whether this statement is not too general. We suggest to revise this statement to make it clear that it really applies to the most important and general assessment results, but certainly not to all the detailed results of the assessments provided in the 14 chapters. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account: The sentence has been removed.
9-449	9	9	24	9	27	This sentence is hard to read and I think incomplete. (... and the parameterisation of ?...) [European Union]	Noted: Comment unclear and likely misplaced!
9-450	9	9	27	9	27	title of subsection could be "Improvements in models parameterizations" [Annalisa Cherchi, Italy]	Rejected: The subsection will start close to the overall section title, which is Model Improvements. We therefore see no need to mention this again the title.
9-451	9	9	27	12	17	would it be an option the revert the ordering of sections 9.1.3.1 and 9.1.3.2? Currently the section on (new) parameterizations precedes the section on new components. However, it seems to us that new "components" could (and will) also include parameterizations and thus it might be more intuitive to first discuss the components and then discuss the parameterizations, incl. those used in these new components. [Thomas Stocker/ WGI TSU, Switzerland]	Noted: Section 9.1.3.1 refers to parametrization improvements in exisiting components. As Section 9.1.3.2 introduces new components and hence truly new parametrizations. We feel that the traditional order works slightly better and decided to keep it.
9-452	9	9	29	9	29	The "explicit representation" is not a self-evident concept; it may also deserve a place in the glossary. [Ramon de Elia, Canada]	Rejected: Box 9.1 in this section explains this concept so we feel no need for a gloassry entry.
9-453	9	9	33	9	33	Parameterization of radiation processes (radiation schemes) is not explained. If this is treated in a different chapter, please provide the reference, if not please comment on advances on this issue as well. [Government of Germany]	There have been no major advances since AR4 in the parametrization of radiation per se. Much of the effort has been on improving the interaction of clouds with radiation and the inclusion of additional gases and aerosol species in the parametrization schemes. The latter is discussed in Chapter 7. We have added "c;oud-radiation" interactions in out pointer to Chapter 7.
9-454	9	9	33			Missing in 9.1.3.1.1 is a reference to aerosols and clouds: could simply mention it and refer to appropriate	Taken into account: this has been added to the

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						section in chapter 7. Cloud:7.2.3, aerosols 7.3.2, or mention section further down in this chapter (new component?) [Andrew Gettelman, United States of America]	sentence referring to Chapter 7.
9-455	9	9	36	9	37	"Advances made in cloud processes are described in Chapter 7". Chapter 7 then refers to chapter 9 in a similar way and skips any idea of judgement. This is weak. In particular if there has been so many advances in understanding clouds, as detailed in Chapter 7, why is the spread in climate sensitivities still so large? This is tackled a bit later in the chapter, when discussing figure 9.44, but the chapter needs more, both for the sake of honesty and for better decisions on research funding. [David Webb, United Kingdom]	Noted: Section 9.1.3.1 and Section 7.2.3 in Chapter 7 discuss advances in the parametrisations themselves, with no reference to the results. This is deliberate as it separates the progress in process representation from the progress in the overall model results. We evaluate the simulation of clouds in Section 9.4.1.
9-456	9	9	37			It might help if you could include a sentence or two about the improvements in clouds and convection here, even though the detail is in another chapter. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Rejected: As space is limited we feel discussing these improvements in a single section (7.2.3) and referring to it from here is sufficient.
9-457	9	9	43	9	43	Does any CMIP5 model use EDMF? If not, it is unclear to me why this development is discussed. If so, it would be good to explicitly cite which GCM uses it. [Timothy Merlis, United States of America]	Taken into account: The IPSL model is at least one model using it. We feel listing particular models is inappropriate in this section, as it would lead to long an unwieldy model list. Table 9.1 is providing this information. We have added a sentence clarifying this to the start of this section.
9-458	9	9	43	9	47	The EDMF approach was developed independently in two teams in the Netherlands and in France. Only the second team is cited here, while similar ideas are at work in the LMDZ GCM for quite a long time (Hourdin, F., F. Couvreux and L. Menut, 2002, Parameterisation of the dry convective boundary layer based on a mass flux representation of thermals, J. Atmos. Sci., 59:1105-1123, http://www.lmd.jussieu.fr/~hourdin/PUBLIS/thermique.ps.gz , see also C. Rio and F. Hourdin, 2008, A thermal plume model for the convective boundary layer: Representation of cumulus clouds, J. Atmos. Sci. 65:407—425, http://www.lmd.jussieu.fr/~hourdin/PUBLIS/Rio2008.pdf or Rio, C., F. Hourdin, F. Couvreux and A. Jam, 2010, Resolved versus parametrized boundary-layer plumes. Part II: Continuous Formulations of Mixing Rates for Mass-Flux Schemes, BLM, 135: 469-483, http://www.lmd.jussieu.fr/~hourdin/PUBLIS/Rio_etal_blm_2010.pdf). Note that this idea of combining eddy diffusion with a mass flux scheme for the convective boundary layer was first proposed in a paper by Chatfield and Brost, 1987, A two-stream model of the vertical transport of trace species in the convective boundary layer, in JGR and rediscovered in our two teams at the end of the 90s. The "thermal plume model" developed in our team, is used in the IPSLCM5B simulations made available on the CMIP5 database (Hourdin et al, 2012, LMDZ5B...). [Frédéric HOURDIN, France]	Taken into account: The recent references to the LMD development effort have been added.
9-459	9	9	43			Better to say the EDMF approach, like the shallow scheme.... (Otherwise it seems to imply a preference for EDMF). [Andrew Gettelman, United States of America]	Accepted
9-460	9	9	44			"like the shallow cumulus scheme of (Park and Bretherton, 2009)": We recommend that the authors edit the phrase to read: "like the shallow cumulus scheme of the University of Washington (Park and Bretherton, 2009)". [Government of United States of America]	Accepted
9-461	9	9	45	9	45	scheme of Park and Bretherton (2009), ... [Hai Lin, Canada]	Editorial
9-462	9	9	45	9	45	Change "(Park and Bretherton, 2009), to "Park and Bretherton (2009)" [SAMIR POKHREL, INDIA]	Editorial
9-463	9	9	51	9	57	These terrain generated features are generally much better handled by the model itself when the horizontal resolution is increased, (see e.g. Jung et al, J Clim 2012) I think that should be mentioned here. [Gunilla Svensson, Sweden]	Taken into account - A sentence to this effect has been added and the reference included.
9-464	9	9	53	9	53	The "recent" parametrizations are not that recent. [Ramon de Elia, Canada]	Taken into account: These old citations have been removed and only progress since AR4 is reported.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-465	9	9	53	9	53	(e.g., Palmer et al., 1986; McFarlane, 1987). [Hai Lin, Canada]	Editorial
9-466	9	9	53	9	53	Change "(e.g., (Palmer et al., 1986); (McFarlane, 1987))" to "(e.g., Palmer et al., 1986; McFarlane, 1987)" [SAMIR POKHREL, INDIA]	Editorial
9-467	9	9	53	9	57	"more recent": the reported reference are ~10 or 15 years old. Were these parameterizations not already included in the CMIP3 models? How could these not so recent efforts be aided by 2010 observations? I recommend to fully revise this discussion. [Elisa Manzini, Germany]	Taken into account: The paragraph has been shortened and only progress since AR4 is reported.
9-468	9	9	55	9	55	Include Rotstayn et al. (2012). Rotstayn, L. D., Jeffrey, S.J., Collier, M.A., Dravitzki, S.M., Hirst, A.C., Syktus, J.I., Wong, K.K., 2012: Aerosol- and 49 greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: a 50 study using single-forcing climate simulations. Atmos. Chem. Phys., 12, 6377-6404. [Mark Collier, Australia]	Noted: This comment appears out of place
9-469	9	9	55	9	56	(e.g., Lott and Miller, 1997; Gregory et al., 1998; Scinocca and McFarlane, 2000). [Hai Lin, Canada]	Editorial
9-470	9	9	55	9	56	Change "(e.g., (Lott and Miller, 1997); (Gregory et al., 1998); (Scinocca and McFarlane, 2000))" to "(e.g., Lott and Miller, 1997; Gregory et al., 1998; Scinocca and McFarlane, 2000)" [SAMIR POKHREL, INDIA]	Editorial
9-471	9	9	55			Although the link between climate models and NWP models is made later in the chapter, it could be introduced earlier. Development of parameterization has come from both the Climate and the NWP communities. For example, in the line in question, the referenced work of Lott and Miller(1997) was carried out at an NWP centre, and the testing was carried out through weather forecasting experiments. Further comments below also relate to the suggestion that a little more could be made of the role of NWP in the development of parametrization and in demonstration of the benefits of increased resolution. Moreover, as climate-model assessment makes use of reanalysis, improvement of parameterization in those NWP models that are also used for reanalysis brings improvements to reanalysis. [Adrian Simmons, United Kingdom]	Noted: We discuss this issue in a later section. We have added a reference to NWP models and resolution dependence of orographic wave representations.
9-472	9	9				Check how citations appear in the main text. A lot of parenthesis. [Massimo Bollasina, Italy]	Editorial
9-473	9	10	2	10	3	The statement claiming that "non-orographic gravity waves is becoming a common feature" is not substantiated. Some references are needed, I only know a few (Sassi and Garcia, JAS, 2010; Geller et al J Clim, 2011) [Gunilla Svensson, Sweden]	Taken into account: The two paragraphs have been rewritten and the references post AR4 have been added.
9-474	9	10	2	10	8	This section contains only "old" references. There are many manuscripts which have been published recently and which should be mentioned here; for example you could find many of these papers in the reference lists of the review paper of Alexander et al. (2010, Q. J. R. Meteorol. Soc. 136: 1103–1124) and the Geller and Gong (2010, JGR) paper. [Martin Dameris, Germany]	Taken into account: The two paragraphs have been rewritten and the references post AR4 have been added.
9-475	9	10	2	10	8	"common": Is this also an "advancement"? How is it judged, this development? Any reference? The mean middle atmosphere circulation is dynamically driven and radiatively damped (Holton 1983). Not sure what is meant with "GCM" here, ambiguous. Maybe "AOGCM" or ESM, is meant? Note that atmosphere general circulation models (with specified SST) that included the middle atmosphere included non-orographic GWD since the mid 1990s. "non-orographic" means that the GW are not of orographic origin (flow over a mountain). [Elisa Manzini, Germany]	Taken into account: The two paragraphs have been rewritten and the references post AR4 have been added.
9-476	9	10	4	10	5	Change "(primarily non-orographic) GWD (Holton, 1983)" to "primarily non-orographic GWD (Holton, 1983)" [SAMIR POKHREL, INDIA]	Editorial
9-477	9	10	8			an unusual citation for the Brewer-Dobson circulation [Peter Braesicke, United Kingdom]	Taken into account: The two paragraphs have been rewritten and the reference has been removed.
9-478	9	10	13	10	13	"admit" is a jargon word here – please explain. [Martin Juckes, United Kingdom]	Taken into account: text clarified
9-479	9	10	13	10	13	The reference of Redi, 1982 is misplaced. [Farahnaz Khosrawi, Sweden]	Editorial
9-480	9	10	13	10	15	change "the (Redi, 1982) neutral diffusion and (Gent and McWilliams, 1990) eddy advection parameterisation (see also (Gent et al., 1995; McDougall and McIntosh, 2001))" to "the Redi (1982) neutral diffusion and Gent	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						and McWilliams (1990) eddy advection parameterisation (see also Gent et al., 1995; McDougall and McIntosh, 2001) [SAMIR POKHREL, INDIA]	
9-481	9	10	15	10	15	“Since AR4, the main focus”: focus of what? [Martin Jukes, United Kingdom]	Taken into account: text clarified
9-482	9	10	17	10	18	change “(Gnanadesikan et al., 2007) (Ferrari et al., 2008; Ferrari et al., 2010) and (Danabasoglu et al., 2008)” to “(Gnanadesikan et al., 2007; Ferrari et al., 2008; Ferrari et al., 2010 and Danabasoglu et al., 2008)” [SAMIR POKHREL, INDIA]	Editorial
9-483	9	10	25	10	26	change “((Boccaletti et al., 2007), (Fox-Kemper et al., 2008), and (Klein and Lapeyre, 2009)), and the parameterisation of (Fox-Kemper et al., 2011)” to “(Boccaletti et al., 2007; Fox-Kemper et al., 2008 and Klein and Lapeyre, 2009) and the parameterisation of Fox-Kemper et al. (2011)” [SAMIR POKHREL, INDIA]	Editorial
9-484	9	10	26	10	26	Please change “(Fox-Kemper et al., 2011)” to “Fox-Kemper et al., (2011)” [HASIBUR RAHAMAN, India]	Editorial
9-485	9	10	27	10	27	Could add references “... models (e.g., Danabasoglu et al., 2012; Bi et al., 2012a).” [Anthony Hirst, Australia]	Taken into account. References added
9-486	9	10	28	10	28	Please add the following paragraph: It is noticed that surface waves can much improve the simulation of the upper ocean through vertical mixing process (Qiao et al, 2004), and then climate system. The too cold tongue, as one of the tropical biases, is reduced by considering surface wave-induced mixing (Song et al, 2012), and surface wave can modulate the water vapor transport (Song et al, 2012). References: Qiao, F., Y. Yuan, Y. Yang, Q. Zheng, C. Xia, and J. Ma (2004), Wave-induced mixing in the upper ocean: Distribution and application to a global ocean circulation model, Geophys. Res. Lett., 31, L11303, doi:10.1029/2004GL019824. Song, Y., F. Qiao and Z. Song (2012), Improved Simulation of the South Asian Summer Monsoon in a Coupled GCM with a More Realistic Ocean Mixed Layer. Journal of Atmospheric Sciences, 69, 1681–1690, DOI: http://dx.doi.org/10.1175/JAS-D-11-0235.1. Song, Z., F. Qiao, and Y. Song (2012), Response of the equatorial basin-wide SST to non-breaking surface wave-induced mixing in a climate model: An amendment to tropical bias, J. Geophys. Res., 117, C00J26, doi:10.1029/2012JC007931. [Bin Wang, China]	Rejected. This section is about documented ocean component improvement in CMIP5 models not about ocean models improvements in general.
9-487	9	10	30	10	32	Emphasize that parameterizations such as Simmons et al, 2004 are a considerable improvement over past parameterizations used in AR4 because they are energetically consistent. [Sonya Legg, United States of America]	Taken into account: text modified to reflect comment
9-488	9	10	31	10	31	(MacKinnon et al., 2009) appears a conference proceedings. Ok? [Anthony Hirst, Australia]	Taken into account. Reference verified
9-489	9	10	32	10	32	Could add references: “... now used in several climate models (e.g., Jayne, 2009; Danabasoglu et al., 2012; Bi et al., 2012) [Anthony Hirst, Australia]	Taken into account. References added
9-490	9	10	32	10	32	Delete the spurious parenthesis before “Jayne” [Sonya Legg, United States of America]	Editorial
9-491	9	10	33	10	33	“transport of dense water down-slope with gravity currents” - change “with” to “within” or “by”. [Sonya Legg, United States of America]	Editorial
9-492	9	10	33	10	33	Delete the suprious parenthesis before “Legg” [Sonya Legg, United States of America]	Editorial
9-493	9	10	34	10	34	“making their way” seems like casual language to use in the report. [Timothy Merlis, United States of America]	Editorial
9-494	9	10	37	10	44	1. This paragraph reads that NPZD-type models are lacking ecosystem dynamics. This is false as lower trophic level dynamics are explicitly simulated. 2. Half of CMIP-5 OBGC models are NPZD-type. Then what about the other half and their differences between NPZD and non-NPZD models? It would be important to	Taken into account: text clarified

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						separate the model complexity and its ability to reproduce observed tracer fields. [Takamitsu Ito, United States of America]	
9-495	9	10	37	10	54	This subsection seems to fit better into section 9.1.3.2. [Government of Germany]	Taken into account: text moved to 9.1.3.2
9-496	9	10	39	10	39	... About half of the CMIP5 OBGC models... [European Union]	Editorial
9-497	9	10	39	10	39	“half of CMIP5” – missing space. [Martin Juckes, United Kingdom]	Editorial
9-498	9	10	39	10	39	Pls give a gap after CMIP5 [HASIBUR RAHAMAN, India]	Editorial
9-499	9	10	39			... ofCMIP5 (typo) [Government of France]	Editorial
9-500	9	10	40	10	40	What are NZP-models? The abbreviation has not been introduced yet. [Farahnaz Khosrawi, Sweden]	Taken into account: text clarified
9-501	9	10	40			explain abbreviation "NPZD" [Barbara Früh, Germany]	Taken into account: text clarified
9-502	9	10	43	10	43	I would prefer PFT to PFP but maybe the abbreviation is not necessary as it is never used. [James Christian, Canada]	Taken into account: text clarified
9-503	9	10	43	10	43	PFTs instead of PFPs? [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Taken into account: text clarified
9-504	9	10	43	10	43	PFP's are introduced as an abbreviation for plankton functional ... - but is not used past this location. Is this really needed? [Government of United States of America]	Taken into account: text clarified
9-505	9	10	43	10	43	What is the abbreviation PFP exactly standing for? Further, it should be placed somewhat earlier in the sentence. [Farahnaz Khosrawi, Sweden]	Taken into account: text clarified
9-506	9	10	43	10	43	PFPs acronyms are not defined earlier, also pls change "(PFPs; (Le Quere et al., 2005)" to "(PFPs; Le Quere et al., 2005)" [HASIBUR RAHAMAN, India]	Taken into account: text clarified
9-507	9	10	46	10	46	Is the sentence correct? It seems there is something missing. [Farahnaz Khosrawi, Sweden]	Editorial
9-508	9	10	46	10	49	Split the sentence into two sentences. [Farahnaz Khosrawi, Sweden]	Editorial
9-509	9	10	51	10	52	"On centennial scales, deep sea carbonate sediments neutralize atmospheric CO2." This reads poorly. Could be construed to read as if all anthropogenic CO2 will be "neutralised" in a few centuries. This is not the case (e.g., Ridgwell and Hargreaves, 2007, doi:10.1029/2006GB002764). I suggest change "centennial scales" to "multi-millennial scales", or otherwise reword. [Anthony Hirst, Australia]	Taken into account: text clarified
9-510	9	10	52	10	53	I am not clear what a "growing" number of models means in this context, since the models have been frozen for several years. I am skeptical that a significant number of CMIP5 models include an interactive sediment module, although some ocean models used in CMIP5 do have versions that include such a module. I could not find any reference to data fields arising from such a module (e.g., dissolution or remobilization of DIC from sediments) in the CMIP5 data request. [James Christian, Canada]	Taken into account: text clarified
9-511	9	10				Neither section 9.1.3.2.8 (land-ice) nor 9.1.3.1.4 (sea-ice) mentions new parameterizations of ice-bergs which are included in several CMIP5 models. See Martin, T, and Alistair Adcroft, July 2010: Parameterizing the fresh-water flux from land ice to ocean with interactive icebergs in a coupled climate model. Ocean Modelling, 34(3-4), doi:10.1016/j.ocemod.2010.05.001. [Sonya Legg, United States of America]	Taken into account: text modified to reflect comment
9-512	9	11	4	11	7	These two examples are too specific, also with respect to other sections [Massimo Bollasina, Italy]	Rejected. We feel that giving published examples to back-up statements is very important.
9-513	9	11	6	11	6	“60% of summer temperature variability” – clarify what aspect of variability this refers to. [Martin Juckes, United Kingdom]	Accepted and taken into account - text clarified.
9-514	9	11	6			I suggest additional references for the European heat wave: Schär et al, 2004, doi:10.1038/nature02300 and Zaitchik et al., 2006, doi:10.1002/joc.1280 [Alexander Loew, Germany]	Rejected. These references are prior to the previous IPCC report and are not required here to back-up the points we make.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-515	9	11	9	11	27	These paragraphs contain a number of very old references (much before 2007). Suggest it is shortened and reworded to avoid their use [Government of United Kingdom of Great Britain & Northern Ireland]	Agreed. This text has been updated.
9-516	9	11	16	11	16	also mention that agricultural land management practices, different crops, seeding and harvest times, irrigation etc., are currently not included but may have profound climate effects [European Union]	Accepted and taken into account - land-use change now mentioned in revised text.
9-517	9	11	21	11	23	Skip the line break and make out of these two paragraphs one paragraph. [Farahnaz Khosrawi, Sweden]	Rejected. Separate paragraphs are required here for clarity.
9-518	9	11	23	11	24	Models can be tested offline, but this removes two-way interactions [Massimo Bollasina, Italy]	Yes, but this means that you can test land-surface models with "perfect" driving data.
9-519	9	11	23	11	27	Validation of land surface models is not easy neither. I would therefore say that this statement is only partly true, as Fluxnet only allows validation at the point scale. At grid scale with mixture of different vegetation classes, satellite data provides only source; often tiling approach is applied in models which require effective parameters at model grid scale (e.g. soil hydraulic parameters) which are simply not known. I recommend to rephrase the paragraph. [Alexander Loew, Germany]	Noted. This statement has been removed.
9-520	9	11	23	11	27	the offline evaluation of land models does not allow interaction with the atmosphere. Tuning or assessing model variables or parameters may be suffering from compensating errors, and other values would have been found to give optimal model performance when the interaction with the atmosphere is allowed. Therefore "more straightforward" is maybe not a fortunate phrase [Bart Van den Hurk, Netherlands]	See response to 9-519.
9-521	9	11	23			I don't think tesing a land component model is any easier or more stratighforward than an ocean model in an offline mode - at least not from a logistics point of view. If you at least argue that there is less feedback from the land locally to the amtosphere and thus it is maybe more sensible than say for the ocean where the latter has a strong impact on the atmosphere (e.g. tropics) then a sentence like this might make sense but it has to be said as such. [duane waliser, United States of America]	See response to 9-519.
9-522	9	11	24	11	24	What is meant with "offline mode"? What is in this case done? [Farahnaz Khosrawi, Sweden]	See response to 9-519.
9-523	9	11	24	11	24	Doesn't one need measured data for the evaluation? [Farahnaz Khosrawi, Sweden]	Yes, and we make the point that flux data is available (line 25/26).
9-524	9	11	29	9	39	There is no discussion of the importance or modelling of melt ponds, crucial to early summer melt [Government of United Kingdom of Great Britain & Northern Ireland]	Taken into account. Brief discussion added.
9-525	9	11	29	11	48	The statement that progress has "apparently slowed" is surprising and inappropriately subjective. E.g. sea-ice representation in the UKMO HadGEM model now has 5 categories of sea-ice, a more physically based advection scheme and representation of basal melting. [Martin Juckes, United Kingdom]	Taken into account. Text modified.
9-526	9	11	30	11	48	Some models fail to simulate some aspects of sea-ice dynamics, such as the coupling between ice state and ice velocity (Rampal et al., 2011, JGR, 116, C00D07) [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Noted, but this paper describes only CMIP3 models.
9-527	9	11	41	11	48	This paragraph is odd and confusingly written - its title is sea ice dynamics and yet the discussion starts with talk of thermodynamics - not good enough to just talk of mechanical deformation in lieu of the word dynamics [Government of United Kingdom of Great Britain & Northern Ireland]	Accepted. Text revised.
9-528	9	11	44	11	44	"thin ice melts more quickly than thicker ice". Are the authors talking about seasonal melting, or interannual changes? On the century time scale, changes in ice thickness are larger for initially thick ice (Bitz, 2008; Holland et al., 2008). As is now, the sentence is therefore ambiguous. [François Massonnet, Belgium]	Noted, but the text was removed because of space limits
9-529	9	11	55	11	55	change "(e.g., (Losch et al., 2011))" to "(e.g., Losch et al., 2011)" [SAMIR POKHREL, INDIA]	Editorial
9-530	9	12	1	12	1	change "(e.g., (Pedersen et al., 2009; Vancoppenolle et al., 2009b).)" to "(e.g., Pedersen et al., 2009; Vancoppenolle et al., 2009b)." [SAMIR POKHREL, INDIA]	Editorial
9-531	9	12	1	12	1	remove extra "." after Vancoppenolle et al., 2009b) [HASIBUR RAHAMAN, India]	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-532	9	12	6	12	6	biogeophysical' written twice [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Rejected here. Wrong place.Relates to p.9-13 lines 18-19 Para 9.1.3.2.4 Dynamic global vegetation models and wildland fires
9-533	9	12	7	12	8	change "(e.g., (Bitz and Lipscomb, 1999))" to "(e.g., Bitz and Lipscomb, 1999)" [SAMIR POKHREL, INDIA]	Editorial
9-534	9	12	13	12	13	By 'better treatments', do you mean more realistic, or more complex? [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Rejected here. Wrong place.Relates to p.9-13 line 26. Para 9.1.3.2.4 Dynamic global vegetation models and wildland fires
9-535	9	12	15	12	17	This sentence is not well constructed and is difficult to read. Suggest re-phrasing. [Government of Australia]	Accepted. Text modified.
9-536	9	12	17	12	17	I would add "positive" before "strong feedbacks". Indeed, negative feedbacks damp the errors that the authors are mentioning. [François Massonnet, Belgium]	Taken into account. Text modified.
9-537	9	12	21	12	33	In the context of the carbon cycle it should be mentioned that it will be strictly necessary to also include the N cycle. A carbon cycle model without N (and P) can hardly provide realistic outputs with regard to ecosystem processes, specifically if used for prognosis. [European Union]	Accepted -- importance of N cycle for complete interactions with ecosystem processes has been added.
9-538	9	12	21	14	24	In section 9.1.3.2. there is little discussion of recent advances towards including a nitrogen cycle in ESMs [Government of United Kingdom of Great Britain & Northern Ireland]	Noted -- progress towards N cycles will be included if there are CMIP5 models with prototype N cycles.
9-539	9	12	23	12	23	Please add "earth's" before climate system [HASIBUR RAHAMAN, India]	Editorial
9-540	9	12	24	12	24	The species have not been introduced yet. [Farahnaz Khosrawi, Sweden]	Accepted -- species are enumerated.
9-541	9	12	35	12	45	"aerosols" is wrong, it is either "aerosol" or "aerosol particles"; aerosol is a suspension of fine solid particles or liquid droplets in a gas (see e.g., http://en.wikipedia.org/wiki/Aerosol) [Barbara Früh, Germany]	Editorial
9-542	9	12	35	12	47	Section 9.1.3.2.2 : advances in the representation of aerosols are assessed in depth in chapter 7. I would suggest to move to chapter 7 elements of this paragraph which are not already discussed in chapter 7. [Sandrine BONY, France]	Noted -- the focus here is on CMIP5 models, not on their process representations in benchmark model calculations. Overlap with Chapter 7 regarding CMIP5 models has been minimized -- section 7.3.2 discusses the treatments of wet/dry deposition and the vertical profiles of aerosol extinction, topics complementary to those discussed here..
9-543	9	12	36	12	36	AOGCMs as well as ESMs. [Massimo Bollasina, Italy]	Taken into account -- combined with comment 9-544.
9-544	9	12	36	12	40	This paragraph gives the impression that the AR5 models that include a better representation of aerosols are all ESMs whereas it is not the case for all the modeling centers (e.g GFDL CM3 has the indirect effect but none of the GFDL ESM has it).This could be reformulated to avoid any ambiguity. [RYM MSADEK, United States of America]	Accepted -- improvements in both AOGCMs and ESMs are now noted.
9-545	9	12	40			This most recent reference on the topic should be added to (or replace) Bauer et al. 2008b: Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., Hess, P., Mahowald, N., Collins, W., Iacono, M. J., Bretherton, C. S., Flanner, M. G., and Mitchell, D.: Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5, Geosci. Model Dev., 5, 709-739, doi:10.5194/gmd-5-709-2012, 2012. [Government of United States of America]	Accepted -- reference has been updated to Liu et al 2012.
9-546	9	12	42	12	42	Strictly, Rotstayn et al. (2010) only raised the hypothesis that interactive aerosols (in particular, dust) improved the simulation of interannual variability, since other things had changed in the model as well as adding aerosols. The follow-up study (Rotstayn et al., 2011, already in the reference list) had a clean comparison of simulations with and without the Australian dust source, and confirmed the hypothesis. So the reference should be amended to (Rotstayn et al., 2010, 2011). [Leon Rotstayn, Australia]	Accepted -- reference has been corrected to latter Rotstayn paper.
9-547	9	12	42	12	44	We feel that it should clearly state here that, in contrast to AR4, most AR5 models have include, at least some,	Accepted -- the addition of aerosol-cloud interaction

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						aerosol-cloud interaction schemes. However, it is still in the development stage, and the improvement of the models due to addition of these schemes has yet to be tested. [Government of United States of America]	schemes to most AR5 models has been clarified.
9-548	9	12	49	12	54	Suggest to include the EMIC study by Stocker et al, Nature CC in revision as also cited in chapter 6 that considers CH4, N2O emissions from land and includes permafrost and peat representations. [Fortunat Joos, Switzerland]	Accepted -- reference to Stocker study added.
9-549	9	12	49			Sec 9.1.3.2.3: as far as I know none of the CMIP5 ESMs have interactive methane emissions that respond to climate. Some have atmospheric methane chemistry, and I know HadGEM2 has a diagnostic of wetland emissions, but not coupled to the chemistry. This paragraph sounds like interactive methane feedbacks are becoming common – but they are not yet [Chris Jones, United Kingdom of Great Britain & Northern Ireland]	Accepted -- Statements re methane feedbacks have been corrected as suggested.
9-550	9	12	50	12	50	Before the abbreviations CH4 and CO2 were used and now it is written carbon dioxide instead of CO2. [Farahnaz Khosrawi, Sweden]	Editorial
9-551	9	12	52	12	54	change "(Khvorostyanov et al., 2008a); (Khvorostyanov et al., 2008b), and in some cases this is integrated with the representation of terrestrial and oceanic methane cycles (Volodin et al., 2010); (Volodin, 2008b)." to "(Khvorostyanov et al., 2008a; Khvorostyanov et al., 2008b), and in some cases this is integrated with the representation of terrestrial and oceanic methane cycles (Volodin et al., 2010; Volodin, 2008b)." [SAMIR POKHREL, INDIA]	Editorial
9-552	9	12	54			(Volodin, 2008b) missing in references [Barbara Früh, Germany]	Editorial
9-553	9	13	1	13	23	but these DGVM can so far not properly handle agriculture or managed forests. Both would be necessary to get a more realistic picture of climate change effects on ecosystems and biosphere feedbacks. [European Union]	Accepted -- omission of agriculture and managed forests from current DGVMs noted in text.
9-554	9	13	1	13	23	Important discussion of fires that should be emphasised, as the information is important for policy makers. A brief mention of how fires are projected to change (or the risk of fires) should be included in the SPM and TS. [European Union]	Rejected -- this is a subject for chapter 12.
9-555	9	13	1	13	23	Forest age is a new evaluation criterion for ESMs. See Sentman et al. 2011 and Shevliakova et al. 2009. [Ronald Stouffer, United States of America]	Accepted -- use of forest age as an evaluation criterion is now noted in the text.
9-556	9	13	6	13	6	... introduce biogeophysical feedbacks... [European Union]	Editorial
9-557	9	13	6			please check "introduce biogeophysical and biogeophysical feedbacks" [Barbara Früh, Germany]	Editorial
9-558	9	13	21	13	22	"Since the frequency..." Please provide a reference with this statement. Also what else drives wildfire risk besides temperature? [Government of United States of America]	Accepted -- references to papers from SIO re fire frequency sensitivity to temperature have been added.
9-559	9	13	21			What is the reference for fires frequency increasing with temperature? [Ronald Stouffer, United States of America]	Taken into account -- combined with comment 9-558.
9-560	9	13	22	13	23	Adding to the immediately preceding comment, it could be noted here that the effects of fires are also being implemented in the data assimilation and modelling systems used for short- and medium-range forecasting and short-term reanalysis of atmospheric composition, so there is again scope for cross-fertilization. [Adrian Simmons, United Kingdom]	Noted
9-561	9	13	22			Shevliakova et al. 2009 could be added to list of ESMs which included fire parameterizations. [Ronald Stouffer, United States of America]	Accepted -- additional reference regarding models w/fire parameterizations added.
9-562	9	13	27			when mentioning the RCPs, we suggest to refer to Chs 1 and 12 where the scenarios are introduced resp. discussed in some detail [Thomas Stocker/ WGI TSU, Switzerland]	Accepted -- references to Chapters 1 and 12 added.
9-563	9	13	35	13	35	Possibly add a paragraph what is still missing for ESM'S: e.g. inclusion of other biogeochemical cycles such N, P, improved representation of agricultural practices [European Union]	Noted
9-564	9	13	36	13	49	A section about chemistry-climate interactions is certainly necessary. But this section should provide some	Rejected - Papers to specific studies on chemistry-

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						more concrete examples (not only ozone is important!) with relevant and more specific references (e.g. lines 45-46) which are directly related to climate research. It is not sufficient to mention only the CCMVal SPARC report or the WMO assessment report which are both huge reports. [Martin Dameris, Germany]	climate interactions are cited in subsequent sections that focus on the results. Here we describe how chemistry is treated in the CMIP5 models rather than providing a broad assessment of chemistry-climate interactions that can be found in the two reports that are cited. However, we agree that examples beyond ozone should be given and have added stratospheric water vapor.
9-565	9	13	38			make a hemispheric distinction, mostly SH impact and link to comment 23 (9-10 / 8 BDC) [Peter Braesicke, United Kingdom]	Accepted and SH explicitly mentioned.
9-566	9	13	39	13	42	The given definition of a CCM is misleading, since so far most CCM are not based on climate models, i.e. are not AOGCMs with fully coupled chemistry. In most CCMs sea-surface temperatures are prescribed; in this sense most CCMs are AGCM with fully coupled chemistry. Currently, many CCM groups are further developing their models systems by coupling them to ocean models. But these AOGCM have not been used in the past, except the Canadian CCM CMAM. [Martin Dameris, Germany]	Noted - CCM definition removed since this subsection focuses on the representation of chemistry in the CMIP5 ensemble.
9-567	9	13	42	13	43	... Several stratospheric chemistry modules have been... [European Union]	Editorial
9-568	9	13	42	13	44	Different phrasing, for example: Several chemistry modules have been incorporated into climate models. Some important composition-climate feedbacks have been identified for stratospheric and tropospheric applications. [Peter Braesicke, United Kingdom]	Editorial
9-569	9	13	47	13	47	I am not sure what this statement adds to the above when rewritten [Peter Braesicke, United Kingdom]	Accepted -- sentence deleted.
9-570	9	13	49			What do you mean by "rest of the climate system"? [Martin Dameris, Germany]	Accepted -- the phrase "rest of the climate system" has been removed.
9-571	9	13	51	13	51	This title rather addresses a technical aspect of the model, the height of the top, than the purpose of the "high top", which is to model realistically the stratospheric dynamics, which is required to capture the stratosphere to troposphere coupling as analysed in observations. I think the title should be changed accordingly, e.g. Stratosphere troposphere coupling [European Union]	Accepted -- title of subsection revised as suggested.
9-572	9	13	51			Section 9.1.3.2.7: Perhaps comment that to increase the height of the top requires changes in the radiation scheme such as inclusion of greater spectral resolution in the UV. [Joanna Haigh, United Kingdom]	Noted - however, details on high-top models have not been added due to length.
9-573	9	13	55			give approx altitude [Peter Braesicke, United Kingdom]	Accepted -- altitudes added.
9-574	9	14	1	14	1	... (i.e., those with a model top below the stratopause)... This is probably a better distinction between high and low top [European Union]	Editorial
9-575	9	14	1	14	2	Manzini et al 2012 actually show that the high and low top CMIP5 sub-ensembles are affected by differences in tropospheric warming (RCP8.4 scenario) that cannot be attributed to the location of the model top (their figure 3a). This result is in agreement with the evidence that clouds processes are dominant factors in determining climate sensitivity. My judgment is that while the high and low top subdivision can be still be of use in the case of pre-industrial or historical runs (without or with relatively small climate change, Charlton-Perez et al, 2012, Reichler et al, Nature Geoscience 2012), it leads to spurious results if applied to scenario runs. Similar problems have been reported previously (McLandress et al, Atmos. Chem. Phys., 12, 2533–2540). So, caution is needed when subdividing CMIP5 model sets to search for stratospheric effects. [Elisa Manzini, Germany]	Accepted -- discussion of the interaction of hi/lo-top model formulations and stratospheric effects has been revised as suggested.
9-576	9	14	4	14	12	This section should also refer to Chapter 13 where sea level rise is discussed. [European Union]	Taken into account -- combined with comment 9-577
9-577	9	14	4			Section 9.1.3.2.8. It is useful to have a short summary section on this subject, but I would suggest that you could refer in the first sentence to 13.4.3 and 13.4.4, where projections for the Greenland and Antarctic ice-sheets are assessed. At line 8, Vizcaino 2008 is not the only relevant post-AR4 reference for coupled models; there are also Huybrechts 2011, Driesschaert 2007, Charbit 2008, Ridley 2010, Robinson 2012 (see 13.4.3.1 and 13.4.3.3). At line 10-11, I don't think this is a general conclusion of all these studies; AMOC weakening	Accepted -- references to chapter 13, and the corrections suggested for the references and summary of their findings have been introduced.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						may be a negative feedback (see 13.4.3.1). In the last sentence, it would be consistent with ch4 and ch13 to emphasise that the dynamical response of the ice sheets is currently the largest model uncertainty in ice-sheet projection (discussed extensively in 13.4.3.2 and 13.4.4.2). [Jonathan Gregory, United Kingdom]	
9-578	9	14	4			Move this subsection to p13 36, thus before section 9.1.3.2.6. [Farahnaz Khosrawi, Sweden]	Rejected -- no reason for resequencing given.
9-579	9	14	5			It's the timing that is uncertain more than the total amount which could be released. We know how much ice there is. [Reto Knutti, Switzerland]	Accepted -- amount replaced by rate of release of water from land ice.
9-580	9	14	17	14	19	Many models had both these 'improvements' already (i.e. Schmidt et al, 2006) . It is important to note that not every detail of a GCM gets written up as a specific sensitivity study. [Gavin Schmidt, United States of America]	Accepted -- reference to Schmidt added and sentence recast in terms of wider adoption.
9-581	9	14	18	14	18	remove the "(" at the beginning of this line. [Hai Lin, Canada]	Editorial
9-582	9	14	19	14	19	remove the ")" at the end of this line [Hai Lin, Canada]	Editorial
9-583	9	14	20	14	20	remove the "(" at the beginning of this line. [Hai Lin, Canada]	Editorial
9-584	9	14	24	14	24	There is an extra '(' at end of sentence. [Martin Juckes, United Kingdom]	Editorial
9-585	9	14	24	14	24	(e.g., Arora et al., 2009). [Hai Lin, Canada]	Editorial
9-586	9	14	24	14	24	change "(e.g., (Arora et al., 2009)" to "(e.g., Arora et al., 2009)" [SAMIR POKHREL, INDIA]	Editorial
9-587	9	14	24			Some models have included all of the above features (the surface current effect on wind stress, the diurnal cycle, and a chlorophyll-mediated penetration of solar radiation) ever since the CMIP3/AR4 (e.g. GFDL CM2.1 in CMIP3, and its CMIP5 progeny ESM2M, ESM2G, CM3, and CM2.5; see Delworth et al, J. Climate 2006; Gnanadesikan et al., J. Climate 2006; and Wittenberg et al., J. Climate 2006). [Andrew Wittenberg, United States of America]	Accepted -- references to GFDL models added and sentence recast in terms of wider adoption.
9-588	9	14	26	15	5	Section 9.1.3.3 : Line 41 : as for horizontal resolution, it would be nice to include some information about the typical vertical resolutions used in current models (e.g. in CMIP5). Line 47 : I would change « may lead » in « can sometimes ». Lines 1-5 : this sections lacks a discussion about the benefits and the limitations of increasing the horizontal resolution in the atmosphere (well below the resolution needed to resolve convective cloud systems). There is a rich literature about it (e.g. Hourdin et al. 2012 : Climate and sensitivity of the IPSL-CM5A coupled model: impact of the LMDZ atmospheric grid configuration). [Sandrine BONY, France]	Taken into account - the typical vertical resolution added, « may lead » changed to « can sometimes ». More discussion about resolution is found in some parts of 9.4 and 9.5 and briefly summarized in the new Box for 'Understanding model performance'.
9-589	9	14	26			This section should be merge with 9.1.3.2.7; horizontal and vertical resolution (extent) are not independent; thinking about waves they are related by a dispersion relationship [Peter Braesicke, United Kingdom]	Rejected - 9.1.3.2.7 is basically focused on representation of stratosphere.
9-590	9	14	26			Section 9.1.3.3: There is very little mention of vertical resolution, whereas the CMIP5 models in Table 9.1 clearly use a range of both horizontal and vertical resolutions, in both atmosphere and ocean. The potential benefits, and drawbacks, of this should be discussed here. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Noted - unfortunately, we had to delete the mention of vertical resolution due to space limitation, rather than elaborating it. The typical vertical resolution of CMIP5 is added.
9-591	9	14	28	14	32	Would it be useful to add in this paragraph that high-resolution models generally improve the representation of small-scale processes, which may feedback on the large-scale to improve it. Roberts et al (2009, J. Clim, 22, 2541-2556) is a good example of such improvement: high-resolution models better simulate tropical instability waves, which leads to an improvement of the mean SSTs in the tropical Pacific as well as an improvement of ENSO. [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Noted - discussion like that is found in 9.4 and briefly summarized in the new Box for 'Understanding model performance'.
9-592	9	14	28	14	32	What does this paragraph actually say? Improvement is "expected"... "generally" ... "better representation of" but only if This is an example where a scientist finds the text very basic and a policy maker finds it uninformative. [Leonard Smith, United Kingdom]	Noted - in this introductory paragraph, general wording is unavoidable.
9-593	9	14	36	14	40	These lines seem to downplay the improvement in resolution. The global climate model ensembles analysed at CSIRO (for long-term change) have a decrease in average grid square side length from 242km for CMIP3 (24 models) to 180km for CMIP5 (25 models). This relates to areas being 55% as large, and suggests simulations are at least 2.5 times as expensive, from this change alone -certainly a significant increase! [Government of Australia]	Noted - we think it is appropriate to call the change from 242km to 180km 'modest change'.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-594	9	14	38	14	38	In an atmospheric GCM when the finer horizontal resolution is used (100 km to 10 km) but not in a resolved cumulus clouds scale, however, the conventional deep convective scheme might violate small cloud fraction assumption. This has to be pointed out here. [Jui-Lin (Frank) Li, United States of America]	Noted - the suggested issue is implied by an earlier statement 'Model resolution needs to be developed in concert with parameterisations and their scaling with resolution in order to realize the expected improvements.' Unfortunately, there is no space for elaboration.
9-595	9	14	42			a large part of the CORDEX RCM runs have a horizontal grid width of about 50 km, it is resigned from a higher resolution for the benefit of a higher number of ensemble members [Barbara Früh, Germany]	Noted.
9-596	9	14	42			At least within the CORDEX project it has been decided to abandon a higher than 50 km horizontal resolution for the sake of a higher number of ensemble members [Government of Germany]	Noted.
9-597	9	14	47	14	47	And, depending how model performance is measured, the steps do not have to be positive. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted.
9-598	9	14	47	14	55	"stepwise, rather than incremental, ...": this is a meaningless distinction here: adding something clearly causes a stepwise improvement in the thing added (here in the representation of small scales) but it is not clear that there is a stepwise improvement in any meaningful measure of performance. [Martin Jukes, United Kingdom]	Noted - it is true a stepwise improvement in small scales doesn't guarantee overall improvement of simulations. But, a modest improvement of resolution doesn't cause any stepwise improvement even in small scales. So, we think the distinction here is meaningful.
9-599	9	14	52	14	52	(e.g., McClean et al., 2006b). [Hai Lin, Canada]	Editorial.
9-600	9	14	52	14	52	change "(e.g., (McClean et al., 2006b))" to "(e.g., McClean et al., 2006b)" [SAMIR POKHREL, INDIA]	Editorial.
9-601	9	14	52	14	52	Pls change "(e.g., (McClean et al., 2006b))" to "(e.g., McClean et al., 2006b)" [HASIBUR RAHAMAN, India]	Editorial.
9-602	9	14	55	14	55	Add reference to submitted paper: F.O. Bryan, P.R. Gent, R. Tomas. 2012: Can Southern Ocean Eddy Effects Be Parameterized In Climate Models?, submitted to JoC. Also add reference to the Delworth et al. 2011 paper that is already cited on page 9-22 line 12. [Stephen Griffies, United States of America]	Noted - the paper doesn't seem to be accepted in time, thus it is not cited.
9-603	9	14	55			Add reference to Delworth et al. (2012): Delworth, Thomas L., and Coauthors, 2012: Simulated Climate and Climate Change in the GFDL CM2.5 High-Resolution Coupled Climate Model. J. Climate, 25, 2755–2781. doi: http://dx.doi.org/10.1175/JCLI-D-11-00316.1 [RYM MSADEK, United States of America]	Accepted.
9-604	9	15	1	15	5	This seems extraneous since no models assessed in chapter 9 resolve convection. Convection-resolving models are discussed in Chapter 7 so you can just point to that. On the other hand, there are climatic benefits (e.g. to humidity) of properly resolving synoptic features as happens at approximately T85 or better (e.g., Sherwood et al. 2010 JGR 115: D09104), which is reached by more CMIP5 ones than CMIP3 ones. Blocking also seems to improve, as you note later in 9.5.2.2.1 – maybe point to that here as well. [Steven Sherwood, Australia]	Noted - this paragraph is indeed to point to Chapter 7, just with pointing out a parallelism between cloud resolving atmosphere and eddy resolving ocean and adding an RCM example. More discussion about resolution is found in Sections 9.4 and 9.5 and briefly summarized in the new Box for 'Understanding model performance'.
9-605	9	15	2			Please note here that even higher resolution models end up parameterizing other processes, even though some things may be explicitly resolved. [Government of United States of America]	Noted - the suggested issue is touched upon in Chapter 7. Unfortunately, we can't repeat it here due to space limitaiton.
9-606	9	15	3			Possibly mention that deep convective cloud resolving global models are unlikely to be applicable for centennial simulations in the near-future. Global cloud resolving models have, however, been applied for process studies, see Chapter 7.2.2.1.2. [Thorsten Mauritsen, Germany]	Noted - we think a general pointer to Chapter 7 is enough here.
9-607	9	15	9	15	39	Section 9.2.1 : This section might be the right place to discuss the complex issue of « How good is good enough ? » for climate models, or « What should the models simulate right to provide reliable projections ? ». Unlike in NWP, model projections can not be evaluated against observations, and the connection between model errors in the current climate and climate projections remains (with a few expectations) difficult to	Taken into account: This section has been removed and the discussion requested by the reviewer is carried out inan extended Section 9.2.3.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						establish. I would not say that « Unlike NWP models, [...], climate models are concerned with climatological distributions ». Climate models are concerned with many different things. The problem is that for many applications, we just don't know how to rank or prioritize the evaluation of these different things. [Sandrine BONY, France]	
9-608	9	15	9	19	17	A discussion of the utility of metrics for evaluating the quality of model projections is missing in this section. There is a very limited discussion in section 9.3.2.1 but it seems insufficient. [Ronald Stouffer, United States of America]	Taken into account: This section has been removed and the discussion requested by the reviewer is carried out in an extended Section 9.2.3.
9-609	9	15	9			The fundamental problem of model evaluation is that we can't evaluate by repeated verification of the actual prediction. We would have to wait for a century, and even if we did a single outcome would be insufficient. So fundamentally the model evaluation is with circumstantial evidence. We try to find things that are relevant for the prediction, but we struggle to identify those which are. We explore a climate never observed with those models, and therefore the evaluation is very difficult. I describe this in Knutti 2008 Phil Trans Roy Soc, but of course there are many other papers as well making that point. [Reto Knutti, Switzerland]	Taken into account: This section has been removed and the discussion requested by the reviewer is carried out in an extended Section 9.2.3.
9-610	9	15	9			This section discusses methods, but I miss an explicit discussion of observation uncertainties on one hand, and on the role of natural variability on the other hand. There has been a lot of work recently on how big the variability locally can be (e.g. Deser 2011 Clim Dyn, Deser 2012 Nature Climate Change, Mahlstein 2012 GRL, Mahlstein 2011 ERL, Hawkins and Sutton, etc.). While for the model this can be avoided by multiple ensembles there is only one realization of the observations. There are certain aspects of climate (e.g. extremes) where even a perfect observations is of little help to evaluate model because the noise is so big. This is even worse with trends. And it's not something that we will get rid of soon even if models get much better. Maybe some of that material is in the detection chapter, but if it is then this section should refer to it. [Reto Knutti, Switzerland]	Taken into account: This section has been removed and the discussion requested by the reviewer is carried out in an extended Section 9.2.3.
9-611	9	15	13	15	13	"perform" instead of "inform"? [Farahnaz Khosrawi, Sweden]	Editorial
9-612	9	15	13	15	13	"judgements"à"judgments" [Jui-Lin (Frank) Li, United States of America]	Editorial
9-613	9	15	16	15	25	I think there would be no harm in explicitly acknowledging "expert knowledge"; metrics provide a synthesis and make results traceable; the choice of metrics is till subjective; please clarify this and do not call performance metrics as such objective. [Peter Braesicke, United Kingdom]	Taken into account: This section has been removed and the discussion requested by the reviewer is carried out in an extended Section 9.2.3.
9-614	9	15	16	15	25	This section is about evaluation of climate models using performance metrics; the Waugh and Eyring (2008) paper mentioned in line 21 does not belong to this topic because it deals with stratosphere resolving AGCM coupled to stratospheric chemistry models. [Martin Dameris, Germany]	Rejected. The Waugh and Eyring (2008) paper is among the pioneering work on performance metrics, so should be cited here. It is also one of the first that applied performance metrics in a process-oriented model evaluation approach. This section is about techniques. These are the same for climate models and chemistry-climate models (which are climate models with interactive chemistry and often prescribed sea surface temperatures).
9-615	9	15	16	15	25	Instead of starting with "In the AR4...", this paragraph is a good place to emphasize what's new in the AR5. For example, we are now able to examine the cloud simulations at different vertical levels, and develop new model performance metric for different pressure levels, because of new observational capabilities by NASA A-Train satellites. Also, we feel it may be appropriate to add a few sentences here about model evaluation using combined/collocated measurements (e.g. clouds and water vapor, Jiang et al. 2012), this is another advantage over the AR4 because the availability of collocated A-Train measurements. [Government of United States of America]	Noted: This section has been removed and its content shifted to 9.2.3 and the new 9.2.1.1 (former 9.2.2.1). This section does not concern itself with any particular detail in model evaluation, but aims at a broad discussion. The use of new cloud observations is discussed at its appropriate place in 9.4.1.
9-616	9	15	16	15	25	This paragraph is a good place to emphasize what's new in the AR5. For example, we are now able to examine the cloud simulations at different vertical levels, and develop quantitative metrics for model performance with observational measurement uncertainty. Also, we are able to examine multiple correlative parameters (for example, clouds and water vapor) using nearly simultaneous satellite measurements enabled by A-Train. At Line 21, add a citation to Jiang et al (2012), Li et al., (2012). At Line 25, add a citation to the	Noted: This section has been removed and its content shifted to 9.2.3 and the new 9.2.1.1 (former 9.2.2.1). This section does not concern itself with any particular detail in model evaluation, but aims at a broad discussion. The use of new cloud observations is

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						paper by Su et al. (2012). [Government of United States of America]	discussed at its appropriate place in 9.4.1.
9-617	9	15	20			delete "objective" [Peter Braesicke, United Kingdom]	Noted: This section has been removed and its content shifted to 9.2.3 and the new 9.2.1.1 (former 9.2.2.1). We do not refer to the metrics as purely objective anymore.
9-618	9	15	21			Pincus et al. 2008a (and all occurrences thereafter) should become Pincus et al. 2008. All references to Pincus et al. 2008b (e.g., p. 9-16, line 32) should also be changed to Pincus et al. 2008, since only one paper exists (the references should be merged). [Government of United States of America]	Editorial
9-619	9	15	21			The paper by Jiang et al. (2012) should be cited here for providing observationally-based diagnostics and quantitative assessment of CMIP5 model performances. Reference: Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	Rejected: This section has been removed and its content shifted to 9.2.3 and the new 9.2.1.1 (former 9.2.2.1). This section does not concern itself with any particular detail in model evaluation, but aims at a broad discussion. The use of new cloud observations including the Jiang et al paper is discussed at its appropriate place in 9.4.1.
9-620	9	15	21			The Waugh and Eyring paper has severe limitations, which need to be mentioned, see: 88. Grewe, V. and R. Sausen, 2009: Comment on "Quantitative performance metrics for strato-spheric-resolving chemistry-climate models" by Waugh and Eyring (2008). Atmos. Chem. Phys. 9, 9101–9110. The Waugh and Eyring method easily can produce artefacts which are statistically not significant. Hence the method should not be abled in an IPCC report without mentioning its caveats. [Robert Sausen, Germany]	Rejected. The statistical significance of the approach is discussed in Waugh and Eyring (2008). We also use a different performance metric in the chapter, why details on the exact metric used in this study are not discussed. Caveats on performance metrics in general are part of this section.
9-621	9	15	22	15	25	This is probably the worst sentence of the entire chapter. "These metrics can also be used to explore the value of weighting projections based on model performance"? This has already been explored, and the verdict is in - the existing metrics are mostly useless as a means for weighting model projections, because they are biased by the types of data that are available and not available and by the desire to distill things into a simple enough form to make it easy to perform a multi-model comparison. Consequently it is not even fair to say that the existing metrics are objective, since they are targets of opportunity rather than the list one would come up with if one had no prior knowledge of what it is possible and not so possible to observe. This enterprise has been a clear failure if the goal is to find a way to weight predictions intelligently. The second part of the sentence is the key thing, but it is relegated to a part of a single sentence. There needs to be much more discussion of process-based metrics and the avenues we should be exploring between now and AR6. What data would we need to have confidence in the physical mechanisms most responsible for Arctic sea ice decline? Do I need to get the MJO right to get northwest US extreme precip right? Does getting the MJO right also allow me to better predict climate changes in the Asian summer monsoon? How could I tell whether increased cloud-top entrainment drying vs. a deepening PBL thickening the cloud layer is more important for climate changes in low clouds? [Anthony Del Genio, United States of America]	Taken into account: This section has been removed and the discussion requested by the reviewer is carried out in an extended Section 9.2.3.
9-622	9	15	25			The paper by Su et al. (2012) should be cited here for providing process-oriented CMIP5 model evaluation for those processes that are related to climatically important feedbacks. Reference: Su, H., J.H. Jiang, C. Zhai, V.S. Perun, J.T. Shen, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, C. Morcrette, J. Petch, M. Ringer, J. Cole, M. Mesquita, T. Iversen, J.E. Kristjansson, A. Gettelman, L. Rotstayn, S. Jeffrey, J.L. Dufresne, M. Watanabe, H. Kawai, T. Koshiro, T. Wu, E.M. Volodin, T. L'Ecuyer, J. Teixeira, and G.L. Stephens, "Diagnosis of Regime-dependent Cloud Simulation Errors in CMIP5 Models Using A-Train Satellite Observations," J. Geophys. Res., in press. [Government of United States of America]	Rejected: This section has been removed and its content shifted to 9.2.3 and the new 9.2.1.1. This section does not concern itself with any particular detail in model evaluation, but aims at a broad discussion. The use of new cloud observations including the Jiang et al paper is discussed at its appropriate place in 9.4.1.
9-623	9	15	27	15	30	The point made in this sentence is true, but it could be added that for parameterization schemes that are common to NWP and climate models, confidence in these schemes may be enhanced if they work well in predictions from NWP models that are tested against subsequent observations. [Adrian Simmons, United Kingdom]	Noted: This paragraph has been moved and rewritten in Section 9.2.3.
9-624	9	15	27	15	39	The concepts of ECVs as model benchmark should be introduced earlier; it is valid for most model validation,	Noted: This paragraph has been moved.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						not only metrics. [Peter Braesicke, United Kingdom]	
9-625	9	15	30	15	33	Please note that reanalyses can also have data homogeneity issues. [Government of United States of America]	Noted: This paragraph has been moved and rewritten in Section 9.2.3.
9-626	9	15	33	15	33	Unclear what is meant in this case with "coupled system". [Ramon de Elia, Canada]	Noted: This paragraph has been moved and rewritten in Section 9.2.3.
9-627	9	15	33	15	34	It is curious that the concept of ECVs (Essential Climate Variables) is brought up, but never discussed again in this chapter. This seems to be an important topic. How is the presentation of the results and associated discussion in this chapter influenced by this concept? [Government of United States of America]	Noted. The concept of ECVs is not introduced in Chapter 2 why it is also not applied in Chapter 9. We have removed 'ECVs'.
9-628	9	15	33	15	36	Data comparison in the frame of the SPARC Data Initiative (www.sparc-climate.org/activities/trace-gas-climatologies) and the ESA CCI (www.esa-cci.org) should be mentioned here as well. [Farahnaz Khosrawi, Sweden]	Rejected. No individual efforts are mentioned here, it is a general statement on ECVs which is not longer introduced as a concept in the chapter since it is not used in Chapter 9, see above.
9-629	9	15	34	15	34	add 'and to assess' after to develop [CLAUDIA STUBENRAUCH, France]	Noted: This paragraph has been moved and rewritten in Section 9.2.3.
9-630	9	15	34	16	51	General comment concerning the activity of the Earth Observation community: Coordinated assessment efforts have been initiated by the GEWEX Radiation Panel (now Data and Assessment Panel). Cloud and radiative flux assessments have just been finished (a water vapour assessment has just started); results of the first (intercomparing 12 cloud global cloud datasets from satellite observations) are summarized in Stubenrauch et al. 2012, accepted and available at http://climserv.ipsl.polytechnique.fr/gewexca , results of the second will also appear in Bull. Amer. Meteor. Soc.. One finding of the cloud assessment is that even if absolute values of cloud properties may differ due to different instrument (or retrieval) capability to identify thin cirrus, relative geographical and seasonal variations in the cloud properties agree in general very well and are therefore valuable for model evaluation. [CLAUDIA STUBENRAUCH, France]	Noted: This paragraph has been removed and uncertainties in cloud observations are discussed in the appropriate paragraphs in Section 9.2.1 and 9.4.1.
9-631	9	15	34			An appropriate reference for ECVs is missing here. I recommend the recent GCOS supplement as an appropriate reference [Alexander Loew, Germany]	Noted. The concept of ECVs is not introduced in Chapter 2 why it is also not applied in Chapter 9. We have removed 'ECVs'.
9-632	9	15	35			Should mention here, that a number of recent papers, e.g. Jiang et al (2012), did include observational uncertainty in the model evaluation. In fact, for quantitative model evaluation, observational uncertainty must be considered. [Government of United States of America]	Rejected: This paper uses very short data sets (in the climate context). We discuss the general issue of observation uncertainty in the revised section 9.2.3.
9-633	9	15	35			Should mention here that some recent papers, e.g. Jiang et al (2012), did include observational uncertainty in the model evaluation. [Government of United States of America]	Rejected: This paper uses very short data sets (in the climate context). We discuss the general issue of observation uncertainty in the revised section 9.2.3.
9-634	9	15	36	15	39	This sentence needs rewriting for clarity and to make clear that it is either the quantity and / or the quality of the observations that is the impediment. There are many observations in the UTLS but the issue is that 90%+ are not of sufficient quality. [Peter Thorne, United States of America]	Noted: We discuss the general issue of observation uncertainty in the revised section 9.2.3.
9-635	9	15	38			Note Microwave Limb Sounder (MLS) observations of water vapor ice clouds, temperature, ozone etc have been available for the upper troposphere / lower stratosphere region continuously for the past 9 years. But the data have been used for AR4 and AR5 model evaluations since 2005 (e.g. Li et al. 2005; 2007; Su et al 2006b; 2011; 2012; Jiang et al. 2011; 2012). References: Li, J-L., D.E. Waliser, J.H. Jiang, D.L. Wu, W.G. Read, J.W. Waters, A.M. Tompkins, L.J. Donner, J-D. Chern, W-K. Tao, R. Atlas, Y. Gu, K.N. Liou, A. Del Genio, M. Khairoutdinov, and A. Gettleman, "Comparisons of EOS MLS Cloud Ice Measurements with ECMWF analyses and GCM Simulations: Initial Results," Geophys. Res. Lett. 32, L18710, doi:10.1029/2005GL023788, 28 September 2005. Li, J-L., J.H. Jiang, D.E. Waliser, and A.M. Tompkins, "Assessing Consistency between EOS MLS and ECMWF Analyzed and Forecast Estimates of Cloud Ice," Geophys. Res. Lett. 34, L08701, doi:10.1029/2006GL029022, 2007. Su, H., D.E. Waliser, J.H. Jiang, J-L. Li, W.G. Read, J.W. Waters, and A.M. Tompkins, "Relationships of upper tropospheric water vapor, clouds and SST: MLS observations, ECMWF analyses and GCM simulations," Geophys. Res. Lett. 33, L22802,	Rejected: This section has been removed and its content shifted to 9.2.3 and the new 9.2.1.1. This section does not concern itself with any particular detail in model evaluation, but aims at a broad discussion. The use of new cloud and water vapor observations including many of the suggested papers are discussed at their appropriate place in 9.4.1.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						doi:10.1029/2006GL027582, 2006b. Su, H., J.H. Jiang, J. Teixeira, A. Gettelman, X. Huang, G. Stephens, D. Vane, and V.S. Perun, "Comparison of Regime-Sorted Tropical Cloud Profiles Observed by CloudSat with GEOS5 Analyses and Two General Circulation Model Simulations," J. Geophys. Res. 116, D0910, doi:10.1029/2010JD014971, 2011. Su, H., J.H. Jiang, C. Zhai, V.S. Perun, J.T. Shen, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, C. Morcrette, J. Petch, M. Ringer, J. Cole, M. Mesquita, T. Iversen, J.E. Kristjansson, A. Gettelman, L. Rotstayn, S. Jeffrey, J.L. Dufresne, M. Watanabe, H. Kawai, T. Koshiro, T. Wu, E.M. Volodin, T. L'Ecuyer, J. Teixeira, and G.L. Stephens, "Diagnosis of Regime-dependent Cloud Simulation Errors in CMIP5 Models Using A-Train Satellite Observations," J. Geophys. Res., in press. Jiang, J.H., H. Su, S. Pawson, H.C. Liu, W. Read, J.W. Waters, M. Santee, D.L. Wu, M. Schwartz, N. Livesey, A. Lambert, R. Fuller, and J.N. Lee, "Five-year (2004-2009) Observations of Upper Tropospheric Water Vapor and Cloud Ice from MLS and Comparisons with GEOS-5 analyses," J. Geophys. Res. 115, D15103, doi:10.1029/2009JD013256, 2010. Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	
9-636	9	15	38			Please note that the observations from Aura Microwave Limb Sounder (MLS) for the UTLS region have been available since August 2004. The measurements include temperature, water vapor, ice water content, ozone, carbon monoxide, HCL, HNO3, N2O, etc. [Government of United States of America]	Noted: since 2004 is not long-term in assessing climate models. Hence we believe our statement to be correct in the context of this chapter.
9-637	9	15	49	15	54	Reanalyses are also often used to validate climate models when observations are not available, too uncertain or too scarce. Reanalyses also have the advantage of providing homogeneous variables in time and space, which is very convenient for evaluating models. [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Taken into account: We now mention reanalyses in this sentence.
9-638	9	15	49	15	56	More references to the studies using CMIP5 model simulations can be added here, such as Jiang et al., 2012; Li et al., 2012; Tian et al., 2012; Klein et al., 2012. [Government of United States of America]	Noted: This section summarizes approaches to evaluation, not its results. The papers mentioned are included in the cloud evaluation part of Section 9.4.1.
9-639	9	15	50	15	50	maybe add after 'global distribution' -> 'global distribution and timeseries' [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Rejected: As these are merely examples, the list needn't be comprehensive.
9-640	9	15	51			A few more references need to be added here to reflect the most recent AR5/CMIP5 model evaluation efforts, such as Jiang et al. (2012) and Su et al. (2012). References: Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. Su, H., J.H. Jiang, C. Zhai, V.S. Perun, J.T. Shen, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, C. Morcrette, J. Petch, M. Ringer, J. Cole, M. Mesquita, T. Iversen, J.E. Kristjansson, A. Gettelman, L. Rotstayn, S. Jeffrey, J.L. Dufresne, M. Watanabe, H. Kawai, T. Koshiro, T. Wu, E.M. Volodin, T. L'Ecuyer, J. Teixeira, and G.L. Stephens, "Diagnosis of Regime-dependent Cloud Simulation Errors in CMIP5 Models Using A-Train Satellite Observations," J. Geophys. Res., in press. [Government of United States of America]	Rejected: These papers are specific to cloud evaluation and are assessed in the appropriate place in the chapter in section 9.4.1.
9-641	9	15	55			Here, as in many parts of the documents, it is not clear what actually an error is. Quite often I would prefer to talk about uncertainties/deviations from a (chose) reference; I appreciate that this is a subtle difference, but an important one I believe! [Peter Braesicke, United Kingdom]	Noted: The discussion of this issue has been enhanced in Section 9.2.3. The particular reference to the root mean square error has been removed here as it is not essential to the discussion.
9-642	9	15				section 9.1: One thing missing in this section is a discussion of the use and benefits of ensembles and the manner they are constructed. While I see this is discussed below - it isn't discussed in the pedagogical way these other aspects are and I think the reader could benefit from this as an "advance" in the way we are able to carry out the science of projections. [duane waliser, United States of America]	Taken into account: Ensemble techniques are now discussed in Section 9.2.2.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-643	9	16	2			Recent model evaluations emphasizing combined/collocated measurements of more than one variables should be mentioned here. This is important since for model improvement, we could not just look at or compare one single variable (e.g. cloud) with observation. We also recommend that other related variables are compared (e.g. water vapor etc that could affect the changes in cloud), e.g. Jiang et al. (2012). [Government of United States of America]	Taken into account: This issue is important but does not fit into this discussion. Instead we added it to the section on isolating processes (now 9.2.1.2), where "combined" evaluation is discussed.
9-644	9	16	4	16	24	Section 9.2.2.2 : In this section, it should also be mentioned that new variables are now considered to evaluate processes in climate models. It is the case in particular of the isotopic composition of water (precip, vapor, vegetation, etc), whose use has exploded since the AR4 owing to the arrival of new observational datasets (satellite and in-situ). As an example : Risi et al, JGR, 2012a,b : Process-evaluation of tropospheric humidity simulated by general circulation models using water vapor isotopic observations: Part 1. Comparison between models and observations ; Part 2. Using isotopic diagnostics to understand the mid and upper tropospheric moist bias in the tropics and subtropics. Note that these papers could be of interest also for section 9.4.1.1.2. [Sandrine BONY, France]	Taken into account: The issue and references have been added.
9-645	9	16	4	16		Sub grid scale variability is another important aspect to evaluate in GCMs. Several studies evaluate subgrid scale distribution (Quaas 2012 , Konsta et al., 2012) [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account: The issue and references have been added.
9-646	9	16	4	16		Quaas, J., Evaluating the "critical relative humidity" as a measure of subgrid-scale variability of humidity in general circulation model cloud cover parameterizations using satellite data, J. Geophys. Res., 117, D9, doi:10.1029/2012JD017495, 2012.Konsta D, Chepfer H, Dufresne JL 'A process oriented characterization of tropical oceanic clouds for climate model evaluation, based on a statistical analysis of daytime A-train observations', Climate Dynamics (2012) [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account: The issue and references have been added.
9-647	9	16	6			What are the "key processes"? How we measure that a model is fit for purpose should be addressed earlier (metrocs are just a tool, they do not necessarily help to understand the underlying physics!) [Peter Braesicke, United Kingdom]	Taken into account: This section discusses a specific set of approaches. The broader discussion of the issues raised is now in Section 9.2.3 and is announced at the start of Section 9.2. We have removed the word "key" as it is not relevant here.
9-648	9	16	10	16	10	consider rephrase "...results are averaged within categories that describe physically distinct regimes of the system under study ". [Jui-Lin (Frank) Li, United States of America]	Noted
9-649	9	16	10	16	13	Tsushima et al.(accepted) should also be cited here. The study is an extension of Williams and Webb (2009) of cloud regime analysis of annual mean climatology, which analyzes cloud regimes in the seasonal variation and develop a metric for the seasonal variation of cloud regimes. There are some other studies about cloud metrics/diagnostics. See comments 13-18. [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Accepted.
9-650	9	16	10	16	13	Tsushima Y, MA Ringer, MJ Webb, KD Williams, Quantitative Evaluation of the Seasonal Variations in Climate Model Cloud Regimes, Climate Dynamics (accepted). [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Accepted
9-651	9	16	10	16	13	Several studies have developed cloud metrics for the globe (Tsushima et al. accepted) and cloud diagnostics over the tropics (Ichikawa et al. 2012, Nam et al. 2012, Florent and Bony, 2012) in cloud/thermodynamic regimes. [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account - The studies have been assessed. As they discuss specific cloud evaluation issues most are included in the cloud evaluation part of Section 9.4.1 or in Chapter 7.
9-652	9	16	10	16	13	Ichikawa H, H Masunaga, Y Tsushima, H Kanzawa, Reproducibility of Climate Models of Cloud Radiative Forcing Associated with Tropical Convection, J. Clim, 25, 1247-1262, doi:http://dx.doi.org/10.1175/JCLI-D-11-00114.1 [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account - The study has been assessed and included here.
9-653	9	16	10	16	13	Nam C, S Bony, J.-L. Dufresne, H. Chepfer, The 'too few, too bright' tropical low-cloud problem in CMIP5 models, GRL, VOL. 39, L21801, 7 PP., 2012 doi:10.1029/2012GL053421 [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account - The study has been assessed. As it discusses specific cloud evaluation issues it has been included in the cloud evaluation discussion of Section 9.4.1.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-654	9	16	10	16	13	Florent S and S Bony (2012) Interpretation of the positive low-cloud feedback predicted by a climate model under global warming Climate Dynamics 10.1007/s00382-011-1279-7 [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account - The study has been assessed and included here.
9-655	9	16	11	16	13	Two recent studies have used a combination of cloud regimes and cyclone compositing to study cloud and radiation biases in the Southern Ocean. Full citations: A. Bodas-Salcedo, K. D. Williams, P. R. Field, and A. P. Lock, The surface downwelling solar radiation surplus over the Southern Ocean in the Met Office model: the role of midlatitude cyclone clouds, J. Climate, 25, 7467-7486, 2012, 10.1175/JCLI-D-11-00702.1. K. D. Williams, A. Bodas-Salcedo, M. Deque, S. Fermepein, B. Medeiros, M. Watanabe, C. Jakob, S. A. Klein, C. A. Senior, and D. L. Williamson, The transpose-AMIP II experiment and its application to the understanding of southern ocean cloud biases in climate models, J. Climate, 2012. [Alejandro Bodas-Salcedo, United Kingdom of Great Britain & Northern Ireland]	Taken into account - The studies have been assessed. The Bodas Salcedo reference is more relevant to Section 9.4.1 and has been included there. The Williams reference is more relevant to Section 9.2.1.4 and has been included there.
9-656	9	16	12	16	13	A recent study published by Tsushima et al. is another example on the use of cloud regimes. This paper could also be cited along Williams and Webb (2009) in sections 9.2.2.3, and 9.4.1.1.2. Full citation: Y. Tsushima, M. A. Ringer, M. J. Webb, and K. D. Williams, Assessment of radiative feedback in climate models using satellite observations of annual flux variation, Clim. Dyn., in press, 10.1007/s00382-012-1609-4. [Alejandro Bodas-Salcedo, United Kingdom of Great Britain & Northern Ireland]	Taken into account - The study have been assessed. As it discusses specific cloud evaluation issues it is included in the cloud evaluation part of Section 9.4.1.
9-657	9	16	13	16	15	A new study, Su et al. (2012), using the conditional sampling approach to evaluate CMIP5 cloud simulations can be added on line 13 and line 15. Su et al., 2012, Diagnosis of Regime-dependent Cloud Simulation Errors in CMIP5 Models Using "A-Train" Satellite Observations and Reanalysis Data, J. Geophys. Res., revised. [Government of United States of America]	Taken into account - The paper has been used to replace the Su 2011 reference, as the new paper refers to models used in this report.
9-658	9	16	13			In addition to Su et al (2011), the new Su et al. (2012) paper that diagnosis CMIP5 model performances at a range of dynamics and thermodynamics states should also be cited here. Reference: Su, H., J.H. Jiang, C. Zhai, V.S. Perun, J.T. Shen, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, C. Morcrette, J. Petch, M. Ringer, J. Cole, M. Mesquita, T. Iversen, J.E. Kristjansson, A. Gettelman, L. Rotstayn, S. Jeffrey, J.L. Dufresne, M. Watanabe, H. Kawai, T. Koshiro, T. Wu, E.M. Volodin, T. L'Ecuyer, J. Teixeira, and G.L. Stephens, "Diagnosis of Regime-dependent Cloud Simulation Errors in CMIP5 Models Using A-Train Satellite Observations," J. Geophys. Res., in press. [Government of United States of America]	Taken into account - The paper has been used to replace the Su 2011 reference, as the new paper refers to models used in this report.
9-659	9	16	17	16	24	This paragraph would benefit from inclusion of an example as was done in the previous paragraph. Please consider providing an example. [Government of United States of America]	Taken into account: Single Column Models of the atmosphere are now mentioned as an example.
9-660	9	16	18	16	19	Suzuki et al.(accepted) propose a Contoured Frequency by Optical Depth Diagram (CFODD), which evaluates the autoconversion timescale of warm rain processes in clouds. The method uses Cloudsat and MODIS data, and demonstrates that global satellite measurements can be used for evaluation of cloud processes. Suzuki K, GL Stephens, MD Lebsock (accepted) Aerosol effect on the warm rain formation process: Satellite observations and modeling, JGR. [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Noted - As the report is not a review, we focus on widely accepted and used techniques. The proposed approach is new and not in wide-spread use. Hence it has not been included.
9-661	9	16	20	16	21	Stephens et al.(2002) for Cloudsat and MODIS data in A-train are used for CFODD and should be cited. Stephens, G. L., et al. (2002), The CloudSat mission and the A-Train. Bull. Amer. Meteorol. Soc., 83, 1771–1790. [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Rejected: The goal of this section is to discuss advances since AR4.
9-662	9	16	32	16	35	Indeed, individual instrument simulators still contain observation biases (identified for 12 global cloud datasets by the GEWEX cloud assessment). Therefore one has to use several instrument simulators, keeping in mind the advantages and flaws of each specific simulation. [CLAUDIA STUBENRAUCH, France]	Taken into account: A statement on the limitations of simulators has been added.
9-663	9	16	32			Change Jiang et al. 2012b to Jiang et al. 2012. Note: In the reference list on Page 9-104, Line 31-35, the two Jiang 2012a, and Jiang 2012b papers are actually the same paper. The correct reference is the following: Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Briant, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-664	9	16	32			Change Jiang et al., 2012b to Jiang et al., 2012. [Government of United States of America]	Editorial
9-665	9	16	32			In addition or a more fitting ref than the Li et al. 2011a ref are the following. The latter is a particularly good reference for supporting the statement following this set of references. Li, J.-F., D. E. Waliser, C. Woods, J. Teixeira, J. Bacmeister, J. Chern, B. W. Shen, A. Tompkins, and M. Kohler, 2008: Comparisons of Satellites Liquid Water Estimates with ECMWF and GMAO Analyses, 20th Century IPCC AR4 Climate Simulations, and GCM Simulations. Geophys. Res. Lett., 35, L19710, doi:10.1029/2008GL035427. Waliser, D. E., J. F. Li, C. Woods, R. Austin, J. Bacmeister, J. Chern, A. Del Genio, J. Jiang, Z. Kuang, H. Meng, P. Minnis, S. Platnick, W.B. Rossow, G. Stephens, S. Sun-Mack, W.K. Tao, A. Tompkins, D. Vane, C. Walker, D. Wu, 2009: Cloud Ice: A Climate Model Challenge With Signs and Expectations of Progress, J. Geophys. Res.- CloudSat Special Section, 114, D00A21, doi:10.1029/2008JD010015. [duane waliser, United States of America]	Taken into account: The references have been updated
9-666	9	16	37	16	40	Apologies for continuing comments in the same vein, but similarly it could be pointed out that the "instrument-simulator" approach was developed for assimilation of satellite radiance data for NWP, where its implementation has been a significant factor in the improvement in the accuracy of forecasts over the past two decades. The benefit carries over into reanalysis. [Adrian Simmons, United Kingdom]	Noted: This is only partially true. Most simulators here were developed for clouds, an area avoided by data assimilation in NWP for a long time. We did not include any new text as we do not see the immediate relevance to the purpose of this subsection.
9-667	9	16	37	16	51	I would have this discussion in the context of the relevant science discussion to assess the physical validity of model approximations. [Peter Braesicke, United Kingdom]	Rejected: This development since AR4 is of sufficient importance to be handled in a separate part of the chapter.
9-668	9	16	37	16	51	It should be emphasized here that the current satellite simulator package is still under development. Especially, the quantitative analyses for error/uncertainty for all the current simulators are still lacking. The uncertainties due to assumptions used in the simulators may not be smaller than the standard satellite retrieval. Thus the simulator approach is currently only an alternative model evaluation tool, not a superior model evaluation tool. [Government of United States of America]	Taken into account: A sentence pointing to limitations of the simulator approach has been added.
9-669	9	16	37	16	51	We recommend that the authors consider whether it's worth mentioning that cloud simulators require knowledge of cloud particle size, which is highly uncertain. Discrepancies in the model assumptions of cloud particle size may impact quantitative assessment of model simulated cloud amounts. [Government of United States of America]	Noted: This is a highly technical point. The models' assumptions on cloud particle size (where relevant) are used in the simulator.
9-670	9	16	41	16	42	"Microphysical assumptions (which differ from model to model) can be included in the simulators, avoiding retrieval inconsistencies". It is not clear what is meant by this statement. This sentence would be more clear if it were rephrased in the following way: "To avoid inconsistencies with the retrievals, the microphysical assumptions of the simulators should ideally be made consistent with the host model." That said, we don't feel this is true in practice because the various modeling groups use the public simulator packages "as is" without the modifications that would make them more consistent with the host model's assumptions about various cloud properties. [Government of United States of America]	Rejected: This statement is untrue. The whole point of the simulator is to use each model's assumptions about microphysical properties, such as drop size distributions, when calculating the observation-equivalents. This is now common practice.
9-671	9	16	44	16	46	change "(Wyant et al., 2009), (Chen and Del Genio, 2009), (Marchand et al., 2009), (Yokohata et al., 2010a), often in conjunction with statistical techniques to separate model clouds into cloud regimes (e.g., (Field et al., 2008); (Williams and Webb, 2009; Williams and Brooks, 2008)" to "(Wyant et al., 2009; Chen and Del Genio, 2009; Marchand et al., 2009; Yokohata et al., 2010a), often in conjunction with statistical techniques to separate model clouds into cloud regimes (e.g., Field et al., 2008; Williams and Webb, 2009; Williams and Brooks, 2008)" [SAMIR POKHREL, INDIA]	Editorial
9-672	9	16	46	16	46	Pls remove extra bracket from the quoted reference [HASIBUR RAHAMAN, India]	Editorial
9-673	9	16	47			Another example is Feldman,D.R.,W.D.Collins,etal,(2011),SimulationsStudiesfortheDetectionof Changes in Broadband Albedo and Shortwave Reflectance Spectra under a Climate Change Scenario. Journal of Geophysical Research-Atmospheres (In Press). [Karen Shell, United States of America]	Noted; While interesting, this study is not concerned with model evaluation and we have not included it.
9-674	9	16	53	17	13	Section 9.2.2.4 : The IPSL model has also been evaluated through this approach (Risi et al., JGR, 2010). [Sandrine BONY, France]	This section has been removed and reference to chapter 5 is done to refer the different paleoclimate

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							modeling approaches
9-675	9	16	53	17	13	suggest to provide specific references to individual Ch5 sections in the paleoclimate studies section here [Thomas Stocker/ WGI TSU, Switzerland]	This section has been moved and merged with Section 9.4.1.4 where the reference to right sections in chapter 5 is made where needed.
9-676	9	17	1	17	2	Is this list of reconstructions for mid-Holocene and LGM complete? What about Peltier ice sheets? [Josephine Brown, Australia]	The reference to the Peltier ice-sheet is not included here, because it is not used for model evaluation in this chapter but as a climate prescribed forcing..
9-677	9	17	4	17	7	Here you could mention also efforts to reconstruct past extreme events such as strikes by strong storms. A possible general reference could be: Frappier, A, T R Knutson, K-B Liu, and K A Emanuel, 2007: Perspective: coordinating paleoclimate research on tropical cyclones with hurricane-climate theory and modelling. Tellus A, 59(4), 529-537. [Government of United States of America]	This section has been removed because of length consideration. Discussions about past climate extremes can be found in chapter 5.
9-678	9	17	7	17	7	There are a large number of other studies that also show limitations of proxy records. Expand references. [Josephine Brown, Australia]	This section has been moved and only references used for benchmarking are kept in section 9.4.1.4 where model-data comparisons is considered. A reference to chapter 5 is done where more references can be found on paleoclimate mode-data comparisons
9-679	9	17	13	17	13	There are a large number of studies on water isotopes in GCMs and other forward modelling methods. Expand references. [Josephine Brown, Australia]	This section has been removed. A larger set of references can be found in chapter 5.
9-680	9	17	14	17	16	This section is good, but some space might be saved by reducing the number of refs. E.g. 9 refs are given for the point here. The report need not be exhaustive on all points. [Government of Australia]	Noted: The comment is out of place as no references are given here.
9-681	9	17	15	17	36	I am not sure this is needed - I cannot see a use of this technique in the chapter; is it relevant (or just nice detail)? [Peter Braesicke, United Kingdom]	Noted: We have considered the removal of the section. However, significant work using this technique has been carried out since AR4 and has provided new insights. We decided to keep the subsection.
9-682	9	17	15	17	36	9.2.2.5 the section title is "Use of data assimilation and initial values technique" for global climate models, but the discussions and all cited references are essentially all about initial value technique. Data assimilation itself are not directly involved in the cited climate model process researches. [Government of United States of America]	Taken into account: The title was changed to "Initial Value Techniques"
9-683	9	17	15	17	37	this subsections should contains more details on the use & benefit of data assimilation and initial value techniques [Annalisa Cherchi, Italy]	Noted: Data assimilation techniques are not accessible to most climate models. A sentence highlighting the prospects of the technique has been added.
9-684	9	17	15	17	37	This is a very interesting section. If I understand the point correctly, it is saying that the atmospheric component of a climate model is essentially a weather prediction model, and could be used as such if the correct initialization data were available. And, as a consequence, it has been shown that systematic errors in model behaviour develop quickly in the forecast interval. Now there has been a lot of discussion about the failure of observed temperatures to go up over the past 10-12 years, and longer in the tropical troposphere. The response of modelers has been to say that these intervals are too short to decide if there are any systematic errors in models. But here you seem to be saying that with regard to some key model processes, if they are going to go wrong, it will show up quickly. In other words, a discrepancy over a relatively short interval of time would be sufficient to indicate an error. So it would be helpful if we were given some kind of guidance as to what is the time scale necessary for testing the major hypotheses embedded in climate models. For example, how many years of observations on the tropical troposphere are necessary to test whether it is represented incorrectly in models? [Ross McKittrick, Canada]	Noted: It is well known that some, but not all, systematic model errors in the atmosphere evolve in the first few days of a forecast. As these fast-developing errors are usually linked to fast parametrised processes initial value techniques are useful to investigate their cause. The statement that modellers deny the existence of systematic errors is incorrect. The studies cited in this section are all concerned with systematic errors, so are many others throughout the chapter. The connection of such errors to a model's ability to simulate climate variability and trends is not well established and is an area of active research.
9-685	9	17	17	17	36	Good points are made here, though I have suggested earlier that the link between NWP- and climate-model	Noted: The difficulty in making this link highlights the

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						development could be made much earlier in the chapter and in other examples. In addition, these paragraphs do not clearly distinguish between using a climate model simply to run short-term forecasts from an initial state produced by an NWP system, and using the model in the data assimilation process that produces the initial state. The latter process confronts the model with the raw observations in the most basic way, and the statistics of data fits and use offer another way of assessing such models. It is of course more difficult in general to test climate models in this latter way, but lessons learnt from using NWP models in data assimilation may be applicable to climate models that have similar parametrizations. And if, as in the case of the Met Office in the UK, a "unified" model framework is used for NWP and climate within a single institution, there may be a more direct link from model use in data assimilation to climate-model improvement. [Adrian Simmons, United Kingdom]	large gaps that still exist between the communities. A sentence on the prospects of data assimilation as well as the seamless or unified approach to modelling has been added.
9-686	9	17	19	17	19	... predictions and projections... [European Union]	Accepted
9-687	9	17	19	17	19	projections -> projections [Bart Van den Hurk, Netherlands]	Accepted
9-688	9	17	22			"some"? Can more detail be provided? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account: The main purpose of the technique is to test parametrizations. The word "some" was unnecessary and has been removed.
9-689	9	17	28	17	28	Specify what "very early on" means - hours? days? [Josephine Brown, Australia]	Taken into account: The wording has been changed to "within a few days"
9-690	9	17	29			Why isn't the official Transpose AMIP project that leverages off of CMIP referenced more explicitly here? This would be the most fitting examples of this methodology for the given readership. [duane waliser, United States of America]	Taken into account: There was no literature on the project at the time of the SOD. A paper has now appeared and a sentence has been added.
9-691	9	17	38	17	48	How do they differ between models? Why do it separately for RCMs? Basically all models are confronted with data to test their validity. I would make this a model independent statement. [Peter Braesicke, United Kingdom]	Noted. Indeed, both global and regional climate models are evaluated against observations, and the intention was to emphasize that there is no substantial difference, while reanalysis-driven RCM runs can be compared to observations in other ways (e.g., specific observed events and sequences of events) than GCMs running in climate mode. NOTE: This section has been moved and merged with Section 9.6.
9-692	9	17	38			This subsection points out that evaluation of RCMs "often involves simulations with global reanalyses" and further down says that "multi-decadal reanalysis-based RCM evaluation runs have become common", but in fact almost exclusively evaluations of RCMs are being done running RCMs with reanalyses data. I find that situation somewhat unfortunate, since evaluations driven by reanalyses do little to test the main purpose of the RCMs, to generate products containing added value compared to the global driving data. The "complication" mentioned in the second sentence of this paragraph, that "biases in an RCM arise from both biases in the boundary conditions and the representation of regional processes in the RCM itself" I feel is not something to happily run away from by doing evaluations against reanalyses, as the biases in the boundary conditions are precisely a feature offering an RCM a chance to demonstrate its ability to do what it is meant to do, improve upon the driver boundary data. [Fedor Mesinger, United States, and Serbia]	Noted. Much of the RCM evaluation is done with global reanalyses providing boundary conditions. However, it is also common to compare RCM-results from GCM-driven runs with observations. NOTE: This section has been moved and merged with Section 9.6.
9-693	9	17	44	17	46	It is in general not true that reanalysis-forced RCM simulation sequences (e.g. synchronicity) can be compared with observations without considering a strong loss in correlation due to internal variability. This depends on the variable and is worse for large domains without large-scale nudging. [Ramon de Elia, Canada]	Noted. Agreed. The significance of internal variability was discussed in Section 9.6. NOTE: This section has been moved and merged with Section 9.6.
9-694	9	17	46	17	46	after "some longer period" add "(e.g., Maraun et al., 2010b)". The paper has a whole section discussing these different validation approaches. The reference is already in the list, Maraun et al., Rev. Geophys. RG3003, 2010. [Douglas Maraun, Germany]	Noted. NOTE: This section has been moved and merged with Section 9.6.
9-695	9	17	48			Suggestion - you could add the following to the end of this paragraph/sentence: "...have become common (e.g. Christensen et al. 2010), and methods and tools are being developed that more specifically target RCM evaluation (e.g. Whitehall et al. 2012)." Whitehall, K., C. Matmann, D. Waliser, J. Kim, C. Goodale, A. Hart, P. Ramirez, P. Zimdars, D. Crichton, G. Jenkins, C. Jones, G. Asrar, and B. Hewitson, 2012: Building	Noted. The additional reference was considered from the assessment point of view. NOTE: This section has been moved and merged with Section 9.6.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						model evaluation and decision support capacity for CORDEX. WMO Bulletin, 61, 29-34. [duane waliser, United States of America]	
9-696	9	17	55	17	57	not clear how evaluation of individual models in an ensemble supports the evaluation of GCMs in general [Bart Van den Hurk, Netherlands]	NOTE: This section has been restructured and merged with 9.8.3.1 and the role of ensembles in model evaluation has been clarified.
9-697	9	17				It should be useful at this place to point to examples of evaluations also against observations or analyses, in which specific improvements compared to driving data have been demonstrated, section 9.6.4 (Value Added). But in addition, I suggest pointing out a demonstration of the ability of an RCM to generate improvements also in large scales (Veljovic et al., 2010; Mesinger et al., 2012), given that with large scales improved added value at small scales should be assured. It is important to stress that driving RCM experiments with reanalyses and evaluations against the driving reanalysis data is not a proper setting for identification of possible degradations of large scales, since RCMs in this experiment design are made unable to show their ability or lack thereof to add value at large scales. [Fedor Mesinger, United States, and Serbia]	Noted. These issues were discussed in 9.6, in the value added -context. NOTE: This section has been moved and merged with Section 9.6.
9-698	9	17				Reference not in the SOD: [Fedor Mesinger, United States, and Serbia]	Noted. The comment seems to belong together with the comments 9-697 and 9-699. NOTE: This section has been moved and merged with Section 9.6.
9-699	9	17				Mesinger, F., K. Veljovic, M. J. Fennessy, and E. L. Altshuler, 2012: Value added in regional climate modeling: Should one aim to improve on the large scales as well? In "Climate Change: Inferences from Paleoclimate and Regional Aspects", Ed. by A. Berger, F. Mesinger, and Dj. Sijacki, Springer, 201-214 (doi:10.1007/978-3-7091-0973-1_15). [Fedor Mesinger, United States, and Serbia]	Noted. The reference was considered from the assessment point of view. NOTE: This section has been moved and merged with Section 9.6.
9-700	9	18	6	18	8	Is it really true that MME members cannot be treated as independant? They are all different combinations of global and regional models, thus each member is unique. Please clarify. [Andreas Walter, Germany]	The discussion of model independence has been clarified in the revision. Multiple analyses have indicated that model responses are not statistically independent and this has been traced to common model components as discussed. NOTE: This section has been restructured and merged with 9.8.3.1
9-701	9	18	8	18	8	Change "Knutti, 2010" to "Knutti et al., 2010" [HASIBUR RAHAMAN, India]	NOTE: This section has been restructured and merged with 9.8.3.1. These references have been adjusted accordingly.
9-702	9	18	9	18	10	Sentence "In contrast...parameter perturbations" is unclear. Please re-phrase the sentence for clarity. [Government of United States of America]	NOTE: This section has been restructured and merged with 9.8.3.1 and this material was removed.
9-703	9	18	14	18	14	(Rougier et al., 2009a).such as EMICs (Forest et al., [Hai Lin, Canada]	NOTE: This section has been restructured and merged with 9.8.3.1 and this material was removed.
9-704	9	18	16	18	17	Please explain what is meant by this sentence. [Government of Germany]	Noted. NOTE: This section has been restructured and merged with 9.8.3.1.
9-705	9	18	17			Explain "structural uncertainty". [Reto Knutti, Switzerland]	Noted and considered in the revision. NOTE: This section has been restructured, merged with 9.8.3.1, and shortened. The discussion of structural uncertainty was removed.
9-706	9	18	19	18	19	...uncertainty, Sanderson (2012) and Sexton et al. (2012) both ... [Hai Lin, Canada]	Editorial. NOTE: This section has been restructured and merged with 9.8.3.1
9-707	9	18	19	18	19	Williamson et al. (2012) [David Sexton, United Kingdom]	Williamson et al. (2012) was not accepted in time to be included. NOTE: This section has been restructured and merged with 9.8.3.1
9-708	9	18	19	18	19	Williamson et al. (2012) (already cited below) should be added to the citations of Sanderson and Sexton et al already here. [David Sexton, United Kingdom]	Williamson et al. (2012) was not accepted in time to be included. NOTE: This section has been

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							restructured and merged with 9.8.3.1
9-709	9	18	20	18	21	It would be good to define "systematic error" somewhere. Systematic error in a simple linear system is easy to detect and correct; systematic error in a complex nonlinear system like the climate is difficult to detect, difficult to attribute to model characteristics, and difficult to correct. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Thank you for the request for clarification. NOTE: This section has been restructured, merged with 9.8.3.1, and shortened.
9-710	9	18	23	18	33	Please simplify this paragraph! What is the message? Accounting for uncertainties at present implies uncertainty for the future, but we do not know how to transfer this knowledge from past to future? [Peter Braesicke, United Kingdom]	Thank you for the request for clarification. NOTE: This section has been restructured and merged with 9.8.3.1 and has been simplified.
9-711	9	18	24			Bayesian methods for PPE were first used in EMICs (Knutti 2002 Nature, Forest 2002 Science, and subsequent work). [Reto Knutti, Switzerland]	Thank you for this request. While we agree that this is the case, this was not included due to space constraints and this qualifies as review material rather than new results. NOTE: This section has been restructured and merged with 9.8.3.1.
9-712	9	18	30	18	30	The term 'ECS' (Equilibrium Climate sensitivity) first appears here and need to be defined (currently it is defined in p71, L4) [Koichi Sakaguchi, the United States]	Noted and the definition will be appropriately placed. NOTE: This section has been restructured and merged with 9.8.3.1
9-713	9	18	30	18	30	What is ECS? [Bart Van den Hurk, Netherlands]	Noted and the definition will be appropriately placed. NOTE: This section has been restructured and merged with 9.8.3.1
9-714	9	18	30			explain abbreviation "ECS" [Barbara Früh, Germany]	Noted and the definition will be appropriately placed. NOTE: This section has been restructured and merged with 9.8.3.1
9-715	9	18	30			ECS has not been defined to this point. [Government of United States of America]	Noted and the definition will be appropriately placed. NOTE: This section has been restructured and merged with 9.8.3.1
9-716	9	18	31	18	33	Unclear sentence. [Ramon de Elia, Canada]	Noted. NOTE: This section has been restructured and merged with 9.8.3.1
9-717	9	18	31			model discrepancies instead of error? [Peter Braesicke, United Kingdom]	This term is used in the statistical literature and but the discussion was removed in the revision. NOTE: This section has been restructured and merged with 9.8.3.1
9-718	9	18	41			here, more detail might be merited: more precise, better measured, more in general boundary constraints [Peter Braesicke, United Kingdom]	This section was shortened considerably. NOTE: This section has been restructured and merged with 9.8.3.1
9-719	9	18	49	19	17	Chapter 9: Section 9.2.3. (Overall Summary of Model Evaluation Approach in this Chapter). This Section presents the contents and structure of the remaining Sections of chapter 9, with the exception of Section 9.7 (Understanding Model Performance and Climate Sensitivity). The justification for this Section 9.7 should be added. [Government of Spain]	Accepted and considered.
9-720	9	18	49	19	17	the overall summary does not refer to Section 9.7 explicitly -- suggest to add it on line p19, line 7. [Thomas Stocker/ WGI TSU, Switzerland]	Accepted and considered.
9-721	9	18	51	18	52	So if we calibrate enthusiastically to reproduce historical climate, how can we be sure that we are not over-constraining the model? (Palaeo data suggest that some models are over-stable; see for example P. Valdes 2011 Nature Geoscience 4(7) 414-416, doi:10.1038/ngeo1200) [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Rejected - we do not calibrate to reproduce all aspects of historical climate. This is explained in the box on tuning.
9-722	9	18	51	18	54	Reanalyses could also be mentioned here. [Marie-Estelle Demory, United Kingdom of Great Britain & Northern	Accepted and considered.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Ireland]	
9-723	9	18				FAQ 9.1: We felt that the FAQ currently lacks a general introduction to the types of uncertainty associated with climate modelling, and how the relative importance of these sources of uncertainty changes with lead time. It would be important to add such a discussion within this FAQ. [Thomas Stocker/ WGI TSU, Switzerland]	taken into account. Focus of this FAQ is on model performance based on comparison to observations. The phrase in the review comment about 'uncertainty changes with lead time' seems more applicable to Chapter 11.
9-724	9	19	7	19	9	I am missing a more detailed statement about the uncertainties when using performance metrics. (see page 6, line 23!) Although it is discussed in other chapters, it needs a clear statement regarding uncertainties in data sets derived from observations in this chapter. The statement given on page 15 (lines 35-36) is OK, but what does it mean for the performance metric if only one data set is available? A misleading result could be the consequence. How reliable is an estimate of model performance if only one data set is available? [Martin Dameris, Germany]	Accepted and uncertainties regarding datasets highlighted.
9-725	9	19	12	19	12	Double comma [Juan Blanco, Spain]	Editorial
9-726	9	19	12	19	12	Typo: Eliminate one comma [Chiara Cagnazzo, Italy]	Editorial
9-727	9	19	12	19	12	change "forcing,," to "forcing," [SAMIR POKHREL, INDIA]	Editorial
9-728	9	19	12	19	12	Pls remove extra "," after forcing [HASIBUR RAHAMAN, India]	Editorial
9-729	9	19	12			... forcing,,are ... (typo) [Government of France]	Editorial
9-730	9	19	15	19	15	change "(e.g., (Santer et al., 2009a)" to "(e.g., Santer et al., 2009a)" [SAMIR POKHREL, INDIA]	Editorial
9-731	9	19	15	19	15	Pls remove extra bracket [HASIBUR RAHAMAN, India]	Editorial
9-732	9	19	21	19	33	Section 9.3.1 : The role of model intercomparisons is not only to assess the ability of models to reproduce observations. This chapter focuses on this kind of evaluation, but also includes a section about climate sensitivity. So the other purposes of model intercomparisons should be discussed here. This includes the assessment of the range of model behaviours (including in climate change), the identification of robust behaviours across models (including in climate change), and the interpretation of inter-model differences (thanks to the many idealized experiments aiming at addressing specific science questions). It is all these different goals that led to the design of CMIP5. [Sandrine BONY, France]	taken into account; text revised accordingly
9-733	9	19	25	19	25	The SPARC CCMVal initiative for the Evaluation of climate simulations on the stratosphere should be mentioned here as well (www.pa.op.dlr.de/CCMVal). [Farahnaz Khosrawi, Sweden]	Taken into account., but because of space limit references to intercomparison projects were not added
9-734	9	19	30			model deficit instead of error? [Peter Braesicke, United Kingdom]	Taken into account. The sentence has been suppressed
9-735	9	19	31	19	32	Except that when features are universal, they are then dubbed "robust" and meriting confidence. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text clarified. However over confidence is also a matter of concern
9-736	9	19	32			Why not note here that once "universal" (or systematic) errors are identified they become high priority targets for model improvement efforts. [duane waliser, United States of America]	Taken into account. Text rewritten
9-737	9	19	44			The two principle requirements for climate models should be emphasized (bolded or highlighted) since otherwise the reader kind of loses it in the text. [Government of United States of America]	Editorial
9-738	9	19	47	19	49	The second should be rewritten as a "requirement" to make sense of the logic of the paragraph. [Ramon de Elia, Canada]	Editorial
9-739	9	19	47	19	49	I don't agree that "many" such relationships can be tested from the observational record. Water vapor (the Soden and Held paper), OK. Clouds and sea ice, not so much, at least at the process level. [Anthony Del Genio, United States of America]	Agreed -- "many" has been replaced by "several climatically significant".

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-740	9	19	49	19	49	Section 9.3.1 : To my knowledge, Soden and Held (2006) do not make any comparisons to the observational record (but many other papers by Soden and others do that..). [Sandrine BONY, France]	Agreed -- references to papers by Soden and Held have been dropped.
9-741	9	19	50	19	50	remove "to" end of line [Annalisa Cherchi, Italy]	Editorial
9-742	9	19	51	19	52	We suggest changing the ending of this sentence to read: "...remaining uncertainties in observations, historical forcing, the models' response to forcing, and the internal variability of the climate system (e.g., Klock et al. 2011c)." We recommend that the authors leave out oceanic heat storage...that seems too specific for this context, which is a list of general reasons for discrepancies. [Government of United States of America]	Agreed -- the sentence has been revised as suggested.
9-743	9	19	52	19	52	Correct Klocke et al., 2011c reference as per my comment (7) above. [Anthony Hirst, Australia]	Editorial
9-744	9	19				Fig. 10.5: Please explain the colors of the lines in each graph (add legends), and stretch the graphs in the vertical direction, the y axis intervals are too small. [Government of Germany]	BELONGS TO CHAPTER 10
9-745	9	20	1			I would suggest to rearrange the ordering of the subsections and put them in a more logical order as going from the top to the bottom (thus from the ocean to the atmosphere or vice versa). E.g. if one starts with the atmosphere as it is done now I would suggest to put the ocean and sea ice chapter at the end and also swap their order (first sea ice and then ocean). [Farahnaz Khosrawi, Sweden]	Rejected -- the current ordering of the sections does not detract from the sense or readability of this section.
9-746	9	20	2	20	3	For clarity, consider replacing the word "hemisphere" with "semicircle". [Government of United States of America]	Editorial
9-747	9	20	6	20	41	I find it rather odd within this section not to note several well known / recently discovered forcing issues viz. recent reduction in stratospheric water vapour, higher aerosol burdens from volcanoes in the stratosphere since 2000 and the apparent under-estimate of tropospheric aerosols in the RCP pathways since 2005. I would have thought that a paragraph to the effect would be warranted here? [Peter Thorne, United States of America]	Accepted -- This section now includes a reference to the "hiatus" box on recent global-mean temperature trends. Solar forcing and natural aerosols have been added.
9-748	9	20	8	20	14	It's not correct to associate the second phase (RCP) and third phase (ECP) terminology with previous assessments - they are specific to CMIP5 and this assessment - and similarly, the 1850-2005 and 2006-2100 time split is characteristic of CMIP5/RCPs, not protocols used in previous assessments. It might aid clarity to merge the final sentence of this paragraph with the following paragraph insofar as Annex II specifically relates to the homogenised CMIP5 time-series. [Tim Johns, United Kingdom of Great Britain & Northern Ireland]	Agreed -- reference to earlier assessments has been removed.
9-749	9	20	10	20	11	It would be good to explain why the HISTORICAL runs end at 2005 and the "future" runs begin at 2006, given that the present year is 2012, so theoretically boundary conditions could have been specified until 2011 at least. [Josephine Brown, Australia]	Noted -- this is a prescribed feature of the CMIP5 archive.
9-750	9	20	10	20	11	Rephrase the sentence as follows: "The first phase covers the start of the modern industrial period through to the present-day, thus covering the years from 1850 to 2005 (van Vuuren et al., 2011). [Farahnaz Khosrawi, Sweden]	Editorial
9-751	9	20	12			Provide some reference to another part of the report that explains what Extension Concentration Pathways are. [Government of United States of America]	Accepted -- The same references for the ECPs cited in chapter 12 (Meinshausen et al 2011 and Taylor et al. 2012) are now cited here.
9-752	9	20	12			It is not clear from this sentence what the third phase is. A reference to an earlier or later chapter where the Extension Concentration Pathways are defined should also be added. [RYM MSADEK, United States of America]	Taken into account -- combined with comment 9-751
9-753	9	20	16	20	16	The word "experiments" is repeated. Remove one. [Hai Lin, Canada]	Editorial
9-754	9	20	16	20	16	change "20th century experiments experiments" to "20th century experiments" [SAMIR POKHREL, INDIA]	Taken into account -- combined with 9-753.
9-755	9	20	16	20	16	Pls remove extra " experiments" [HASIBUR RAHAMAN, India]	Taken into account -- combined with 9-753.
9-756	9	20	30	20	30	Pls change "(Hibbard et al., 2007)" to "(Hibbard et al., 2007)." [HASIBUR RAHAMAN, India]	Editorial
9-757	9	20	43	20	52	Could drop this (small) section, as not relevant to current chapter. Check for duplication with Chapter 11.	Rejected -- this section is required to explain the left-

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						[Government of Australia]	hand panel of Figure 9.1.
9-758	9	20	48			"are comprised of" could be replaced simply by "comprise" [Adrian Simmons, United Kingdom]	Editorial
9-759	9	20	49	20	52	This describes the original layout of the CMIP5 decadal hindcasts. As this research is still young, it was recognized only later that the 5-yearly intervals were not sufficient and WGCM asked the modelling groups to make hindcasts for every year from 1960 to present. For this reason a number of groups extended their CMIP5 hindcasts with additional start dates. [European Union]	Accepted -- The interval has been corrected from 5 to 1 years.
9-760	9	20	54	20	55	Chapter 9: Section 9.4 (Simulation of Recent and Longer Term Records in Global Models). The title of this section is directly followed by the title of next subsection (9.4.1) without an introduction to subsections 9.4.1 to 9.4.6 in which this section is structured. It would be desirable to have a paragraph for it to make easier the reading of the whole Section 9.4. [Government of Spain]	Rejected. This is a reasonable suggestion that was considered seriously by the chapter team. Ultimately no change was made because strong constraints on the length of the chapter.
9-761	9	20				9.4 Section 3.1.A of the paper "North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology" by Sheffield et al. could inform the discussion of model fidelity to spatial and temporal patterns of temperature and precipitation in the CMIP5 models. Section 9.4's discussion of precipitation patterns appears to be limited to the mean state and patterns associated with climate variability patterns; however, a discussion of seasonal contributions could fit, and the Sheffield et al. paper's findings with respect to storm tracks could be integrated into section 9.4.1.3.3. [Government of United States of America]	Taken into account. The Sheffield et al. paper has been assessed in the section on teleconnections over North America.
9-762	9	21	1	21	7	One should perhaps note that reanalysis products can be used with most confidence for variables that are well constrained by the assimilated observations. Where data is persistently sparse for a particular variable, and cannot be inferred well from other observed variables through the multi-variate assimilation process, the reanalysis may inherit flawed characteristics from the assimilating model. [Adrian Simmons, United Kingdom]	Taken into account. This point is now made in the section on metrics where the reanalysis are used to evaluate the models.
9-763	9	21	1	21	58	Comment: While in situ seasonal mean surface air temperature data are computed from daily maximum and minimum temperatures, it is unclear how the mean is computed in Figures 9.2 and 9.3 from reanalysis (e.g., from 3-hourly, 6-hourly, every time step, or maximum/minimum) and from models (e.g., from every time step or maximum/minimum). Therefore, it is necessary to provide the clarification in the discussion on p. 9-21. Such a clarification is also needed for all other figures and discussions on the model-data-reanalysis comparison of surface air temperature in this Chapter. [Xubin Zeng, United States of America]	Rejected. The key point of this figure is that in most areas models are within 2 degree Celsius of observationally-based estimates. ERA-INT Reanalysis have been demonstrated to be consistent with in-situ surface temperature measurements to a few tenths of a degree.
9-764	9	21	3	21	3	Near-global observations of winds are not really available for winds, at least not in the same amount, coverage and temporal resolution as the other fields mentioned here. Maybe you would like to change it to surface wind-stress which is available from satellites over oceans. [Gunilla Svensson, Sweden]	Noted. Upper air winds are a key product physically-consistent data assimilation, which are more reliable than surface wind stress, which, although good coverage is offered by satellites substantial uncertainties remain.
9-765	9	21	3			presumably just ozone (the models will use height resolved information, obviously total ozone helps, but is just part of the story) [Peter Braesicke, United Kingdom]	Taken into account. Total column ozone replaced with ozone.
9-766	9	21	4	21	4	reanalyses' instead of 'reanalysis' [CLAUDIA STUBENRAUCH, France]	Editorial. Corrected
9-767	9	21	4	21	4	I think you mean Box 2.3 here which discusses atmospheric reanalyses products, in which case this should be referred to rather than a Chapter (and the wrong chapter at that). [Peter Thorne, United States of America]	Editorial. Corrected
9-768	9	21	4	21	5	I wonder if this is the appropriate wording. Observations are so scarce that the issue is less to make them consistent among themselves than to obtain a realistic state of the atmospheric system consistent with them. [Ramon de Elia, Canada]	Rejected. Satellite measurements do not suffer from the same data sparsity limitations as in-situ data and physical consistency is important added value provided by reanalysis.
9-769	9	21	4			"Chapter 3" should be "Chapter 2". The chapter in question moved up one place from AR4 to AR5. [Adrian Simmons, United Kingdom]	Editorial. Corrected
9-770	9	21	4			"reanalysis" should be "reanalyses" [Adrian Simmons, United Kingdom]	Editorial. Corrected
9-771	9	21	4			Chapter 3 --> Chapter 2 [Thomas Stocker/ WGI TSU, Switzerland]	Editorial. Corrected

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-772	9	21	5	21	5	fundamental' instead of 'instrumental'? [CLAUDIA STUBENRAUCH, France]	Taken into account with text revision.
9-773	9	21	7			Add "and transient climate response" after "climate sensitivity". [Ronald Stouffer, United States of America]	Taken into account with text revision.
9-774	9	21	7			The persistent difficult bias in the Indian Ocean should be called out specifically too. [duane waliser, United States of America]	Rejected. This remark is not relevant to the identified line number, however, the persistent Indian Ocean bias is discussed later.
9-775	9	21	8	32	9	Check the headings in this section. I think 9.4.1.1.2 is supposed to be 9.4.1.2; also there are 2 9.4.1.3's. [Ross McKittrick, Canada]	Taken into account with text revision.
9-776	9	21	8	32	9	This is an important section since it discusses the ability of models to get temperatures correct. Section 9.4.1.1 is on the spatial patterns of the mean state. I would have thought 9.4.1.2, or some other subsequent section, would be on spatial patterns of trends. But I didn't see any such section. There is only a discussion of the large, global-scale trends (9.4.1.3) but not one on the spatial pattern of trends, which seems to me important for justifying the "fingerprint" approach to signal detection. The McKittrick and Tole paper in Climate Dynamics (see ref. in cell 34) tests how well each CMIP3 model reproduces the observed spatial pattern of trends over land in the CRU data set from 1979-2002. We used classical and Bayesian methods, testing the models independently, and in every possible linear combination, as well as against alternatives consisting of a vector of indicators of socioeconomic change and exogenous geographic processes. We found that 20 of the 22 models either have no explanatory power or are anticorrelated with the observed pattern of trends; that the socioeconomic variables have unique explanatory power that is not accounted for or encompassed by any GCM; and that a Bayesian Model Averaging exercise involving some 357 million combinations of explanatory variables shows that only 3 of the 22 GCMs (IAP-China, INM-Russia and NCAR CCSM 3.0) account for all the explanatory power among all the climate models. I'm not sure if these findings are of any interest to the chapter authors, but if so, the article citation is in the next cell. [Ross McKittrick, Canada]	Noted. The spatial pattern of trends are addressed in CH10.
9-777	9	21	8	32	9	McKittrick, Ross R. and Lise Tole (2012) "Evaluating Explanatory Models of the Spatial Pattern of Surface Climate Trends using Model Selection and Bayesian Averaging Methods" Climate Dynamics, 2012, DOI: 10.1007/s00382-012-1418-9 [Ross McKittrick, Canada]	Noted
9-778	9	21	9	22	19	This section describes changes only for the multi-model mean. Some indication how large the biases can be for individual models would be very useful here (or at least a reference to section 9.6.1.1). [Stefan Fronzek, Finland]	Taken into account with performance summary in 9.4.1.3
9-779	9	21	9			9.4.1.1 From here, the following few pages are about the mean state of the atmosphere. We feel that clouds and water vapor fields should also be considered as important mean state parameters. Also different parts of the atmosphere, e.g. boundary layer, upper troposphere, tropopause etc. [Government of United States of America]	Taken into account with text revision. Clouds and water vapor are assessed in 9.4.1.2.
9-780	9	21	11	21	46	Surface temperatures are invalid guides as to the likely quality of model predictions because of the concerns outlined in Chapter 9, page 9, lines 1 to 17. [David Webb, United Kingdom]	Rejected. As discussed, there are stronger tests, but tuning primarily addresses global means, not spatial patterns.
9-781	9	21	16	21	16	Replace "sunlight" with "solar radiation" as the former implies visible wavelengths, whereas all wavelengths are important. [Josephine Brown, Australia]	Taken into account with text revision.
9-782	9	21	20	21	22	Why is figure 9.2 shown when the assessment can only be made from a figure which is not shown? [Farahnaz Khosrawi, Sweden]	Taken into account with text revision.
9-783	9	21	20	21	32	This paragraph should focus on Figure 9.2 and discuss biases in annual mean temperature. Instead, it refers to a range of other variables and studies, which is very confusing. [Josephine Brown, Australia]	Taken into account. Paragraph has been revised with more emphasis on surface air temperature
9-784	9	21	20	21	32	To address concerns about the 'realism of ERA-INT' it would seem pertinent to reference the two Simmons et al. papers which compared ERA reanalyses to CRUTEM and concluded a very strong similarity. This would strengthen this analysis which otherwise may get critiqued unfairly for comparing models to 'a model'. The ERA surface T has been quite well validated against those nasty horrible real world observations and that	Taken into account with text revision.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						should be mentioned so as to help increase confidences in your comparisons herein. [Peter Thorne, United States of America]	
9-785	9	21	21	21	21	Which model ensemble is "presently available"? Maybe identify in Table 9.1 [Bart Van den Hurk, Netherlands]	All models available in the CMIP5 database are used, except where problems are found in the data.
9-786	9	21	21	21	22	Maybe it would help to move this sentence to end of the paragraph. [Farahnaz Khosrawi, Sweden]	Taken into account with text revision.
9-787	9	21	22	21	22	Are the observations shown in another chapter? Add a reference to the appropriate figure here. [European Union]	Rejected The global maps of the MMM and reanalysis are so similar that it would be duplicative to show both.
9-788	9	21	22	21	22	Starting from ", which" please rephrase. It is not clear what is meant. [Government of Germany]	Taken into account with text revision.
9-789	9	21	22	21	22	Are the observations for the same time period? [Government of Germany]	Taken into account in the figure caption
9-790	9	21	22	21	22	"... which although not shown is evident from the CMIP5 climatology bias." The connection between a meridional gradient being qualitatively similar to observations, and the climatology bias is not not obvious to me [Bart Van den Hurk, Netherlands]	Taken into account with text revision.
9-791	9	21	24	21	24	... with the ERA-interim reanalysis... [European Union]	Taken into account with text revision.
9-792	9	21	24	21	24	suggest to remove "observationally-based". The reanalysis IS NOT observation. In particular, the Ts might be more from model products than from assimilation due to lacking good observation fed into the reanalysis data such as over oceans and highland and two poles. [Jui-Lin (Frank) Li, United States of America]	Rejected. The surface temperature in the reanalysis have been compared to state-of-the in-situ measurements which were assimilated and include satellite data over the oceans.
9-793	9	21	24	21	32	In chapter 2.4.1.1 it is stated that the regional reanalyses of the magnitude in special trends in temperature for the Antarctic region have very low confidence. This circumstance should be accounted for when evaluating climate models in chapter 9. [Government of Germany]	Rejected. This section has nothing to do do with trends.
9-794	9	21	25			Are these the differences between model mean and obs, or individual models and obs? If the former then the wording on line 25 is incorrect. [Reto Knutti, Switzerland]	Taken in to account. Text is clarified.
9-795	9	21	27			"Arctic surface inversion in models have been compared with radiosonde data and identify significant model bias". We recommend editing this in the following way: "... and significant model bias has been identified". Nevertheless, the link of this statement to the context is very weak. Suggest to re-word like "Bias in model surface inversions may be an important reason", then cite the reference. [Government of United States of America]	Taken into account. Text has been modified as suggested.
9-796	9	21	28	21	28	Here it is not explained whether the MME absolute error (Fig. 9.2 lower left) is computed for the absolute differences of time means over all years available in ERA-interim - (I think this is what is shown) - or if the time mean of the absolute differences of annual means is shown. Probably one should write:... The mean absolute error of the individual CMIP5 model climatologies... [European Union]	Taken into account. Text has been modified as suggested.
9-797	9	21	28	21	28	Change "identify significant model bias" to "significant model biases have been identified" [Hai Lin, Canada]	Taken into account. Text has been modified as suggested.
9-798	9	21	28	21	28	Make explicit that these individual models are not shown in fig 9.2 [Bart Van den Hurk, Netherlands]	Taken into account. Text has been modified as suggested.
9-799	9	21	28	21	30	This information is not clearly deducted. [Massimo Bollasina, Italy]	Taken into account. Text has been modified to be more clear.
9-800	9	21	28	21	30	I did not understand the logic in this sentence. [Thorsten Mauritsen, Germany]	Taken into account. Text has been modified to be more clear.
9-801	9	21	28	21	32	The interpretation of the model mean bias vs. Individual bias is not that simple. By construction the former is smaller than the latter. See the work by Knutti 2010 JC and subsequent work by Annan and Hargreaves, JC and GRL. The bias of the model mean man also not be informative, in principal all models could have large	Taken into account. These issues are raised in the following paragraph and in 9.4.1.3.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						biases but they could compensate such that the mean looks ok. A further complication is the natural variability which implies that even if the models were perfect they will agree with the observations for precipitation for example because the record is too short. Those are difficult issues but at least they should be noted. [Reto Knutti, Switzerland]	
9-802	9	21	30	21	32	why do you gain confidence in the mean bias, as a measure of model quality, from the fact that the inconsistency between reanalyses is smaller than the mean absolute bias? -- unclear to us and probably needs more explanation. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text has been removed.
9-803	9	21	30			Should it not be mentioned which three re-analysis products were compared? [Stefan Fronzek, Finland]	Taken into account. Caption has been modified to explicitly mention the reanalysis.
9-804	9	21	34	21	40	Here, two different time periods are compared during a time with a known increase in near surface air temperatures. Do these results change, if model results for 1990-2005 are used? Does it make sense to compare two different time periods? [Government of Germany]	Noted. The time period in Fig 9.2(d) is different results from the period of availability in the three reanalysis, and in any case make these results conservative by, if anything, making the MME bias appearing larger.
9-805	9	21	35	21	40	color scale in bottom-left panel (and, to a lesser extent bottom-right): why does the scale go to 20deg? Can only be single grid-points that have such large values. At least, I cannot see them. Furthermore, colors are difficult to distinguish Figure could be improved by having the color bar from 1 to 8 (or 10), and one color for values exceeding that limit. [Andreas Sterl, Netherlands]	Taken into account. Figure has been improved.
9-806	9	21	35	21	40	Upper-left panel: unit is given as degC, but values are in K! [Andreas Sterl, Netherlands]	Taken into account. Units made consistent.
9-807	9	21	35			Figure 9.2: units of the top left panel are K, not oC (see color bar) [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Units made consistent.
9-808	9	21	38	21	38	change "(1990–2005, (Dee et al., 2011))." to "(1990–2005; Dee et al., 2011)." [SAMIR POKHREL, INDIA]	Taken into account. Text has been modified as suggested.
9-809	9	21	38	21	38	The time span 1990-2005 seems unnecessarily short. [Ian Watterson, Australia]	Taken into account. The limited time span is determined by the availability overlap of the 3 reanalysis which assimilate surface temperature.
9-810	9	21	39	21	39	Might consider to include the other two reanalysis into the figure caption in FIGURE 9.2. [Jui-Lin (Frank) Li, United States of America]	Taken into account. Other reanalysis have been identified.
9-811	9	21	39	21	39	Which reanalysis products are used here? Why is a reanalysis used as reference, and not CRU observations? [Bart Van den Hurk, Netherlands]	Noted. CRU observations are SST over ocean and ~2 metre temperature over land, whereas model fields are surface air temperature. Additionally, several recent papers by Simmons et al. have demonstrated how well ERA-INT agrees with surface observations.
9-812	9	21	39			It says "Mean inconsistency between three different reanalysis" Which three reanalyses does it refer to? [Government of Spain]	Taken into account. Text has been clarified.
9-813	9	21	42	21	46	This paragraph should summarise Figure 9.3 and model biases in seasonal cycle. It does not contain enough detail on model biases and is repetitive. [Josephine Brown, Australia]	Taken into account. Paragraph has been substantially revised.
9-814	9	21	42	21	46	In contrast to the previous paragraph (lines 20-32), there is no assessment of model performance regarding the seasonal cycle of near-surface air temperatures. Is the model seasonal cycle realistic/satisfactory? [Government of United States of America]	Taken into account. Paragraph has been substantially revised.
9-815	9	21	42	21	46	This paragraph seems open-ended, or misses the point. Consider removing. [Thorsten Mauritsen, Germany]	Taken into account. Paragraph has been substantially revised.
9-816	9	21	42	21	53	the analysis of biases in terms of difference between the seasons needs further discussion. Why the colors in the bottom panels of fig. 9.3 are exactly opposite in the northern hemisphere but not in the southern hemisphere? [Annalisa Cherchi, Italy]	Taken into account. Paragraph has been substantially revised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-817	9	21	43			"extreme season" meaning DJF and JJA sounds odd to me; wouldn't it be better to write "diff between winter and summer seasons" [Barbara Früh, Germany]	Taken into account. Paragraph has been substantially revised.
9-818	9	21	45	21	46	What is the main message of this Figure? [Government of Germany]	Taken into account. Paragraph has been substantially revised.
9-819	9	21	46			The statement on the bias in the seasonal cycle should be followed up by some interpretation: e.g. for the northern hemisphere, are the DJF temperatures too cold or are the JJA temperatures too warm, or both? [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Paragraph has been substantially revised.
9-820	9	21	46			I would have expected a comment about the possible source of bias for the seasonal cycle. Do we know what are the possible cause for the bias, particularly over the oceans? Is it due to problems in simulating the annual variability near oceanic fronts? Does the bias at 60°N result from the albedo bias discussed p.22 1.42-43? If yes a link to a further section could be added. [RYM MSADEK, United States of America]	Taken into account. Paragraph has been substantially revised.
9-821	9	21	48	21	53	Please indicate the time-period of the Era-Interim data used for this analysis. [Government of Germany]	Taken into account. Figure caption has been improved.
9-822	9	21	52	21	52	I don't understand the large mean bias in the absolute seasonality, given the fact that the mean seasonality seems to be underestimated considerably over many northern hemisphere land areas (fig 9.3c). Fig 9.4d is not at all discussed in the text, so could be deleted [Bart Van den Hurk, Netherlands]	Taken into account. Paragraph has been substantially revised.
9-823	9	21	55	21	56	The sentence seems to imply that 2-m temperature is not dependent on unresolved processes. Consider arguing differently. [Thorsten Mauritsen, Germany]	Taken into account. Paragraph has been substantially revised.
9-824	9	21	55	22	13	The discussion and results presented are about precipitation "rates" (units mm/day) not simply "precipitation" (which implies totals, in mm/cm/m). Please clarify in the text. [Government of United States of America]	Taken into account with change in text and caption of Figure 9.4
9-825	9	21	55	22	13	It needs noted that problems/uncertainty in the observed precipitation distribution. [Ronald Stouffer, United States of America]	taken into account - This is mentioned in the second sentence of this paragraph, which we have slightly rephrased for clarity.
9-826	9	21	55	22	19	Important discussion of ability of models to reproduce observed rainfall patterns and amounts - very relevant for policy makers who can judge the models' performance. [European Union]	Noted
9-827	9	21	55	22	35	The poor precipitation and cloud fields and the biases in humidity are probably a better indication of the quality of the predictive models and help explain the factor of two scatter in climate sensitivity. This needs to be stated. [David Webb, United Kingdom]	Noted - This is already discussed in the section on climate sensitivity and in Chapter 7.
9-828	9	21	55			How is the simulated precipitation improved with respect to CMIP3? [Massimo Bollasina, Italy]	Taken into account - a sentence has been added at the end of the paragraph to indicate that there has been little change and a reference to the FAQ has been added.
9-829	9	21	55			This point is linked to the remark #4 above. There are biases in precipitations. But are we sure we can say (as we all often do) that simulating precipitation is a « much tougher » test than simulating temperature ? Is it not so clear to me also that precipitation simulation is more dependent on moist parameterized physics than temperature. On large scales and long time constant, rainfall is strongly constrained by large scale advection of water. There are also very strong relationships between rainfall and clear-sky longwave cooling to space in the tropics. By comparison, the cloud radiative feedbacks may be much less directly constrained by the climate system. I was quite surprised to realize during the AMMA campaign that the radiative fluxes were in fact sometimes more dispersed than the rainfall itself, with surface SW radiation varying in the ITCZ region over the continent from 70 to 200 W/m ² for the various models (see for instance Fig 2 in Ruti et al., 2011, already in the bibliography). [Frédéric HOURDIN, France]	Taken into account - We have changed the wording to imply that precipitation is a more stringent test than temperature, rather than that it is harder to simulate.
9-830	9	21	57			It says " ... error relative to observations", but, which observations dataset is being considered? [Government of Spain]	Taken into account - the reference to the observations has been added.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-831	9	21				Figs 9.2, 9.3, 9.4, 9.5: choice of map projection and continent placement makes it difficult to visualize Europe. [Jean-Christophe Golaz, United States of America]	Taken into account - the map projections have been unified across the entire report.
9-832	9	22	4	22	5	"These biases contribute to the low pattern correlation for precipitation climatology (Figure 9.4)". This sentence, as it stands, implies that a pattern correlation plot is shown in Fig. 9.4. We recommend correcting the text to reflect what is shown in the plot. [Government of United States of America]	Taken into account - The reference to pattern correlation has been removed.
9-833	9	22	4	22	9	too long sentence [Bart Van den Hurk, Netherlands]	Taken into account - The sentence has been simplified.
9-834	9	22	5	22	5	Can you give a value for the low pattern correlation here or as an annotation to the figure? [Government of Germany]	Taken into account - The reference to pattern correlation has been removed.
9-835	9	22	9	22	9	Should also mention the overly zonal orientation of the South Pacific Convergence Zone in CMIP5 models (Brown, J. R., A. F. Moise and R. A. Colman (in press), The South Pacific Convergence Zone in CMIP5 simulations of historical and future climate, Climate Dynamics, doi:10.1007/s00382-012-1591-x.) This is distinct from the excess precipitation south of the equator in the eastern Pacific, as it mainly occurs in western and central Pacific. [Josephine Brown, Australia]	Taken into account - text and reference added.
9-836	9	22	9	22	13	The relevance of the last sentence ("Regional scale...") is not evident. This paragraph should focus on Figure 9.4. [Josephine Brown, Australia]	Rejected - The paragraph is about the evaluation of the mean state of precipitation and is hence broader than just Fig 9.4.
9-837	9	22	11	22	11	"improvements" used too many times in the same sentence [Annalisa Cherchi, Italy]	Taken into account - Text has been changed
9-838	9	22	12			Please reword this. Delworth et al. (2012) explore higher resolution, not different convection schemes. I'm not familiar with the Neale et al. paper, so please check that this is correct with regard to Neale et al. at least: "...improvements of resolution (Delworth et al. (2012) and of representation of subgrid scale processes, in particular of convection (Neale et al. 2008)." [Government of United States of America]	Taken into account - Text has been changed
9-839	9	22	12			The reference to Delworth et al. (2011) needs to be updated to Delworth et al. (2012) given in comment #5 above. [RYM MSADEK, United States of America]	Rejected - we feel this is the correct paper to reference here
9-840	9	22	13			Can you say here whether the CMIP5 models' precipitation is generally better than in CMIP3? It is mentioned in later sections but could be forward-referenced here. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Taken into account - a sentence has been added at the end of the paragraph to indicate that there has been little change and a reference to the FAQ has been added.
9-841	9	22	16	22	19	Fig 9.4 - why not show observed climatology ? [Peter Stott, United Kingdom of Great Britain & Northern Ireland]	Rejected - for consistency with the other figures we only show model errors.
9-842	9	22	18	22	18	Write out what the observations are used in Figure 9.4 caption. [Jui-Lin (Frank) Li, United States of America]	Accepted - observations are now explicitly named.
9-843	9	22	18	22	18	delete "and" [Bart Van den Hurk, Netherlands]	Taken into account - Text has been changed
9-844	9	22	18			bias instead of error [Peter Braesicke, United Kingdom]	Rejected - Bias refers to the mean error not the mean absolute error shown here.
9-845	9	22	18			The use of the word "observations" in this figure caption raises a point for consideration. The "observations" in this case are not direct observations, but the GPCP analysis of various sources of observational information. Over land it is based on the GPCP analysis of gauge data, which may be regarded as "observations" in regions of good data coverage, but which involve extrapolation into data void areas, and values are subject to a correction for undercatch. Estimates over the sea and (I believe) over otherwise data-sparse land areas come from satellite measurements from which the precipitation has to be inferred, which works less well at higher latitudes, as I understand it. So just a reanalyses are best not referred to as "observations" so this is true also of datasets such as GPCP. So in this figure caption "with and observations" should be replaced by something such as "and the observation-based GPCP estimates". "analyses" could be used instead of "estimates". [Adrian Simmons, United Kingdom]	Taken into account - "observations" has been replaced with "precipitation analyses from the Global Precipitation Climatology Project" both in the text and the figure caption.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-846	9	22	21	22	21	Chapter 9: Section 9.4.1 (Simulation of Recent and Longer Term Records in Global Models. Atmosphere). Numbering of Subsection 9.4.1.1.2 (Atmospheric moisture, clouds and radiation) seems to be wrong. Please revise it. [Government of Spain]	Accepted
9-847	9	22	21	22	21	9.4.1.1.2 Atmospheric moisture, clouds, and radiation should be: 9.4.1.2 Atmospheric moisture, clouds, and radiation [Koichi Sakaguchi, the United States]	Accepted
9-848	9	22	21	22	35	In addition to talking a lot about CMIP3 biases etc., we feel that it would be useful to emphasize new findings about CMIP5 here. For example biases and spread in model simulated clouds and water vapor is found to be largest in the upper troposphere among CMIP5 models (Jiang et al. 2012). Near the tropopause, some CMIP5 model simulated water vapor field are anti-correlated with the observations (Jiang et al. 2012). References: Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	Taken into account - the paper has been assessed and included in the text
9-849	9	22	21	22	46	will these paragraphs be updated with CMIP5 results if they become available in time? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - CMIP5 studies have been added where available.
9-850	9	22	21	24	16	Section 9.4.1.1.2 : Several recent papers evaluating CMIP5 clouds should be mentioned here. It includes Cesana and Chepfer (GRL, 2012) who compared 2D and 3D fields of cloud fraction derived from CALIPSO and from CMIP models, and Nam et al. (GRL, 2012) who compared CMIP5 model outputs with different satellite observations (CERES, CALIPSO and PARASOL) and showed that the 'too few, too bright' low-cloud problem was still present in CMIP5 models. [Sandrine BONY, France]	Taken into account - these studies were assessed and included in the text
9-851	9	22	21	24	16	We should add descriptions of CMIP5 model performance of clouds and water vapor simulations. Using A-Train satellite observations, Jiang et al. (2012) quantified model biases in water vapor and clouds. The key findings are "The model spreads and their differences from the observations are larger in the upper troposphere (UT) than in the lower or mid-troposphere (L/MT). The modeled mean cloud water content over tropical oceans range from ~3% to ~15× observations in the UT and 40% to 2× observations in the L/MT. For modeled water vapor, the mean values over tropical oceans range from ~1% to 2× of the observations in the UT and within 10% of the observations in the L/MT. " Also, it is good to mention that CMIP5 models show improvements from CMIP3 models in terms of water vapor path, ice (liquid) water path. See Jiang et al., (2012). Figure 1 and Figure 5 from Jiang et al. (2012) can be used to the model improvements from CMIP3 to CMIP5. [Government of United States of America]	Taken into account - these studies were assessed and included in the text
9-852	9	22	21	24	16	9.4.1.1.2 This section contains a thorough discussion of radiative fluxes, but does not give much attention to moisture fluxes. Section 3.2 of Sheffield et al. ("North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology") could be used to inform such a discussion. [Government of United States of America]	Noted - A global assessment of nmositure fluxes does not exist and as the focus of this chapter must be on global behaviour, no assessment can be carried out. Chapter 14 deals with regional issues.
9-853	9	22	21			Instead 9.4.1.1.2 must say 9.4.1.2 [Government of Spain]	Accepted
9-854	9	22	21			In Subsection 9.4.1.2 a quite extensive description of CMIP3 results. This is not done in Subsection 9.4.1.1. Please, try to uniform subsections [Government of Spain]	Taken into account - The section has been modified to include CMIP5 results where available.
9-855	9	22	24	22	24	It is unprecise to say "ERA" reanalysis, we have ERA-15, ERA-40 and ERA-Interim, do you mean all of them? [Gunilla Svensson, Sweden]	Taken into account - Text has been changed
9-856	9	22	24			ECMWF instead of ERA [Peter Braesicke, United Kingdom]	Accepted
9-857	9	22	26	22	27	introduce the TTL here; relates to 9-23 / 50 [Peter Braesicke, United Kingdom]	Rejected - we do not see a particular reason to highlight the TTL at this point
9-858	9	22	27	22	32	It is confusing whether relative or specific humidity is being referred to when the text says A is 100% higher than B or X varies by a factor of three. Also, the statements about large discrepancies in specific humidity at	Taken into account - This discussion has been changed and updated with new CMIP5 findings.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						higher levels are misleading because the main reason for them is discrepancies in temperature. The relative humidity profiles are not all that different among models, so the discrepancies do not point to hydrological errors so much as to energetic ones. Moreover the relative humidity is what matters most both for infrared radiation and the formation of clouds and precipitation. Finally, it would seem important to point out that biases in relative humidity do not affect the water vapour feedback to first order, because of the logarithmic nature of water vapour greenhouse effect (John and Soden 2007). [Steven Sherwood, Australia]	
9-859	9	22	30			Change Su et al . 2006 to Su et al. 2006b. [Government of United States of America]	Editorial
9-860	9	22	37	22	38	Why are the radiative fluxes "most important observable properties"? More is needed. [Ronald Stouffer, United States of America]	Taken into account - Text has been changed
9-861	9	22	37	22	42	Breakdown of the centred RMS error in the annual variation of cloud regimes shows that CMIP5 models tend to capture the temporal variation of the CRE reasonably well, with the main differences between models coming from the variation in amplitude. On the other hand, in the extra-tropics, most models fail to correctly represent both the amplitude and time variation of the CRE of congestus, frontal and stratocumulus regimes.(Tsushima et al. accepted) [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account - This new study and its findings are included further down in the discussion of CRE.
9-862	9	22	37	22	46	This paragraph does not sit well, because it describes the annual variation of CRE by tropical stratocumulus before annual climatology of CRE is discussed in p23 18-3. The placement should be reconsidered. [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account - the text has been deleted here as it referred to CMIP3 models only
9-863	9	22	37	23	10	This discussion only mainly touches on CMIP3 but there is a study on surface and TOA radiative flux biases on CMIP5 & CMIP3 that was submitted in time by Li et al. It was just resubmitted with revisions and is expected to be accepted. The reference is: Li, J.-L. F., D. E. Waliser, G. Stephens, S. Lee, T. S. L'Ecuyer, S. Kato, and N. Loeb (2012), Characterizing and Understanding Radiation Budget Biases in CMIP3/CMIP5 GCMs, Contemporary GCMs and Reanalyses, Journal of Geophysical Research, Submitted with revisions. The first author's email is "Jui-Lin (Frank) Li" <Juilin.F.Li@jpl.nasa.gov> [duane waliser, United States of America]	Rejected - The paper acceptance date (27 March 2013) is after the cutoff for assessment.
9-864	9	22	37	24	12	There are a number of references to Li et al. papers and in the item just above this one, I've provided some clarifications on the reference information. [duane waliser, United States of America]	Noted
9-865	9	22	41	22	41	"reanalysis" should be "radiosonde"? [Melissa Free, United States of America]	Rejected. Misplaced comment; re-interpreted as page 28. "Reanalysis" is correct.
9-866	9	22	42	22	42	insert "planetary" before "albedo" [Bart Van den Hurk, Netherlands]	Taken into account - the text has been modified to include newer results
9-867	9	22	46	22	46	Could include Seidel et al., GRL, doi:10.1029/2012GL053850 (in press) in support of uncertainties. [Melissa Free, United States of America]	Accepted. Misplaced comment, re-interpreted as page 26. Seidel et al. is now included.
9-868	9	22	48	22	55	But the bias in surface LW is also due to the bias in observed precip from climatologies such as GPCP, which are incompatible with the atmospheric radiative cooling inferred from CloudSat. See comment on row 11 above. [Anthony Del Genio, United States of America]	Noted
9-869	9	22	49			Add "global" before "downward all-sky" [Ronald Stouffer, United States of America]	Editorial
9-870	9	22	50			Add "global" before "downward lognwave" [Ronald Stouffer, United States of America]	Editorial
9-871	9	22	53	22	53	suggest to include Li et al. (2012). Li, J-L F., D. E. Waliser, G. Stephens, T. S. L'Ecuyer, S. Kato, N. Loeb, et al., (2012d) Characterizing and Understanding Cloud Water and Radiation Budget Biases in CMIP3/CMIP5 GCMs, Contemporary GCMs and Reanalyses, minor revision, JGR. [Jui-Lin (Frank) Li, United States of America]	Noted
9-872	9	22	53	22	53	"systematic omission of precipitating ice "à"systematic omission of precipitating and/or convective core ice" [Jui-Lin (Frank) Li, United States of America]	Editorial
9-873	9	22	53			Apart from the general overall comparisons of radiation in the above study, it would provide additional/useful	Taken into account.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						support along with the Waliser et al. citation already given as this connects the bias to those seen in CMIP3 and CMIP5. [duane waliser, United States of America]	
9-874	9	22	54	22	56	It may be appropriate to mention here that GCM or GCM-like radiative transfer codes have been shown to systematically underestimate atmospheric solar absorption even for pure gaseous atmospheres relative to line-by-line calculations (e.g., "b" subcases in Fig. 2 lower right panel of Oreopoulos, L., et al. (2012), The Continual Intercomparison of Radiation Codes: Results from Phase I, J. Geophys. Res., 117, D06118, doi:10.1029/2011JD016821). [Government of United States of America]	Accepted -- reference to CIRC article added.
9-875	9	22	57			An appropriate reference on the topic of solar absorption by aerosols is: Kim, D. and V. Ramanathan (2008), Solar radiation budget and radiative forcing due to aerosols and clouds, J. Geophys. Res., 113, D02203, doi:10.1029/2007JD008434. [Government of United States of America]	Accepted -- reference to Kim and Ramanathan added.
9-876	9	23	1	23	2	The way this sentence is written it implies that Wild et al. 2006 attribute the underestimate of SW atmospheric absorption to weak water vapor lines, which is not true (no such discussion can be found in that paper). A more appropriate reference for this topic would be: Learner, R. C. M., W. Zhong, J. D. Haigh, D. Belmiloud, and J. Clarke (1999), The contribution of unknown weak water vapor lines to the absorption of solar radiation, Geophys. Res. Lett., 26(24), 3609–3612. The authors paragraph should also consider mentioning possible missed absorption due to water vapor continuum, see for example: Ptashnik, I. V., R. A. McPheat, K. P. Shine, K. M. Smith, and R. G. Williams (2011), Water vapor self-continuum absorption in near-infrared windows derived from laboratory measurements, J. Geophys. Res., 116, D16305, doi:10.1029/2011JD015603. [Government of United States of America]	Taken into account. -- while the reviewer has misread the dependent clause (the reference to Wild was in the context of important of water vapor as a solar absorber), references to the effects of weak lines have been added.
9-877	9	23	3	23	6	We feel that broadband surface albedo comparisons aren't necessarily very illuminating since they the broadband albedo cannot be uniquely defined (e.g., what spectral weighting has been used?). Moreover, is ISCCP the best source for surface albedos? What about MODIS-derived surface albedos which provide some spectral breakdown and also a distinction between "white-sky" (diffuse incidence) and "black sky" (direct incidence)? Finally (and more importantly), are the surface albedo values quoted over land only? Because the Donohoe and Batisti 2011 paper referenced gives (their Tables 1 and 2) a global surface albedo of 12.3% (0.123) which is consistent with the value inferred from the previously quoted Stephens et al. (2012a) paper (their figure B1 implies an effective surface albedo of $23/(165+23)=0.122$). [Government of United States of America]	Accepted -- this material has been deleted.
9-878	9	23	3	24	4	The numbers for surface albedo seem large, given that most of the Earth is covered by oceans with an albedo of about 0.07. A quick calculation of the solar weighted surface albedo in ECHAM6 gives about 0.13. Maybe the numbers given are planetary albedo? In which case they seem somewhat large. [Thorsten Mauritsen, Germany]	Taken into account -- combined with comment 9-877
9-879	9	23	4	23	4	what is the uncertainty of the ISCCP albedo of 0.334? [CLAUDIA STUBENRAUCH, France]	Accepted -- this material has been deleted.
9-880	9	23	8	23	31	As noted lines 24-26 and apparent in Figure 9.5, there are serious problems in the subtropical stratocumulus regions, notably west of South and North America and southern Africa. Also in regions of trade cumulus. These almost certainly reveal difficulties in the modeling of the diurnal cycles of such clouds and probably also diurnal and angular sampling biases in the satellite observations. This would seem to be worth mentioning. Cf. also Fig. 9-30 for precipitation diurnal cycles. [Robert Kandel, France]	Taken into account.
9-881	9	23	11	23	11	CMIP5 -> CMIP3 [Bart Van den Hurk, Netherlands]	Rejected - CMIP5 is correct, it's the figure caption that was wrong and has been corrected.
9-882	9	23	12	23	22	I suggest referencing Figure 7.6 which shows the maps of CRE [Steven Sherwood, Australia]	Accepted -- reference to figure 7.6 included.
9-883	9	23	12			insert "(right panels)" after Figure [Barbara Früh, Germany]	Accepted
9-884	9	23	18	23	18	units of the fields in the maps are missing [Bart Van den Hurk, Netherlands]	Accepted -- units to fields shown in the maps have been added.
9-885	9	23	18			errors' in the figure caption, 'biases' in the text; I prefer biases, unless sources of error are discussed. [Peter Braesicke, United Kingdom]	Accepted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-886	9	23	19			Instead of "... CMIP3 multimodel mean" it must say "CMIP5". The text says that Figure 9.5 is referred to CMIP5. [Government of Spain]	Accepted
9-887	9	23	24	23	31	It is hard to follow the descriptions of Fig. 9.5. Please identify the stratocumulus and trade cumulus regions explicitly. "Too weak an effect of the model clouds on shortwave radiation" refers to what color in the figure. Clarify which data set is newer: EBAF2.6 or ES-4 (for those who are not familiar with these). [Government of United States of America]	Accepted -- Descriptions of figure 9.5 have been clarified as suggested.
9-888	9	23	26	23	26	"too weak cloud influence": in the figure I mainly see red colors over the indicated areas. It is suggested that this implies a too strong CRE [Bart Van den Hurk, Netherlands]	Accepted -- the sign of the biase in CRE has been clarified.
9-889	9	23	27	23	27	... Northern... [European Union]	Editorial
9-890	9	23	27	23	27	typo in "Northern" [Bart Van den Hurk, Netherlands]	Editorial
9-891	9	23	33	23	42	Should you also mention here model errors in the treatment of aerosol secondary effect? [Chiara Cagnazzo, Italy]	Rejected -- the signal associated with the anthropogenic indirect effect is an order of magnitude smaller than the biases under discussion here.
9-892	9	23	33			... the errors - that lead to a bias in - CRE ... [Peter Braesicke, United Kingdom]	Editorial
9-893	9	23	33			"errors" (here and elsewhere): I suggest to use biases rather than errors. For many people an error is something that is wrong in the sense of a coding error, faulty, defect, etc. It's not that the model is wrong in that sense, it's inaccurate or biased because of the simplifications and assumptions that inevitably need to be made in a model. [Reto Knutti, Switzerland]	Accepted -- "errors" replaced with "biases" where appropriate. As bias has a particular statistical meaning, not all occurrence of "error" can simply be replaced by it.
9-894	9	23	36	23	38	Tsushima et al.(accepted) also show that their Cloud regime error metric(CREM, Williams and Webb,2009) in the annual mean climatology show better score in the CMIP5 models compared to their CMIP3 counterparts, hence should be cited. [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Noted -- reference added.
9-895	9	23	38	23	39	We find the statement: "Particular examples include the improved simulation of vertically integrated ice water path (Jiang et al., 2012b; Li et al., 2012c)." contradicts these later statements: "Global mean values of simulated ice and liquid water path vary by factors of 2 to 10 between models (Jiang et al., 2012a; Li et al., 2012b).", p. 9-23, lines 45-46. "A particular challenge for models is the simulation of the correct phase of the cloud condensate, although very few observations are available to evaluate models particularly with respect to their representation of cloud ice (Li et al., 2012c; Waliser et al., 2009b).", p. 9-24, lines 7-9. [Government of United States of America]	Accepted -- these apparently contradictory statements have been reconciled.
9-896	9	23	38	23	41	About the improvement of vertically integrated ice water path, cloud regime analysis shows that many of the high cloud regimes are better represented in the CMIP5 versions both in the tropics and extra-tropics (anvil and deep convective regimes in the tropics, and cirrus and thin cirrus regimes in the extra-tropics) (Tsushima et al. accepted). [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Taken into account - we integrated this finding in the text
9-897	9	23	38	23	41	As for the reduction of overabundant optically thick clouds, breakdown of the cloud regime error metric(CREM) into the relative frequency of occurrence (RFO) error and cloud radiative property error shows that it is the cloud radiative property error which has been reduced in optically thick cloud regimes. (Tsushima et al., accepted). [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Accepted -- reference to CREM Tsushima paper repeated
9-898	9	23	38			Change Jiang et al. 2012b to Jiang et al. 2012. [Government of United States of America]	Noted-It's automatically generated by the IPCC endnote style.
9-899	9	23	39	23	39	It might not be proper to draw the conclusion that there is an improvement from CMIP3 to CMIP5 for liquid water path stated in line 39: "...Particular examples include the improved simulation of vertically integrated... cloud liquid water path". To date, there are still significant questions about the fidelity of the observational LWP references for GCM cloud liquid water and path. The main huddle in the observation side is that the observational data for liquid water path are poorly characterized and uncertainties are not well known (Li et al., 2012). Li, J.-L. F., S. Lee, D. E. Waliser, S. Lee., B. Guan, G. Stephens, Matt Christensen, J. Teixeira, 2012, Assessment of Cloud Liquid Water in CMIP3, CMIP5, and Contemporary GCMs and Analyses in Comparison with Observations, J. Geophys. Res., minor revision. [Jui-Lin (Frank) Li, United States of America]	Taken into account -- statement removed.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-900	9	23	39	23	40	Change Jiang et al. 2012a to Jiang et al. 2012 [Government of United States of America]	Noted-It's automatically generated by the IPCC endnote style. --combined with comment 9-898
9-901	9	23	39			Change Jiang et al. 2012b to Jiang et al. 2012 [Government of United States of America]	Noted-It's automatically generated by the IPCC endnote style. -- combined with comment 9-898
9-902	9	23	41	23	42	last sentence of paragraph: is this statement true? The coordinated assessment of cloud properties from satellite observations (Stubenrauch et al. 2012, available at http://climserv.ipsl.polytechnique.fr/gewexca/) has shown that most of the differences can be understood by instrument sensitivity or retrieval filtering and the resulting biases have been quantified. Though some variables are determined with smaller uncertainty than others. [CLAUDIA STUBENRAUCH, France]	Accepted -- This sentence has been removed.
9-903	9	23	41	23	42	this paragraph, related to Figure 9.5, mentions that "many of the above variables are now falling within the often large observational uncertainties". Yet these observational uncertainties are not included in the Figure, it seems. Suggest to add these to the Figure in order to support the statement [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - The sentence has been removed.
9-904	9	23	44	23	51	The reference Jiang et al. 2012a at line 45-46 and Jiang et al 2012b at line 50 all need to be changed to Jiang et al 2012. The CMIP5 model evaluation study by Jiang et al. (2012) not only includes the vertically integrated ice and liquid water path, but also includes the vertically resolved ice and liquid water content, or the mass mixing ratios for ice and liquid. The key findings are the following: (1) The CMIP5 model spreads and their differences from the observations are larger in the upper troposphere (UT) than in the lower or mid-troposphere (L/MT). (2) The CMIP5 modeled mean ice water content over tropical oceans range from ~3% to ~15× observations in the UT and 40% to 2× observations in the L/MT. The above three key findings should be included here since they are important findings related to the AR5 models, but were not available in AR4, since back in the AR4 time, the vertically resolved A-Train cloud observation was not available. [Government of United States of America]	Taken into account - The findings have been incorporated more directly in the text keeping in mind that a single study can only provide limited evidence for the assessment. The upper tropospheric issue was already specifically mentioned later in the text.
9-905	9	23	44	23	51	This paragraph suggests that it is mainly in the upper troposphere where models and observations disagree. Is that really true? I thought models were underestimating middle and low cloud amounts by more than any bias in the UT. Mid-level cloud is not even mentioned in this section. [Steven Sherwood, Australia]	Taken into account - the remaining sizeable errors in low and mid-level clouds are now also mentioned
9-906	9	23	44	23	51	Systematical underestimate of the relative frequency of occurrence (RFO) of cloud regimes of anvil cloud regime in the tropics and cirrus cloud regimes in the extra-tropics (i.e. cloud regime low cloud top pressure and medium cloud optical thickness) are found in CMIP5 models (Tsushima et al., accepted). [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Noted - These findings are too detailed to be included here.
9-907	9	23	45	23	46	This sentence unclear. What does 2 to 10 mean? Water path varies by 10 and ice by 2? [Government of Australia]	Taken into account - Text has been changed
9-908	9	23	47	23	47	change "(optical thickness >1.3, (Pincus et al., 2012))" to "(optical thickness >1.3; Pincus et al., 2012)" [SAMIR POKHREL, INDIA]	Editorial
9-909	9	23	47	23	47	The global mean fraction of clouds is estimated to about 0.68 +0.03, for cloud optical depth > 0.1 and not > 1.3 (Stubenrauch et al. 2012, accepted and available at http://climserv.ipsl.polytechnique.fr/gewexca/). Does the underestimation in models of 5-10% correspond to an underestimation for clouds with optical depth > 1.3? Does this mean that if considering clouds with opt depth > 0.1, the model underestimation of cloud fraction would increase? [CLAUDIA STUBENRAUCH, France]	Noted - The statement refers to the findings of the two studies quoted. Pincus et al. show a very large observational uncertainty in cloud fraction estimates for tau<1.3. Hence, they recommend not to evaluate model clouds with lower tau.
9-910	9	23	48			link to model resolution as well [Peter Braesicke, United Kingdom]	Rejected - There is no study that demonstrates this link for clouds.
9-911	9	23	50	23	51	Compensation of errors deserves further discussion (add to Box 9.1?) in the context of predictability: although the errors compensate in the current regime to produce semi-realistic results, they are likely to be affected in different ways by climate change, so the magnitude and even sign of response to forcings is less reliable where compensation of errors is suspected. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account - A discussion of this issue has been added to the new Section 9.2.3, which describes the overall model evaluation philosophy adopted here, including its limitations, one of which is potential error compensation.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-912	9	23				Figure 9-5: The caption of this figure refers to CMIP3 while the corresponding text passage (p. 9-23 lines 11 & 13) relates to CMIP5. Please check. [Government of Germany]	Taken into account - this has been corrected in the caption.
9-913	9	24	5	22	7	In Svensson and Karlsson (JCLIM, 2011) a more in-depth discussion on processes important in the Arctic and especially over the sea-ice such as turbulent fluxes and radiative fluxes. These does not give confidence that the models are good in representing the Arctic sea-ice (see comment 4 &5) [Gunilla Svensson, Sweden]	Noted - This section is specifically about clouds and the papers cited do provide sufficient evidence for the statements made. We therefore did not include the additional reference.
9-914	9	24	6	24	6	Pls change "(Karlsson and Svensson, 2010)" to "Karlsson and Svensson, 2010" [HASIBUR RAHAMAN, India]	Editorial
9-915	9	24	12			We recommend the addition of a citation to Su et al. (2012) on line 11. This paper provides regime-dependent error diagnosis for CMIP5 simulated clouds. Suggest to add the following on line 12: "Su et al. (2012) showed that the cloud simulation errors in the CMIP5 models are predominantly contributed by the errors in cloud parameterization schemes in the models, while the imperfect simulations of large-scale dynamic and thermodynamic states account for a relatively small fraction of total errors on the tropical average, but can be sizable on regional scales." [Government of United States of America]	Taken into account - This section now quotes a discussion on regime-oriented approaches in Section 9.2.1 and the paper has been included there.
9-916	9	24	14	24	16	Section 9.4.1.1.2 : Two comments : (1) Errors in the simulation of clouds are not only a concern for climate change projections but also for simulating many other aspects of the climate system (e.g. circulation, MJO, ENSO, etc) and biochemical processes (e.g. aerosols, cf Stier et al. ACP, 2012 showing that inter-model differences in cloud fraction contribute significantly to uncertainties in aerosol radiative forcing estimates). (2) Briant and Bony (GRL, 2012) actually showed that in the IPSL-CM5A model, errors in the simulation of present-day low-cloud radiative effects translated into different low-cloud feedbacks in climate change. [Sandrine BONY, France]	Noted - These issues are discussed in Chapter 7 and the main focus here is on evaluating the model simulations.
9-917	9	24	14	24	16	Talk about biases and sources of errors. [Peter Braesicke, United Kingdom]	Rejected - Bias has a particular connotation in statistics as th emean error, we do not refer to mean error here. Sources of the errors are discussed above where they are known. Many errors are not understood in their details.
9-918	9	24	14	24	16	Summary sentence should also indicate that there has been improvement in representation of a number of aspects of the climate system. [Government of Australia]	Rejected - This section discusses clouds and an overall summary of the chapter is provide at its conclusion.
9-919	9	24	14	24	16	"... reported in Chapter 12". This is passing the buck. Are the humidity, cloud, rainfall and other predictions of Chapter 12 likely to be "excellent", "very good", "good" or what? [David Webb, United Kingdom]	Rejected - This statement reflects our current inability to conclusively link errors in present-day simulations to effects on future smulations in more than a few examples. This is extensively discussed in many places in the Chapter and does not to be repeated here.
9-920	9	24	14	245	16	Important point. [Robert Kandel, France]	Noted
9-921	9	24	15	24	15	there is no sec 9.7.4 in this chapter [HASIBUR RAHAMAN, India]	Taken into account - the proper section number is now cited.
9-922	9	24	17			section 9.4.1.2 is missing [Barbara Früh, Germany]	Editorial
9-923	9	24	18	24	18	This section number should be 9.4.1.2 [Ian Watterson, Australia]	Editorial
9-924	9	24	18	25	46	this subsection is more on techniques than on performance of atmospheric variables. It should be more appropriate to have it in section 9.3 rather than 9.4 [Annalisa Cherchi, Italy]	Taken into account. Most of the paragraph is about results rather than technique, so it is appropriate to keep the material (and figure) here. But text has been revised, to make it more concise and result-oriented.
9-925	9	24	18	25	46	Section 9.4.1.3 The section is about model evaluation metrics. The satellite observational metrics developed for AR5/CMIP5 model evaluation (e.g. Jiang et al. 2012; http://climatesciences.jpl.nasa.gov/seminar/2011-09-	Taken into account. The suggested paper is assessed, but given the tight constraints on the

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						ipcc-models) have the following unique features that should be included in this report: (1) The metrics are bi-variate or multi-variate metrics that are designed to evaluate model performance in simulating more than one parameters simultaneously. The metrics are build upon collocated observations from multiple satellite instruments. (2) The metrics are designed to evaluate model performances in terms of spatial mean, variances, and correlation with the observations at different vertical pressure levels from boundary to mid-troposphere to upper troposphere and to tropopause; (3) Observational uncertainties are included in the formula for quantitative computation of model performance scores. The above model performance metrics are unique for AR5/CMIP5 model evaluations due to the fact that collocated A-Train observations were not available during the AR4 time, and thus should be highlighted in this chapter or this section. To illustrate this metrics, we suggest inclusion of Figure 9 or Figure 11 from Jiang et al. (2012) in this chapter 9. Reference: Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	chapter length no figure based on these new results was added.
9-926	9	24	18	25	46	We recommend that the authors consider adding Figure 11 from Jiang et al. (2012), which shows quantitative scores of model performance. [Government of United States of America]	Taken into account. The Jiang et al. paper is assessed, but no figure based on these new results is included.
9-927	9	24	18	25	48	There are two sections 9.4.1.3. I think the first of these should be 9.4.1.2 [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Editorial
9-928	9	24	20			tone done, emphasis traceability over the model does well; it depedns on the subjective choice of metrics [Peter Braesicke, United Kingdom]	Takken into account - language changed to reflect concern raised.
9-929	9	24	21	24	21	Write "there have" instead of "the have" [Juan Blanco, Spain]	Editorial
9-930	9	24	21	24	21	I would suggest to start the sentence with "These" or "The metrics" instead of just "The". [Farahnaz Khosrawi, Sweden]	Editorial
9-931	9	24	21	24	21	Change "The have been" to "They have been" [Hai Lin, Canada]	Editorial
9-932	9	24	21	24	21	"The have been ..." should be " They have been ..." [HASIBUR RAHAMAN, India]	Editorial
9-933	9	24	21			It says "... The have been used ..." it should say -->".. They have been used..." [Government of Spain]	Editorial
9-934	9	24	24			It says" .. All CMIP3 models ..." It should say -->"... All CMIP3 and CMIP5 models...". [Government of Spain]	Taken into account. Comment addressed.
9-935	9	24	26	24	26	longware --> longwave [François Massonnet, Belgium]	Editorial
9-936	9	24	28	24	30	Section 9.4.1.3 : The sentence « This example illustrates...effects). » doesn't make sense. The regional distributions of precipitation and clouds are very much connected to the large-scale circulation patterns, while the temperature varies much more smoothly. Fields that vary smoothly (e.g. temperature, water vapour) agree better with observations generally than the fields that exhibit large small-scale variations (e.g. precipitation)... especially when considering pattern correlations. [Sandrine BONY, France]	Taken into account. Text has been completely revised to reflect concern raised.
9-937	9	24	28	24	30	Precipitation (in models) is both a function of circulation and of parameterised cloud and convective processes. [Josephine Brown, Australia]	Taken into account. Text has been completely revised to reflect concern raised.
9-938	9	24	28	24	30	the 2m air temperature is also a parameterized quantity; therefore the sentence "This example illustrates...." is misleading [Barbara Früh, Germany]	Taken into account. Text has been completely revised to reflect concern raised.
9-939	9	24	28	24	30	As for point #4 and #11, I feel that this statement is quite misleading, both beacause of the comparison of things which are not comparable and because of the interpretation given. [Frédéric HOURDIN, France]	Taken into account. Text has been completely revised to reflect concern raised.
9-940	9	24	28	24	34	A slightly odd example, temperature depends more on radiation than circulation; fields dominated by parameterised fields can agree well, you mention the QBO in a later section ... I agree that there is a scale	Taken into account. Text has been completely revised to reflect concern raised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						issue, but it is more complex - please rephrase and select a better example (in parts) [Peter Braesicke, United Kingdom]	
9-941	9	24	30	24	34	meaning of sentence unclear, mainly in its last part [Annalisa Cherchi, Italy]	Taken into account. Text has been revised to be more clear.
9-942	9	24	32			"... validate against point-based observation". This gives an impression that surface "point-based" measurements having to be relied upon to validate simulated precipitation. Figure 6. focuses on the period of 1980-1999. Satellite-derived products have also been commonly used for validation. [Government of United States of America]	Taken into account. Text has been completely revised to reflect concern raised.
9-943	9	24	33	24	33	Model precipitation is generally validated against gridded data e.g. CMAP, GPCP not "point-based observations". [Josephine Brown, Australia]	Taken into account. Text has been completely revised to reflect concern raised.
9-944	9	24	33			Comments relating to Fig. 9.6: It would be helpful to divide the precipitation and SW CRE analysis in Fig. 9.6 into land and sea, as this may highlight both differences in uncertainty in the two regimes, which could be large, and also the differences in uncertainty in the observations (correlations between other obs datasets and the default products could be included on the Figure). (Some information on this is included in Section 9.6.1.1, so an alternative would be to forward-reference that section here.) [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Correlations with alternate observations are now shown, but the figure remains focus on global (land +ocean) because of space limitations for discussion
9-945	9	24	34	24	34	"smaller errors" -> "higher correlations" [Bart Van den Hurk, Netherlands]	Taken into account. Text corrected as suggested.
9-946	9	24	37	24	37	what does a "centred" correlation mean? Anomaly correlation? [Bart Van den Hurk, Netherlands]	Taken into account. An attempt has been made to make this more clear, although the term "centered" was suggested in First Order Draft Comments.
9-947	9	24	37	24	43	In the figure you show the average result for each model ensemble. What about the confidence in the multi-model means? Do model deficiencies tend to cancel each other out, i.e. the correlation pattern of the average model fields being larger than the average correlation? Please comment. [Government of Germany]	Taken into account. Applying these and other statistics on the multi-model mean does generally yield to an improved score. This is now addressed in the next Figure, 9.7
9-948	9	24	37			Fig. 9.6: Again here what is the role of natural variability? For a noisy field we would never expect perfect correlation even for a perfect model. It would be nice to add the values typically obtained by correlating two ensemble members of the same model. [Reto Knutti, Switzerland]	Taken into account. We considered including this, but found that with 20 year climatologies, the pattern correlations between combinations realizations are always above 0.99 which would provide little value on the figure.
9-949	9	24	42	24	42	Why are data upscaled to such a low resolution (5x5) grid? [Massimo Bolasina, Italy]	Taken into account. It is now clarified why 4x5 resolution was used.
9-950	9	24	42	24	42	The 'default' product is not highlighted in Table 9.2. Presumably first on the list is intended. The 'period' is not evident. [Ian Watterson, Australia]	Taken into account. Issue raised now addressed in text and Table 9.2.
9-951	9	24	45	25	2	Maybe it would be easier to highlight first where the observations (assimilation models!) deviate before classifying the climate models; I find the figure overloaded [Peter Braesicke, United Kingdom]	Taken into account. Observational differences are now discussed before model differences.
9-952	9	24	47	24	47	In this context also the Waugh and Eyring paper should be cited which applies the metrics for stratospheric resolving chemistry climate models: D. W. Waugh and V. Eyring, Quantitative performance metrics for stratospheric resolving chemistry-climate models, Atmos. Chem. Phys., 8, 5699-5713, 2008. [Farahnaz Khosrawi, Sweden]	Taken into account. Waugh and Eyring paper is assessed.
9-953	9	24	48	24	48	CMIP3 runs do not extend to 2005? Should this be CMIP5? [Josephine Brown, Australia]	Taken into account. Figure is now based on CMIP5
9-954	9	24	48	24	48	Typo? CMIP3 ---> CMIP5 [KIYOSHI TAKAHASHI, Japan]	Taken into account. Figure is now based on CMIP5
9-955	9	24	48	24	48	"...CMIP3 simulations" should be "...CMIP5 simulations" [Bin Wang, China]	Taken into account. Figure is now based on CMIP5
9-956	9	24	48			shouldn't it be CMIP5 instead of CMIP3? [Barbara Früh, Germany]	Taken into account. Figure is now based on CMIP5

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-957	9	24	52	24	52	Remove sentence "The results in this figure are illuminating". It does not add information to the discussion. [Josephine Brown, Australia]	Taken into account. Figure is now based on CMIP5.
9-958	9	24	55	25	12	Recently Jha et al. (2012) documented the diversity in the historic experiments of 10 CMIP5 models in simulating different aspects of SST, particularly those associated with ENSO, as well as the impact of low frequency variations on the ENSO variability and its global connection. It is shown that the majority of the CMIP5 models capture the relative large SSTA variance in the tropical central and eastern Pacific, as well as in North Pacific and North Atlantic. Meanwhile, the frequency of ENSO is hardly captured by almost all models, particularly for the period of 5-6 years. The models reproduce the global averaged trends, particularly since 1970s. However, almost no model correctly simulates the spatial pattern of the trends. These results suggest that it is still a challenge to reproduce the features of global historical SST variations with the state-of-the-art coupled general circulation model. The low frequency variations caused by external forcing's enhance the SST variability and also modify the global connection of ENSO. Jha, B., Z.-Z. Hu, and A. Kumar, 2012: SST and ENSO variability and change simulated in historical experiments of CMIP5 models. Clim. Dyn. (submitted). [Zeng-Zhen Hu, United States of America]	Rejected. This comment has nothing to do with this section.
9-959	9	24				The section 9.4.1.3 should be 9.4.1.2 [HASIBUR RAHAMAN, India]	Editorial
9-960	9	24				Section 9.4.1.3: The use of statistical methods like this is a fine first step, but I get the impression that by concentrating on it here and elsewhere the report gives the the impression that it is sufficient. A section like this really needs to lead onto one which shows how more focussed studies of the physics had led to model improvement. If this has not happened then it is a waste of time. [David Webb, United Kingdom]	Taken into account. Text has been revised to be more clear that this is a "first step" to complement the more in-depth diagnosis which is in this assessment. This is now the subject of the final paragraph of this section.
9-961	9	25	4	25	4	the "multi-model mean" is not shown in fig 9.7. Is it inferred from seeing more red colors than blue? [Bart Van den Hurk, Netherlands]	Taken into account. Multi-model mean now added to figure.
9-962	9	25	4	25	5	This refers to Figure 9.7 showing the mean and median agreement between multi-model outputs and observations, however mean and median are not shown in Figure 9.7. [Government of Australia]	Taken into account. Multi-model mean now added to figure.
9-963	9	25	4	25	5	Text states that Figure 9.7 shows that the multi-model mean agrees better with the observations than any individual model - but this figure only shows errors for individual models. [European Union]	Taken into account. Multi-model mean now added to figure.
9-964	9	25	4	25	5	"another notable feature of Figure 9.7 is multi-model mean (median) agree more favourably..." -- please specify where this can be seen in figure? Are there entries for multi-model mean/median or are you referring to the pattern of positive/negative bias between the models for a particular quantity, thus inferring a mean/median by eye? Please clarify [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Multi-model mean now added to figure.
9-965	9	25	4	25	6	The multi-model mean that is discussed here suggests that it is shown on Fig9-7 but it is not. Check if a column was forgotten in the figure or modify the figure's description. [RYM MSADEK, United States of America]	Taken into account. Multi-model mean now added to figure.
9-966	9	25	4	25	9	How can I see that mean and median compare better with observations than any individual model? This should be explained. And, more importantly, is it right? The first model (ACCESS1-0) is all blue. If I understand the caption correctly this means that for every variable this model is better than the median of all models. [Andreas Sterl, Netherlands]	Taken into account. Multi-model mean now added to figure.
9-967	9	25	8	25	9	Would better read 'compares better with observations than most individual models'. Cannot assume that multi-model mean agrees 'well' with all observations. [Government of Australia]	Taken into account. Text corrected as suggested.
9-968	9	25	8			I suggest that the authors mention that part of the fact that the model mean agrees better with observations than a single model is fundamental, as pointed out by Annan and Hargreaves. See also Knutti 2010 J. Clim. [Reto Knutti, Switzerland]	Taken into account. Comment addressed.
9-969	9	25	11	25	21	Figure 9.7 did not include the results of FGOALS-g2, one of CMIP5 models (see http://cmip-pcmdi.llnl.gov/cmip5/availability.html). Please revise this figure and include the results of FGOALS-g2. It will be better to evaluate the model comprehensive performance if more variables such as Q850, PSL(sea level	Taken into account. All models available now included in the figure.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						pressure), and LWCF(TOA longwave cloud radiative effect), are included in Fig 9.7(Glecker et al., 2008). [Bin Wang, China]	
9-970	9	25	12	25	12	the colors of the extremes (purple and dark red) are too similar [Bart Van den Hurk, Netherlands]	Editorial
9-971	9	25	12			Change "CMIP5" to "CMIP3". (also in line 17 of the same page). [Government of United States of America]	Taken into account. Comment addressed.
9-972	9	25	14			instead of error: better agreement was achieved [Peter Braesicke, United Kingdom]	Taken into account. The spirit of this comment was applied other than the line number identified, which was the figure caption.
9-973	9	25	17			Change "CMIP5" to "CMIP3". (also in line 17 of the same page). [Government of United States of America]	Taken into account. Comment addressed.
9-974	9	25	23	25	39	These two paragraphs appear to overlap with discussion in Section 9.2.2.7. There should be more focus here on the CMIP5 model biases and implications for model reliability, rather than abstract comparison of methods. [Josephine Brown, Australia]	Taken into account. Comment addressed.
9-975	9	25	27	25	27	typo: take out last word: 'to' [CLAUDIA STUBENRAUCH, France]	Editorial
9-976	9	25	27			I think the questions what is the models purpose and what means fit for purpose must be answered first. I have no doubts that metrics improve traceability of change, but metrics are not necessarily helping to understand the physics better! [Peter Braesicke, United Kingdom]	Taken into account. The metrics presented in this section are not target a specific model application, however the text now points to Section 9.8.3 which does.
9-977	9	25	32	25	33	Here it is said that "...it does reduce the chance that a poorer performing model will get the right answer for the wrong reasons". It is an important point but nowhere is discussed the consequences --or how bad it is for our aims-- to get the right answer for the wrong reasons. Some people believe it is immaterial for our aims, while others consider it fundamentally wrong. Clearly it is a very tricky issue since if you look hard enough probably most right results are for the wrong reasons. It will be great if this can be discussed to make this topic more mainstream. It reappears in page 62 line 33-34 and page 68 line 52 [Ramon de Elia, Canada]	Taken into account. This is indeed an important topic. Because of tight space requirements it is not elaborated upon in this section continues to be discussed elsewhere as identified by the reviewer.
9-978	9	25	35	25	37	This sentence needs clarification. Also suggest rewriting with point made, then reference given, rather than starting "Author xxx said" [Government of Australia]	Editorial
9-979	9	25	44	25	46	I could not disagree more strongly with this statement. Confidence in metrics is greatest when the metrics tell us something about the climate changes that we are using the models to predict. The metrics that we have mostly do not do that, so we should have no confidence in them. It is precisely the desire for simplicity and robustness (which makes life easier for us) that has led us down this unproductive path. The lesson to be learned is that apparently some things that are difficult to calculate across a range of models and difficult to observe accurately (which make life difficult for us) are the very things that control climate change. Climate change is a set of subtle, small changes in the Earth system - why should we think that simple, robust observables make the best metrics for understanding it? [Anthony Del Genio, United States of America]	Taken into account. It is now indicated that while the metrics used in this section succinctly quantify the differences between models and observations, there is little evidence that these relate to the reliability of projections. Additionally, the reader is referred to Section 9.8.3 which does directly address the points made in this reviewer comment.
9-980	9	25	45			The robustness of the performance metric mode of evaluation also benefits from a wide range of observation-based products. [duane waliser, United States of America]	Taken into account. Comment addressed.
9-981	9	25	48	25	48	9.4.1.3 Long-Term Global-Scale Changes should be: 9.4.1.4 Long-Term Global-Scale Changes [Koichi Sakaguchi, the United States]	Editorial
9-982	9	25	48	31	40	Chapter 9: Section 9.4.1 (Simulation of Recent and Longer Term Records in Global Models. Atmosphere). Numbering of Subsections from 9.4.1.3 (Long Term Global Scale Changes) onwards seems to be wrong. Please revise it. [Government of Spain]	Editorial
9-983	9	25	48			section 9.4.1.3 is double (see p.24 l.18) [Barbara Früh, Germany]	Editorial
9-984	9	25	48			9.4.1.3 The number of this subsection is repeated. It also appears in Page 24 line 18. [Government of Spain]	Editorial
9-985	9	25	55	25	55	9.4.1.3.1 Global surface temperature and humidity should be: 9.4.1.4.1 Global surface temperature and humidity [Koichi Sakaguchi, the United States]	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-986	9	25	56			The section discusses the agreement between the observations and the MME, but nothing specific is explained about the largest disagreements: (1) following Krakatau the models respond much more than observed; (2) after 1995, the MME starts to diverge from the observations. [European Union]	Thank you for raising these issues. The disagreement between model simulations and observational data following Krakatau is partly explained by uncertainties in both the observations and the estimated forcing by the volcanic eruption. The MME divergence from observations since 1995 is discussed in Box 9.2.
9-987	9	25	56			The legend of Figure 9.8 does not explain the upper part of the Figure. [Government of Spain]	Taken into account. The discussion of the global mean absolute temperatures is now provided.
9-988	9	25				Figure 9.7: Multi-model mean (and median) is referred in the text, but not included in the figure. [Koichi Sakaguchi, the United States]	Noted. A column for MM mean and median has been added in Figure 9.7.
9-989	9	25				Section 9.4.1.3.1: Again a section starts with the global surface surface temperature although this is the field that was most likely used to tune the model (see page 9 lines 1 to 17). To repeat: surface temperature is a poor check on the quality of each model. [David Webb, United Kingdom]	Rejected: Figure 9.8 shows GMST evolution and therefore, the transient response of the model is being evaluated as an independent diagnostic compared with the GMST climatology.
9-990	9	26	1	26	12	The model mean tends to show a too strong response to volcanoes. This is not referred to in the text, but a remark would be welcome [Bart Van den Hurk, Netherlands]	Rejected. It is not evident that the model mean responds too strongly to volcanoes in Figure 9.8.
9-991	9	26	1			How is the difference in the 1961-1990 temperature in the models (inset of Fig. 9.8) affecting the time series? [Massimo Bollasina, Italy]	Noted. The temperature time-series are shown as anomalies with respect to the 1961-1990 mean. The 1961-1990 mean values for each model are shown to indicate the spread among the models' basic states.
9-992	9	26	2	26	4	Because every model has its own interannual variability, we feel that this statement is misleading. One needs to evaluate the frequency and the amplitude of the modes of variability of the models individually and then compare them to observations, rather than averaging all the multi-models and concluding that the models have little interannual variability. [Government of United States of America]	Taken into account. The role of internal variability has been clarified in Box 9.2. This main point is that the random phase of the interannual variability will tend to cancel out for the model mean whereas cancellation will not occur for the observations.
9-993	9	26	2	26	4	This is a misleading statement because every model has its own interannual variability. One needs to evaluate the frequency and the amplitude of the modes of variability of the models individually and then compare them to observations, rather than averaging all the multi-models and concluding that the models have little interannual variability. [Government of United States of America]	Taken into account. The role of internal variability has been clarified in Box 9.2. This main point is that the random phase of the interannual variability will tend to cancel out for the model mean whereas cancellation will not occur for the observations.
9-994	9	26	2			"although again there are some important differences among models": What are the "some important differences"? There are no hint or reference in the context. Some elaboration will be useful. [Government of United States of America]	Taken into account. These differences, including the realized internal variability, have been clarified and elaborated in Box 9.2.
9-995	9	26	3	26	3	insert "average" after "ensemble" [Bart Van den Hurk, Netherlands]	Editorial
9-996	9	26	3			Add "mean" after "ensemble". [Government of United States of America]	Editorial
9-997	9	26	4			Another relevant factor is that the variability in the CMIP model simulations is not timed to match that of the real atmosphere and differs from simulation to simulation. [Dian Seidel, United States of America]	Noted. Thank you for the clarification. A discussion of internal variability has been included in Box 9.2.
9-998	9	26	10	26	12	See comments above (page 3 line 45). The models have not broadly captured the latest temperatures - yet look OK earlier due to tuning. [John Christy, United States of America]	Noted. This issue is now addressed in Box 9.2.
9-999	9	26	11	26	11	As has been widely publicised, the observed warming in the last 15 years is small, and the model mean clearly rises more in Fig. 9.8. Some discussion of this is needed, including specific reference to other chapters that consider the issue. [Government of Australia]	Noted. This issue is now addressed in Box 9.2.
9-1000	9	26	11	26	11	"Warming of recent decades" is not very explicit, and might be confusing given that the trend in observed global mean temperature is not significantly different from zero since 1998. It would be better to be explicit	Taken into account.: This now makes reference to Ch11 and Box 9.2.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						here and give a time frame, egg, "Warming since 19XX" or "warming over the past XX years". Regarding the temporal evolution over the last decade or so, we suggest to add a brief statement on this here with potential reference to the Ch11 assessment (Figure 11.33) [Thomas Stocker/ WGI TSU, Switzerland]	
9-1001	9	26	11	26	12	Model tuning or selection also influences this level of agreement. As stated by Mauritsen et al 2012 (doi:10.1029/2012MS000154) "Climate models ability to simulate the 20th century temperature increase with fidelity has become something of a show-stopper as a model unable to reproduce the 20th century would probably not see publication" [Jean-Christophe Golaz, United States of America]	Noted.
9-1002	9	26	14	26	23	What about differences between upper and bottom panels in fig. 9.8? Why EMIC appear to have less spread in the simulations (i.e. less models in the overestimation part of the plot)? [Annalisa Cherchi, Italy]	Taken into account.. This has been clarified to explain that EMICs do not show the same level of internal variability.
9-1003	9	26	14	26	23	Figure 9.8 did not include the results of FGOALS-g2, FGOALS-s2 and FIO-ESM. They are CMIP5 models ((see http://cmip-pcmdi.llnl.gov/cmip5/availability.html). Please revise this figure and include the results of these models. [Bin Wang, China]	This could be done provided that the model data are available as noted. The figure now contains models that were available prior to the IPCC deadlines.
9-1004	9	26	14			Figure 9.8: has there been any attempt to quantify/remove model drift? [Jean-Christophe Golaz, United States of America]	Noted. Model drift has not been removed for global mean surface temperature time-series.
9-1005	9	26	19	26	19	There is a newer Vose et al. reference (see Chapter 2) for NCDC's MLOST product, which should also be called NCDC MLOST for consistency. This comment applies to several other segments of text and figure captions within the chapter. [Peter Thorne, United States of America]	Taken into account.Thank you for the clarification and update for the reference.
9-1006	9	26	20			Please remove "absolute" since the temperature is shown in degrees C. [Government of United States of America]	Taken into account.
9-1007	9	26	22	26	23	delete "vertical ... eruptions" (double) [Bart Van den Hurk, Netherlands]	Editorial
9-1008	9	26	25	26	26	An additional reference (Carvalho and Jones, "CMIP5 Simulations of Low-Level Tropospheric Temperature and Moisture over tropical America") could be added to support the claim that "Simulated changes in near surface specific humidity over land have been examined and found to be broadly consistent with observational estimates." Please note that the Carvalho and Jones paper does not compare against observations, but instead compares against reanalysis, so the language would have to be adjusted accordingly. Temperature results from the paper noted here could also be described in the language in this section. [Government of United States of America]	Noted, however, this paragraph on humidity was removed.
9-1009	9	26	34	26	34	This section discusses lower troposphere and middle troposphere, but is labelled upper troposphere -- make title more appropriate. [Government of Australia]	Accepted -- have changed title of subsection to "tropical tropospheric temperature trends"
9-1010	9	26	34	26	34	9.4.1.3.2 Upper tropospheric temperature trends should be : 9.4.1.4.2 Upper tropospheric temperature trends [Koichi Sakaguchi, the United States]	Accepted -- thank you.
9-1011	9	26	34	26	51	Whilst very important, this paragraph is poorly constructed and badly draft to make it difficult readily to understand [Government of United Kingdom of Great Britain & Northern Ireland]	Taken into account -- however, in the absence of any concrete suggestion or pointer, it is difficult to act upon this comment except to point out that owing to the level of contention on this issue, precision and traceability are more important than ease-of-reading. Nevertheless, we have attempted to be more reader-friendly.
9-1012	9	26	34	26	51	Again I would note that some of the debate regarding the derivation of satellite products (in particular tropospheric temperatures) has concerned satellite sampling corrections for diurnal variations. [Robert Kandel, France]	Noted -- while this section must be cognizant of observational uncertainty, the latter is assessed in Chapter 2, to which frequent reference is made. No change.
9-1013	9	26	34	27	29	Section 9.4.1.3.2: In the upper troposphere to lower stratosphere region, the humidity, ice clouds and temperature is not well simulated by models (e.g. Jiang et al 2012) or even the analysis models (e.g. Figure 5 and 5 in Jiang et al. 2010). So large errors in CMIP5 models in terms of simulating the upper tropospheric	Noted -- while these papers are relevant of other sections of this chapter, the current subsection deals with uncertainties that are likely to be dominated by

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						tends are expected. References: Jiang, J.H., H. Su, S. Pawson, H.C. Liu, W. Read, J.W. Waters, M. Santee, D.L. Wu, M. Schwartz, N. Livesey, A. Lambert, R. Fuller, and J.N. Lee, "Five-year (2004-2009) Observations of Upper Tropospheric Water Vapor and Cloud Ice from MLS and Comparisons with GEOS-5 analyses," J. Geophys. Res. 115, D15103, doi:10.1029/2009JD013256, 2010. Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	other processes. No change.
9-1014	9	26	34	27	33	The section should mention and take into account Seidel et al., GRL, doi:10.1029/2012GL053850 (in press) which shows that uncertainties in observations make the differences between models and observations less clear than this section states. The conclusions here should be softened accordingly. [Melissa Free, United States of America]	Taken into account, thank you. This paper is now included in the assessment, and the stated confidence level has been adjusted to low, owing to the large observational uncertainty, consistent with Section 2.4.4.
9-1015	9	26	34	27	33	This section should note that many of the studies here use old versions of the observed data products. In particular, none have to date considered the recent RICH analyses {Haimberger et al., 2012} which show much greater amplification within the tropics than the older radiosonde datasets did. This should be noted appropriately in redrafting this section. [Peter Thorne, United States of America]	Taken into account, thank you. The SOD text, which referred to the FOD text of Ch02, already included a discussion of RICH. The Seidel et al. (see comment 9-1014) indeed shows that model stratification change is more consistent with the newer observational analyses, and this results leads to a less robust assessment of models overestimating static stability.
9-1016	9	26	34			The 9.4.1.3.2 is a very descriptive subsection but with no Figures. [Government of Spain]	Taken into account -- however, in the absence of any concrete suggestion or pointer, it is difficult to act upon this comment except to point out that owing to the level of contention on this issue, precision and traceability are more important than ease-of-reading. Nevertheless, we have attempted to be more reader-friendly.
9-1017	9	26	34			This section is all about the tropics, which has been a focus of interest. The rest of the world should be addressed too. Also, the results a new study addressing the tropical trends (Seidel, D. J., M. P. Free, and J. S. Wang, Reexamining the warming in the tropical upper troposphere: Models versus radiosonde observations, Geophys. Res. Lett., doi:10.1029/2012GL053850, in press, accepted 16 October 2012) suggests that the model discrepancies might not be as serious as suggested by earlier work, so the statement (page 27, line 31) of high confidence should probably be softened. [Dian Seidel, United States of America]	Taken into account, thank you. This paper is now included in the assessment, and the stated confidence level has been adjusted to low (medium evidence, low agreement). However, since the tropics have been a particular point of contention and space constraints are severe, we have not included an explicit assessment of the extratropics.
9-1018	9	26	35	26	51	Regarding the stated range of tropical lower- and mid-troposphere trend estimates from McKittrick et al. (2010), they were updated in McKittrick et al. (2011), where the range is LT: 0.08 - 0.15 C/decade, and MT: 0.03 - 0.11 C/decade. [Ross McKittrick, Canada]	Taken into account, thank you The reference has been added, but the quoted range is taken from Chapter 2 and does not change as a result of this inclusion.
9-1019	9	26	35	26	51	Reference: R. McKittrick, Stephen McIntyre, Chad Herman, Corrigendum, Atmospheric Science Letters, 2011, 12, 4 [Ross McKittrick, Canada]	Taken into account, thank you.
9-1020	9	26	35	26	51	The point made in the last line of this paragraph is extremely important. As you indicate, the evidence for warming, or lack thereof, in the tropical troposphere has been subject to intense controversy and seems to have considerable importance for assessing climate models. The models all indicate that there should be amplified warming in the tropical troposphere, and it is a vast part of the atmosphere with apparently great importance for understanding planetary-scale climatic processes. So it seems inadequate to observe that some data sets show no significant warming in that region over the 1979-2009 interval and then move onto another topic. Although the next paragraph goes on to compare observed trends to the model projections,	Noted. There are really two issues here. The first is whether the troposphere has warmed; the overall assessment is of course left to Ch02. The second is whether, apart from folklore, there is evidence that indeed the warming of the tropical troposphere has anything to do with water vapour feedback. In the absence of published evidence for the latter, an

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						shouldn't there first be some discussion of the importance of this region? Wouldn't the absence of any significant warming in this region indicate something about, say, water vapour feedbacks? Recall that the context of this issue is your declaration of "very high confidence" in the models, on the basis of which many of the major conclusions of the IPCC report will be based. If there was a significant warming trend in the tropical troposphere you would no doubt highlight it as a reason for your high confidence in the models. Since you maintain that you have very high confidence in the models it must mean you can explain the lack of warming. This would be a suitable place to do so. [Ross McKittrick, Canada]	assessment is not included.
9-1021	9	26	35	27	29	This is a somewhat tortured explanation of things. The better way to approach it was done in Christy et al. 2010 - by means of the amplification factor (i.e. ratio of tropospheric to surface trends). This way you can compare a metric that removes the impact of the natural variation differences, so intermodel and model vs. obs comparisons are apples to apples.. Observations run about 0.8 to 1.0, while models cluster around 1.4 (tropics) - highly significant difference. Also, talking about the ratio as the metric allows the authors to save a lot of text about comparisons of individual trends themselves in the face of differing natural fluctuations (and save some misleading remarks as mentioned below). The ratio takes care of that. There is a new paper from Seidel et al (GRL) that looks at trend differentials between the upper troposphere (200, 300 hPa) and lower troposphere (700 hPa). Though interesting, it does not address the trend differentials between the surface and upper troposphere - which is the critical issue at hand. I have asked that they provide their Fig 3, but with the surface trend (not 700 hPa) as the lower boundary. The result will be consistent with our previous work showing models portray too much warming aloft. Also, while on this topic, Fu et al. 2011 utilized T2LT and T24 - two quantities with enhanced error characteristics (we don't produce T24 because it is unphysical.) Though I agree with their result, observations and models disagree, the measurement error of these quantities was not really considered, which then would weaken the final conclusions. [John Christy, United States of America]	Noted. However, this is an assessment of the published literature and moreover needs to be consistent with the approach taken in Ch02. Hence, to provide maximum traceability with the published papers, we decided to stick with the individual trend estimates because this allows the reader immediately to compare the numbers. Indeed it would have been useful had Seidel et al. and Po-Chedley and Fu also computed to difference and ratio relative to surface temperature, but we need to make do with what is given to us. Finally, it is not correct that using trend ratios eliminates the effects of natural variability, see response below.
9-1022	9	26	35	27	29	This is a very nice discussion of a difficult issue (although in the final paragraph it does get a bit hard to follow). The failure to mention any of this in the executive summary seems a conspicuous omission. [Steven Sherwood, Australia]	Taken into account, thank you. We have aimed at making the text more reader-friendly and have elevated our summary assessment to the Executive Summary.
9-1023	9	26	36	26	36	(e.g., McKittrick et al., 2010; Santer et al., 2012). [Hai Lin, Canada]	Editorial
9-1024	9	26	36	26	36	change "(e.g., (McKittrick et al., 2010) (Santer et al., 2012))" to "(e.g., McKittrick et al., 2010; Santer et al., 2012)" [SAMIR POKHREL, INDIA]	Editorial
9-1025	9	26	36	26	36	Pls change "(e.g., (McKittrick et al., 2010) (Santer et al., 2012))" to "(e.g., McKittrick et al., 2010; Santer et al., 2012)" [HASIBUR RAHAMAN, India]	Editorial
9-1026	9	26	40	26	51	The discussion separates the different temperature trend measurements from the consistency in the 95% confidence level. While it makes sense, it made this reviewer re-read the passage several times to eliminate any confusion, because the trends themselves are not expressed with confidence limits attached. [David Bader, United States of America]	Taken into account, thank you. Sentence has been removed.
9-1027	9	26	44	26	44	Chapter 2 changed the section ordering and hence figure indexing so you will need to check with Chapter 2 what the final version of this figure is indexed as (its not Figure 2.15 any longer and it may still change). [Peter Thorne, United States of America]	Accepted -- figure and subsection references have been updated.
9-1028	9	26	50	26	50	This sentence is not clear, perhaps it could be omitted, or stated 'zero trend is not inconsistent with some estimates'? [Government of Australia]	Taken into account, thank you. Sentence has been removed.
9-1029	9	26	50	26	51	This sentence does not make sense. Note that "trend" appears to mean different things in different locations throughout the report. The definition given in the glossary is insufficient. [Leonard Smith, United Kingdom]	Taken into account, thank you. Sentence has been removed.
9-1030	9	26	51	26	51	This sentence does not make grammatical sense. Rewrite along the lines of: "A trend of zero is within the 95% confidence interval of some observational trend estimates, but not of others". It might be useful to give numbers or describe the assumptions that lead to a smaller or larger estimate? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account, thank you. Sentence has been removed.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1031	9	26	51	26	51	"consistent with some ... others": a very vague and quite non-informative statement [Bart Van den Hurk, Netherlands]	Taken into account, thank you. Sentence has been removed.
9-1032	9	26	57	26	57	suggest replacement of "...who, however considered the period.." with simply "for the period..." [David Bader, United States of America]	Accepted, thank you.
9-1033	9	26	57	26	57	We recommend that the authors rephrase the text within the parentheses. [Government of United States of America]	Taken into account. The text is not meant to be within parantheses; proper formatting of references will solve this particular issue.
9-1034	9	26				Figure 9.8: The graph a) is not consistent with Fig. 10.1., a), in which there is much more agreement between models and obs for the last decade. Please check this critical inconsistency. [Government of Germany]	Taken into account. Figure 9.8 has been completely revised and model data is now sampled to be consistent with observational coverage as in CH10.
9-1035	9	27	1	27	3	This statement seems to insinuate that the studies finding a statistically significant discrepancy between models and observations in the tropical troposphere did so by averaging across the observational series. But McKittrick et al. (2010; 2011-ref. in row 26) reported results for individual series as well as for multi-series averages. The discrepancy between models and observations is statistically significant either way: in the case of series averages and for every individual series as well. If you are going to mention the distinction then you need to mention these findings as well. [Ross McKittrick, Canada]	Rejected. The sentence makes it clear that it is the model-ensemble averaging that is meant here. No change.
9-1036	9	27	4	27	6	The idea that observational trends should be compared to the extrema of model trends, rather than to the confidence interval around the mean of model trends, is a statistically and methodologically incoherent. It is noteworthy that you have no supporting citations for this position. It amounts to a recommendation to engage in cherry-picking, and it is contradicted by your own methodologies elsewhere. You have already stated (p. 9_8) that a single model run can follow any one of many pathways. So to characterize the behaviour of a model you usually use ensemble means, presumably to draw out the underlying common aspects of the model runs in a forcing scenario. And you claim that the models are based on fundamental physical laws, implying that there is an underlying core theory common to the models. Presumably that core theory is revealed in the average behaviour across ensemble members. Yet here you say something different: that the proper way to compare models and observations is not to use the means but to use the endpoints of the full spread of the model runs. This is a bit too convenient: you can make that spread as wide as you like simply by adding more and more runs. Given enough runs, even from biased and incorrect models, eventually one will have a trend that coincides with the observed data. This proves nothing in a set-up where you can generate infinitely many model runs. It is not evidence in support of the "fundamental laws of nature" upon which the models are based, nor does it validate their parameterization and tuning, nor does it support anything to do with the models as a genre. All it says is that if you roll the dice often enough, eventually you get snake eyes. To say something about models as a group you have to test their average/common trend against the observational counterpart, which is precisely what McKittrick et al. (2010) does. Your argument, in effect, tries to have things both ways. You want to claim very high confidence in "the models" as a unified methodological entity or genre, but then you propose testing them as independent, atomistic single runs. Leaving aside the problem of cherry-picking, even if you got a perfect match between the observed trend and that from model run #1008, it tells you nothing whatsoever about the model as a scientific tool, because you are not testing the model, you are just testing a list of numbers that came out of it. If you want to draw a conclusion about the model, you have to treat it as a data generating process and test it as such, which means taking account of the distribution of what it produces, i.e. the moments. The paper you cite as supporting evidence for your method does not address the issues under discussion. In fact the logic of Bayesian Model Averaging goes completely counter to what you are proposing to do, since it is used to neutralize cherry-picking (or "model selection") bias in situations where researchers can pick from an extremely large number of models. [Ross McKittrick, Canada]	Taken into account. The references have now been repeated from the sentence preceding the sentence in question. It remains a fact that averaging over model realisations showing chaotic climate variability leads to a gross underestimate of natural variability in the result.
9-1037	9	27	6	27	10	This sentence makes it sound like the model-observation discrepancies were only found due to an improper uncertainty metric. No, the uncertainty metric was correct. The discrepancies were found because the models over-predict warming in the tropical troposphere, and robust trend estimators indicate that the difference is statistically highly significant, so that the models on average predict a trend that is significantly higher than any individual observational series or all observational series averaged together. You cite no published papers in	Taken into account. See response to 9-1036

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						support of your claim that a better method would use "the standard deviation or some other appropriate measure of ensemble spread." In fact you can't even say what alternative measure you would prefer! Much less do you cite a paper that argues for it and uses it. So you simply are not in a position to ignore or set aside the findings in the published literature. As stated above, if you are going to appeal to the spread of individual model runs then you have to abandon any claim to have validated the models as a genre, or as a collective embodiment of the core theory of how the climate works. To the extent you want readers to think of climate models as a single collective genre, as represented by a core set of processes "based on fundamental laws of nature" it makes sense to talk about the behaviour of the average of model runs. And, as you point out, on that basis the model-observation discrepancies are very highly significant. [Ross McKittrick, Canada]	
9-1038	9	27	6	27	17	This explanation is incorrect. Douglass et al. 2008 effectively removed the issue of differences in interannual variations by comparing the tropospheric trends in observations and models in a situation where BOTH MODELS AND OBSERVATIONS HAD THE SAME surface trend. Santer et al. 2008 continue to misrepresent what we did (and this carries forward here). (Santer et al. also used obsolete datasets.) Thus, no matter what sequence of El Ninos and La Ninas might have occurred, the fact that the surface trends were the same between models and observations was the key pre-condition for the test of agreement/disagreement between tropospheric trends. It was therefore obvious to use the standard error of the mean for this test since we wanted the best representation of models with which to compare. To say otherwise indicates a misunderstanding of the question being addressed in Douglass et al. 2008. Happily, this issue was further confirmed in Christy et al. 2010 where the direct comparison of amplification factors was done (completely eliminating the variability issue) and, again, models were shown to be significantly different from observations. Thus, in both cases (Douglass et al. 2008 and Christy et al. 2010) the concern about interannual noise was normalized away. This must be understood. What is TL and TM in line 17? Should these be LT and MT in line 17? [John Christy, United States of America]	Taken into account. It is not clear which of the many explanations in this paragraph are claimed to be incorrect. The comment appears to imply that the averaged tropospheric temperatures are completely determined by averaged SST, which is incorrect (e.g., Bengtsson and Hodges, 2011). TL and TM are typos, thank you.
9-1039	9	27	13	27	33	These paragraphs make clear that there is no apparent explanation for the tendency of models to over-predict warming in the tropical troposphere. So you shrug and say that the cause of the bias remains "elusive". Given that this section refers to the lower- and mid-troposphere in the tropics, which makes up about half the Earth's atmosphere and is the place where all the models predict the most rapid and pronounced GHG-induced warming should be observed, how can you then state at the beginning of the chapter (p. 9-3, lines 48-51) that you have "very high confidence" that the models correctly simulate the atmospheric response to greenhouse gas forcing? The most prominent feature of climate models' response to GHG forcing is rapid, amplified warming in the tropical troposphere. Yet in this section you have conceded that it's not even clear whether there has been any warming there at all, and to the extent there has been it is far less than what the models predicted in response to the observed increase in GHG levels, and the model-observational discrepancy is highly significant, and you have no explanation for it. I can't see how you nevertheless have "very high confidence" in the validity of models' handling of GHG forcing, especially in light of the similar discrepancies at the surface as shown in Figures 1.4 and 9.8, in which models predict a lot of warming that does not show up in the surface temperature record over the past 2 decades. How can you have "very high confidence" in models that get a key prediction wrong over multiple decades for reasons you are unable to explain? [Ross McKittrick, Canada]	Taken into account. The statement in the ES has been toned down. Surface temperature of the last 15 years is now dealt with in the new Box 9.2. The statement "you have conceded that it's not even clear whether there has been any warming there at all" is cherry-picking the WG1 text. The model-data discrepancy is exactly NOT highly significant although occurring in many model solutions, owing to natural and chaotic climate variability. A summary of potential causes for the mismatch are now given.
9-1040	9	27	16	27	16	change "(including the 95% confidence interval, (Fu et al., 2011))" to "(including the 95% confidence interval; Fu et al., 2011)" [SAMIR POKHREL, INDIA]	Editorial
9-1041	9	27	17	27	17	LT and MT [Peter Thorne, United States of America]	Editorial
9-1042	9	27	17	27	18	Correct: 'LT and MT'. Further, I am not sure how biases in SST's can be separated from atmospheric model errors? [Thorsten Mauritsen, Germany]	Taken into account. Sentence removed.
9-1043	9	27	17			explain abbreviation "TL" and "TM" [Barbara Früh, Germany]	Editorial. Typo corrected
9-1044	9	27	23	27	23	... of MT against LT variations... [European Union]	Editorial. Typo corrected
9-1045	9	27	23	27	23	LT and MT [Peter Thorne, United States of America]	Editorial. Typo corrected
9-1046	9	27	27	27	27	Pls put "." after "...statistically significant" [HASIBUR RAHAMAN, India]	Editorial. Typo corrected

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1047	9	27	31	27	32	I am not clear how the evidence can be deemed "robust" if there is only medium agreement. Surely the lack of agreement is an indicator of some problems either with the model or the uncertainty quantification? In addition, the object of this sentence "most, though not all, models overestimate the warming trend" is not very well defined. Is the clause "though not all" added to reflect the medium agreement, and in that case should there simply be high confidence? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. See response to 9-1014.
9-1048	9	27	31	27	33	It would be nice not to close with an ambiguous sentence here but rather to spell out very briefly what the potential causes are and why it is elusive. Otherwise it feels like a weak ending. [Peter Thorne, United States of America]	Noted. Text modified.
9-1049	9	27	33	27	33	The summary of 'quality low', 'confidence high', for UTT in Fig. 9.45 should be explained. It seems odd. The assessment should be added to the executive summary. What is the impact of this on the assessment of modelled sensitivity to GHG forcing? [Government of Australia]	Noted. The confidence level applies to the assessment, not to the model. Assessment will be elevated to Executive Summary. No published study exists linking fidelity of simulated tropospheric temperature trends to climate sensitivity.
9-1050	9	27	35	28	14	Important discussion of models' ability to simulate extratropical cyclones. Relevant for policy makers concerned with future storms. Needs to be emphasised in SPM. [European Union]	Noted - Inclusion in the SPM has been suggested
9-1051	9	27	42	27	43	A regional study by Zappa (2012) finds ... [Hai Lin, Canada]	Editorial
9-1052	9	27	42			There is an additional reference that should be considered in the discussion of extratropical cyclone tracks in CMIP5 models instead of only relying on Chang et al.: Colle et al. ("Historical and Future Predictions of Eastern North America and Western Atlantic Extratropical Cyclones in CMIP5 During the Cool Season."). This paper finds that resolution is important in terms of locating the storm track and properly simulating extratropical cyclone intensity. [Government of United States of America]	Taken into account - The paper has been assessed and its findings included as appropriate.
9-1053	9	27	45	27	45	...intensity. Chang et al. (2012) also found... [Hai Lin, Canada]	Editorial
9-1054	9	27	45	27	45	change "(Chang et al., 2012)" to "Chang et al., (2012)" [SAMIR POKHREL, INDIA]	Editorial
9-1055	9	27	45	27	45	Pls change "(Chang et al., 2012)" to "Chang et al., (2012)" [HASIBUR RAHAMAN, India]	Editorial
9-1056	9	27	53	27	53	"more strongly": more than what? [Bart Van den Hurk, Netherlands]	Taken into account - "more" has been removed
9-1057	9	27	54	27	54	Zappa 2012 should be Zappa et al. 2012 [HASIBUR RAHAMAN, India]	Editorial
9-1058	9	27	57	27	57	insert "in this area" after "extremes" [Bart Van den Hurk, Netherlands]	Accepted
9-1059	9	27	57	28	2	Recent studies are helping to address this knowledge gap, as follows. A reduction in the growth rate of the leading storm track mode crossing southern Australia has been identified, with most CMIP3 models capturing the associated changes in the zonal wind but only about a third capturing the changes in baroclinic instability (Frederiksen, C. S., J. S. Frederiksen, J. M. Sisson, and S. L. Osbrough, 2011: Australian winter circulation and rainfall changes and projections. Int. J. Climate Change Strategies Management, 3, 170-188, doi:10.1108/17568691111129002). Another recent study found significant differences between CMIP3 models in their ability to represent extratropical cyclones in the western South Pacific, including differences in their ability to accurately represent the seasonal and spatial variability of these storms (Dowdy, A. J., Mills, G. A., Timball, B. and Wang, Y., 2012: Changes in the risk of extratropical cyclone occurrence in eastern Australia. Journal of Climate, doi:10.1175/JCLI-D-12-00192.1, in press.). [Government of Australia]	Taken into account - The papers have been assessed and their findings included as appropriate.
9-1060	9	28	1	28	2	Suggest adding the following: However, recent studies have shown that only about a third of the CMIP3 models were able to capture the observed changes and trends in Southern Hemisphere baroclinicity responsible for a reduction in the growth rate of the leading winter storm track modes (Frederiksen et al., 2010; Frederiksen et al., 2011a). Those models that were able to capture the changes only showed about half of the magnitude of the observed changes and trends. Another study (Frederiksen et al., 2011b) showed that when the atmospheric component of models was forced with greenhouse gas and observed sea surface temperature forcing the models were able to simulate the observed changes, including the magnitude of the changes. An evaluation of the CMIP5	Taken into account - The papers have been assessed and their findings included as appropriate.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						models has yet to be completed in this regard. [Jorgen Frederiksen, Australia]	
9-1061	9	28	1	28	2	Suggest adding the following: However, recent studies have shown that only about a third of the CMIP3 models were able to capture the observed changes and trends in Southern Hemisphere baroclinicity responsible for a reduction in the growth rate of the leading winter storm track modes (Frederiksen et al., 2010; Frederiksen et al., 2011a). Those models that were able to capture the changes only showed about half of the magnitude of the observed changes and trends. Another study (Frederiksen et al., 2011b) showed that when the atmospheric component of models was forced with greenhouse gas and observed sea surface temperature forcing the models were able to simulate the observed changes, including the magnitude of the changes. An evaluation of the CMIP5 models has yet to be completed in this regard [Government of Australia]	Taken into account - The papers have been assessed and their findings included as appropriate.
9-1062	9	28	1		2	Suggest adding the following: "However, recent studies have shown that only about a third of the CMIP3 models were able to capture the observed changes and trends in Southern Hemisphere baroclinicity responsible for a reduction in the growth rate of the leading winter storm track modes (Frederiksen et al., 2010; Frederiksen et al., 2011a). Those models that were able to capture the changes only showed about half of the magnitude of the observed changes and trends. Another study (Frederiksen et al., 2011b) showed that when the atmospheric component of models was forced with greenhouse gas and observed sea surface temperature forcing the models were able to simulate the observed changes, including the magnitude of the changes. An evaluation of the CMIP5 models has yet to be completed in this regard." References: Frederiksen, C.S., J.S. Frederiksen, J.M. Sisson and S.L. Osbrough, 2011a: Australian winter circulation and rainfall changes and projections. Int. J. Clim. Change Strat. Mang., 3, Issue 2, 170-188. Frederiksen, C.S., J.S. Frederiksen, J.M. Sisson and S.L. Osbrough, 2011b: Changes and Projections in the Annual Cycle of the Southern Hemisphere Circulation, Storm Tracks and Australian Rainfall. Int. J. Clim. Change Impacts Responses, 2, 143-162. Frederiksen, J.S., C.S. Frederiksen, S.L. Osbrough and J.M. Sisson, 2010: Causes of changing Southern Hemispheric weather systems. Chapter 8, Managing Climate Change, Eds. I. Jupp, P. Holper and W. Cai, CSIRO Publishing, pp85-98. [Carsten Frederiksen, Australia]	Noted - As the report provides an assessment on is not a review and as the main focus is on the new CMIP5 ensemble we only included the first of the references.
9-1063	9	28	4	28	14	This is an interesting statement, but what does it mean for our confidence in future projections? Models have been tuned to the present (e.g. QBO periodicity), does it work for projections? [Peter Braesicke, United Kingdom]	Noted - This paragraph is about assessing model errors. It cannot singularly address this question, which is discussed in various places in the chapter
9-1064	9	28	4	28	14	Biases and trends in cimp5 eddy-driven jet is also analysed in Anstey et al., submitted [Chiara Cagnazzo, Italy]	Noted - there was insufficient information to find the referenced paper in the accepted literature
9-1065	9	28	4	28	14	suggest to refer here to Chapter 2 [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - a reference has been added
9-1066	9	28	5	28	5	"earlier studies": which ones? [Annalisa Cherchi, Italy]	Taken into account - A reference has been added
9-1067	9	28	8	28	8	after 'extratropics', please add a reference [Chiara Cagnazzo, Italy]	Taken into account - A reference has been added
9-1068	9	28	10	28	14	Suggest rewriting in the following way: More recently, it has been noted that the degree of poleward shifting of the jets may be systematically affected by model error (Kidston and Gerber, 2010, for the eddy-driven polar jet; Frederiksen et al., 2010, for the Southern Hemisphere winter-time subtropical jet) with some recent studies indicating that the previous consistency between models may still not indicate robust response due to shared limitations in model domain and associated processes occurring in the middle atmosphere (Morgenstern et al., 2010b; Scaife et al., 2010). [Jorgen Frederiksen, Australia]	Noted - As the 2010 reference is a CSIRO publication it is unclear if it was peer-reviewed. As other references make similar points (e.g., Ceppi et al. 2012) we decided not to include the reference.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1069	9	28	10	28	14	Suggest rewriting in the following way: More recently, it has been noted that the degree of pole ward shifting of the jets may be systematically affected by model error (Kidston and Gerber, 2010, for the eddy-driven polar jet; Frederiksen et al., 2010, for the Southern Hemisphere winter-time subtropical jet) with some recent studies indicating that the previous consistency between models may still not indicate robust response due to shared limitations in model domain and associated processes occurring in the middle atmosphere (Morgenstern et al., 2010b; Scaife et al., 2010). References: Frederiksen, C.S., J.S. Frederiksen, J.M. Sisson and S.L. Osbrough, 2011a: Australian winter circulation and rainfall changes and projections. Int. J. Clim. Change Strat. Mang., 3, Issue 2, 170-188. Frederiksen, C.S., J.S. Frederiksen, J.M. Sisson and S.L. Osbrough, 2011b: Changes and Projections in the Annual Cycle of the Southern Hemisphere Circulation, Storm Tracks and Australian Rainfall. Int. J. Clim. Change Impacts Responses, 2, 143-162. Frederiksen, J.S., C.S. Frederiksen, S.L. Osbrough and J.M. Sisson, 2010: Causes of changing Southern Hemispheric weather systems. Chapter 8, Managing Climate Change, Eds. I. Jupp, P. Holper and W. Cai, CSIRO Publishing, pp85-98 [Government of Australia]	Noted - As the 2010 reference is a CSIRO publication it is unclear if it was peer-reviewed. As other references make similar points (e.g., Ceppi et al. 2012) we decided not to include the reference.
9-1070	9	28	10	28	14	Karpechko and Manzini (2012) have shown that models without well-resolved stratosphere may underestimate the Brewer-Dobson circulation response to greenhouse gas concentration increases, which would result in overestimation of the polarward shift response of the mid-latitude jet in the Northern Hemisphere. This study can be cited here. Missing reference: Karpechko A. Yu., and E. Manzini, Stratospheric influence on tropospheric climate change in the Northern Hemisphere (2012), J. Geophys. Res., 117, D05133, doi:10.1029/2011JD017036. [Alexey Karpechko, Finland]	Taken into account - The paper has been assessed and its findings included as appropriate.
9-1071	9	28	10		14	Suggest rewriting in the following way: "More recently, it has been noted that the degree of poleward shifting of the jets may be systematically affected by model error (Kidston and Gerber, 2010, for the eddy-driven polar jet; Frederiksen et al., 2010, for the Southern Hemisphere winter-time subtropical jet) with some recent studies indicating that the previous consistency between models may still not indicate robust response due to shared limitations in model domain and associated processes occurring in the middle atmosphere (Morgenstern et al., 2010b; Scaife et al., 2010)." Reference: Frederiksen, J.S., C.S. Frederiksen, S.L. Osbrough and J.M. Sisson, 2010: Causes of changing Southern Hemispheric weather systems. Chapter 8, Managing Climate Change, Eds. I. Jupp, P. Holper and W. Cai, CSIRO Publishing, pp85-98. [Carsten Frederiksen, Australia]	Noted - As the 2010 reference is a CSIRO publication it is unclear if it was peer-reviewed. As other references make similar points (e.g., Ceppi et al. 2012) we decided not to include the reference.
9-1072	9	28	11	28	11	"affected by model" -> "in" [Bart Van den Hurk, Netherlands]	Editorial
9-1073	9	28	12	28	13	This is an important general point - there is no way to distinguish, a priori, between "robustness" and "shared limitations". Could this be mentioned in FAQ 9.1? The step from "robustness" to confidence is used throughout this report and it would be helpful to discuss in more detail the justification. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account - general limitations to model evaluation are now discussed in Section 9.2.3
9-1074	9	28	13	28	13	how can "limitations in model domain" be applicable to global models? [Bart Van den Hurk, Netherlands]	Taken into account - this refers to the vertical domain, which has been clarified.
9-1075	9	28	16	28	25	suggest to refer here to Chapter 2 [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - a reference has been added
9-1076	9	28	18			This should add "interannual and decadal variability" - please add "interannual". - some satellite records are less than a decade so interannual is still a substantive concern. [duane waliser, United States of America]	Accepted
9-1077	9	28	29	28	30	"..observational support for this is found from ocean salinity observations (Durack et al., 2012)" I would also add the references: Boyer et al., 2005 (Geophys Res Lett), Hosoda et al., 2009 (Journal of Oceanography), Durack & Wijffels 2010 (Journal of Climate). It would also be fair to state "..found from ocean near-surface and subsurface salinity observations.." [Paul Durack, United States]	Noted - As this report is not a review not all papers referring to a particular point need to be cited. The 2012 reference provides sufficient evidence for the statement made.
9-1078	9	28	30	28	30	insert "and" before "precipitation" [Bart Van den Hurk, Netherlands]	Accepted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1079	9	28	30	28	33	The final sentence ("There is medium evidence...") is confusing. How can there be medium confidence in model performance if there is limited agreement on observed trends (i.e. the target)? [Josephine Brown, Australia]	Taken into account. Text made more clear.
9-1080	9	28	31	28	33	"Medium confidence that models correctly simulate precipitation increases..." could be changed to "Medium confidence that models are correct in their simulation of precipitation increases..." (Current wording implies that we know what the precipitation change will be and are checking the models against it, rather than the intended meaning that we are assessing the projection). Use of the word "correct" is also optimistic - perhaps use "realistic" as above, once that has an accepted definition? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text made more clear.
9-1081	9	28	32	28	32	what is the relevance of "high agreement among models" in this context? [Bart Van den Hurk, Netherlands]	Taken into account. Text made more clear.
9-1082	9	28	32	28	33	Concerning "limited evidence that this is been detected in observed trends" please give a reference to Chapter 2.5.1. [Government of Germany]	Accepted
9-1083	9	28	33	28	33	Change "this is been detected" to "this has been detected" [Hai Lin, Canada]	Accepted
9-1084	9	28	35	28	47	What is written here is OK, but the relationship in other chapters of this WG1 assessment is quoted as that between total column water vapour (or precipitable water) and sea-surface temperature, and the correlation of TCWV variations from either the RSS SSML retrievals or the from ECMWF AMIP runs I have checked is particularly strong with the saturation specific humidity computed using sea-surface temperature (see also Uppala et al., 2005 for discussion of this in the context of ERA-40). A direct link of column water vapour with SST rather than lower tropospheric temperature is a more physical thing to look for, in view of the evaporative source of water vapour over the oceans. I thus wonder why the link with SST is not discussed here. [Adrian Simmons, United Kingdom]	Taken into account. The Figure is an update of Mears et al. (2007), which is why it is based on lower tropospheric temperature rather than SST, although it is now noted that the two are highly correlated.
9-1085	9	28	38	28	38	"(updated from Maers al, 2007) could be omitted here since it is already mentioned in the figure caption. [Farahnaz Khosrawi, Sweden]	Editorial
9-1086	9	28	40	28	41	The sentence is a bit complicated to read as it is now and I would suggest to put the text in parantheses into the figure caption. [Farahnaz Khosrawi, Sweden]	Editorial
9-1087	9	28	43	28	43	The abbreviations RSS and UAH should be defined and explained, either directly here or by a reference to another chapter or section. [Andreas Sterl, Netherlands]	Taken into account. Acronyms now clarified in Figure caption.
9-1088	9	28	43	28	43	what is "RSS"? [Bart Van den Hurk, Netherlands]	Taken into account. Acronyms now clarified in Figure caption.
9-1089	9	28	44			explain abbreviation "UAH" [Barbara Früh, Germany]	Taken into account. Acronyms now clarified in Figure caption.
9-1090	9	28	45	28	45	Clarify what is meant by "indicating that these results are inconsistent with model physics". [Government of Australia]	Taken into account. Wording changed
9-1091	9	28	45	28	45	"the line". Which line? Be more precise. [Farahnaz Khosrawi, Sweden]	Editorial
9-1092	9	28	45	28	45	Change "these results are inconsistent with model physics" to "model physics is inconsistent with these results"? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Wording changed
9-1093	9	28	45	28	47	The sentence "All of the observational and reanalysis points lie at the lower end of the model distribution, consistent with the findings of Santer et al. (2012)." is incomprehensible [Government of Germany]	Taken into account. Text made more clear.
9-1094	9	28	45			- I do not understand why it is possible to conclude from the far distance of the reanalysis points to the line "that these results are inconsistent with model physics"; - please explain the impact of inconsistency [Barbara Früh, Germany]	Taken into account. Wording changed
9-1095	9	28	46			"the points lie at the lower end of the model distribution" what does it mean and why it is important? [Barbara Früh, Germany]	Taken into account. Comment addressed.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1096	9	28	49	28	54	Why just 19 CMIP5 models are selected? What criterion is used for this selection? I suggest adding as many as possible other CMIP5 models, including other 4 models from China (BNU-ESM, FGOALS-g2, FGOALS-s2 and FIO-ESM) that are not used in this figure. Please revise this figure. [Bin Wang, China]	Taken into account. One realization from many models available were used for the final version of this figure. Some were not included because accurate splicing between "historical" and "RCP8.5" was not available in time.
9-1097	9	29	1	29	23	Link and align with 9-28 / 4-14; I don't agree with the statement that the models with interactive chemistry agree well [Peter Braesicke, United Kingdom]	Accepted. The sentence was referring to the multi-model mean which has been clarified in the revised version.
9-1098	9	29	1			This section could be expanded just a bit to consider the middle and upper stratosphere, referencing the following new publication. Thompson, DWJ, DJ Seidel, WJ Randel, C-Z Zou, A Butler, C Mears, A Osso, C Long and R Lin, The mystery of recent stratospheric temperature trends, Nature, in press. [Dian Seidel, United States of America]	Rejected. Because of large uncertainty in mid-and upper stratosphere temperature trends the agreement between observed and simulated trends is not assessed (see Chapter 2).
9-1099	9	29	21	29	23	"... agree well with observations, although some models show substantial deviations". Needs more specific explanation! What do you mean? Some models are outside of the blue area - is this what you mean by "substantial deviations"? The observed trend (from 1970 to 1995 for the global mean) is neither matched by the CMIP5Chem multi-model mean nor any single model; the increase of ozone after 1995 is also clearly underestimated by all models. For the south polar region: The agreement regarding the trend is much better here, but the year-to-year variability is clearly overestimated by many models. [Martin Dameris, Germany]	Noted. For details we refer to the paper where deviations for individual models are discussed in more detail.
9-1100	9	29	21	29	23	From the Figure 9.10 it appears that the total column of ozone that is prescribed (in models without chemistry) agrees less with observations than the total column that is simulated with the chemistry models. Is there a reason for the underestimation of the prescribed ozone climatology? [Government of Germany]	Accepted. This has been clarified in the caption of Figure 9.10: "Note that the IGAC/SPARC dataset over Antarctica (and thus the majority of the NOCHEM models) is based on ozonesonde measurements at the vortex edge (69°S) and as a result underestimates Antarctic ozone depletion compared to the observations shown."
9-1101	9	29	38	30	2	I would start with an explanation why ozone is important for climate before going into detail, merge with everything before, make it shorter and more to the point; it will help the reader! [Peter Braesicke, United Kingdom]	Accepted and changed as suggested.
9-1102	9	29	42	29	42	The figure indexing in Chapter 2 has been changed as a result of Chapter restructuring [Peter Thorne, United States of America]	Accepted
9-1103	9	29	45			There are other papers which must be cited here: e.g. Ramaswamy et al. (2006, Science 311) and especially Dall'Amico et al. (2010; Clim. Dyn. 34) - see also the very detailed discussion in Chapter 4 of WMO (2011). [Martin Dameris, Germany]	Noted. These references are cited in Chapter 10 (section 10.3.1.2.2) where the attribution of temperature trends is assessed.
9-1104	9	29	48	29	48	et al. Instead of and co-authors [Farahnaz Khosrawi, Sweden]	Editorial change considered
9-1105	9	29	54	29	56	Among the potential causes there are also changes in stratospheric circulations [Chiara Cagnazzo, Italy]	Noted. However, circulation changes might be associated with temperature errors, but they cannot cause them. No change.
9-1106	9	29	56	29	56	et al. Instead of and co-authors [Farahnaz Khosrawi, Sweden]	Editorial
9-1107	9	30	2	30	2	et al. Instead of and co-authors [Farahnaz Khosrawi, Sweden]	Editorial
9-1108	9	30	4	30	11	suggest to refer here to Chapter 2 [Thomas Stocker/ WGI TSU, Switzerland]	Accepted and reference to Chapter 2 added in the first paragraph of this section.
9-1109	9	30	8			"... agrees generally well with satellite observations and ozone sondes...biases exist for individual models". See my comment #8. Such statements are worthless. Concrete examples must be given to make the statement clear. [Martin Dameris, Germany]	Rejected. For details we refer to the paper where deviations for individual models are discussed in more detail.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1110	9	30	13	30	13	This section would be easier to follow if mid-Holocene and LGM were discussed separately. Suggest two sub-sections. [Josephine Brown, Australia]	Taken into account. However the two time periods are included in the same subsection to provide an overall assessment based on the results of different periods
9-1111	9	30	13	30	13	9.4.1.4 Model Simulations of the Last Glacial Maximum and the Mid-Holocene should be: 9.4.1.5 Model Simulations of the Last Glacial Maximum and the Mid-Holocene [Koichi Sakaguchi, the United States]	editorial. Taken into account
9-1112	9	30	13			Explain upfront why this might improve our confidence in model projections. [Peter Braesicke, United Kingdom]	Taken into account an text modified accordingly
9-1113	9	30	21	30	21	Insolation seasonality changes in both NH and SH in mid-Holocene. Clarify that lack of data in SH means that the validation is focused on NH only. [Josephine Brown, Australia]	taken into account. However the limit space do not allow to explain everything. Reference to chapter 5 is also done
9-1114	9	30	26	30	26	change “(PMIP, (Joussaume and Taylor, 1995), (Braconnot et al., 2007c).” to “(PMIP; Joussaume and Taylor, 1995; Braconnot et al., 2007c). [SAMIR POKHREL, INDIA]	editorial
9-1115	9	30	28	30	45	In this paragraph (and entire section), it is not clear whether mid-Holocene or LGM is being discussed. For this reason, it would be much clearer to separate the two discussions. [Josephine Brown, Australia]	taken into account. The text has been clarified to better identify which period is considered
9-1116	9	30	28	30	45	Are the uncertainties for the MH simulations ascribed as well to the lack of vegetation feedbacks? [Annalisa Cherchi, Italy]	Noted. A subgroup of models have interactive vegetation and carbon cycle. These uncertainties are considered as well as the uncertainties in the past climate reconstructions from observations
9-1117	9	30	31	30	31	change “Hargreaves et al (Hargreaves et al., 2011)” to “Hargreaves et al (2011)” [SAMIR POKHREL, INDIA]	Editorial. Corrected
9-1118	9	30	31	30	31	Pls change "Hargreaves et al (Hargreaves et al., 2011)" to "Hargreaves et al., (2011)" [HASIBUR RAHAMAN, India]	Editorial. Corrected
9-1119	9	30	42			Actually, Annan and Hargreaves (CPD, 2012) show that the land-ocean contrast in the proxy data is higher than for all of the (PMIP2) models that they considered, albeit not by a huge margin. [James Annan, Japan]	taken into account, even though given space limit the paragraph has been reduced
9-1120	9	30	47	30	51	Regional model-data comparison for the LGM and mid-Holocene climate changes over China should be added in this paragraph. It had been revealed that 25 PMIP1 and PMIP2 models underestimated surface cooling at the LGM (Jiang et al., 2011), and 36 PMIP1 and PMIP2 models could not reproduce annual and winter warming as suggested by multi-proxy data over the country at the mid-Holocene (Jiang et al., 2012). These results are directly relevant to the theme of discussion here and provide additional information to authors. Hence they should be addressed, for example, as “In China, models underestimated the reconstructed cooling at the LGM (Jiang et al., 2011) and could not reproduce annual and winter warming as suggested by multi-proxy data at the mid-Holocene (Jiang et al., 2012)”. This comment may be added before “In the southern Hemisphere” on line 49 of Page 30. [References: (1) Jiang, D., Lang, X., Tian, Z., and Guo, D., 2011. Last glacial maximum climate over China from PMIP simulations. <i>Palaeogeography, Palaeoclimatology, Palaeoecology</i> , 309: 347–357. (2) Jiang, D., Lang, X., Tian, Z., and Wang, T., 2012. Considerable model–data mismatch in temperature over China during the mid-Holocene: Results of PMIP simulations. <i>Journal of Climate</i> , 25: 4135–4153.] [Dabang Jiang, China]	Taken into account. However given the limited space it is not possible to site all the literature on regional analyses in this chapter, and additional reference are considered in chapter 5 (paleo chapter)
9-1121	9	30	54	30	54	Pls change "(Harrison et al., 2012)" to "Harrison et al., (2012)" [HASIBUR RAHAMAN, India]	Editorial. Corrected
9-1122	9	31	4			Consider mentioning that resolution also does not either seem to matter. [Thorsten Mauritsen, Germany]	taken into account. Note however that due to space limit we had to simplify and shortened slightly the paragraph.
9-1123	9	31	5	31	6	"some consistently reproduce past changes better than others" -- we suggest to guide the reader more through these two rather rich figures 9.11 and 9.12. It would be helpful to briefly indicate where this can be derived from [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. This has been rewritten
9-1124	9	31	9	31	38	Figure 9.11 and 9.12: These figures are very complicated (really many figures in one). Can they be simplified to convey the key points only? [Josephine Brown, Australia]	Taken into account. The figures have been updated. In particular figure 12 has been simplified and the

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							figure caption better explain the normalisation that has been applied to the data
9-1125	9	31	25	31	38	Figure 9.12 uses some incorrect model names, such as FGOALS2 and FGOALS1. FGOALS model system has two versions in PCMDI website (see http://cmip-pcmdi.llnl.gov/cmip5/availability.html): FGOALS-g2 and FGOALS-s2. Please make sure what do FGAOLS2 and FGOALS1 mean respectively and then revise the figure with correct model names. [Bin Wang, China]	Taken into account. Standard model names are now used in the legend.
9-1126	9	31	25			Figure 9.12: Please also ensure that variables/acronyms used along the top axis of the plots are described in the caption. In addition, it is necessary to introduce methods like the Kendall rank correlation in the caption, or are/can these be introduced in the main text? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. The figure has been simplified and the different variables explains in the figure caption.
9-1127	9	31	26	21	38	labels in the figure are too small. Even with glasses I cannot read them. [Andreas Sterl, Netherlands]	Taken into account. The figure has been updated accordingly
9-1128	9	31	26	31	38	This graphic needs more explanation. What to the acronyms SSTann, MAT, MTCO, MTWA, GDD5, MAP and alpha mean? What do the signs indicate? [Government of Germany]	Taken into account, The figure has been simplified and the acronymes expanded
9-1129	9	31	26	31	38	Fig. 9.12 is hard to understand and interpret. What are the variables (columns)? [Government of United States of America]	taken into account. See 9-1128
9-1130	9	31	26			In the legend of Figure 9.12, the reader does not know the meaning and name of the variables MAT, MTCOPlease explain [Government of Spain]	taken into account. See 9-1128
9-1131	9	31	40	32	9	It feels very odd that there is no bullet here for tropospheric data agreement and I suspect that this will get picked up on. Why is nothing pulled through from that section? [Peter Thorne, United States of America]	Accepted, thank you. Added.
9-1132	9	31	42	31	43	Same remark as for point #3. I would be a « sceptical », I would really use this kind of statement to attack the community of climate modelers. [Frédéric HOURDIN, France]	Noted. Unclear what is meant. The sentence is very strongly backed by existing literature. No change.
9-1133	9	31	42	32	9	In the summary for "Atmosphere", we recommend mention the simulations of clouds and water vapor, especially improvements from CMIP3 to CMIP5, and current model spread. [Government of United States of America]	Taken into account -The summaries have been moved to the end of the cloud section (9.4.1.2). Progress is mentioned and the statement has been made consistent with one added to the Executive Summary.
9-1134	9	31	44	31	45	High confidence with modest agreement ? [Peter Stott, United Kingdom of Great Britain & Northern Ireland]	Noted. Consistent with evidence and agreement among studies -- sometimes we know with high confidence that models do not do a very good job. No change.
9-1135	9	31	44	31	47	This sentence ("There is also high confidence...") is very confusing! [Josephine Brown, Australia]	Noted. Consistent with evidence and agreement among studies -- sometimes we know with high confidence that models do not do a very good job. No change.
9-1136	9	31	45	31	45	Terms such as "modest agreement", "reasonable agreement", "fairly well" etc, are used throughout Chapter 9 when comparing models with observations. It is not clear how these terms rate sequentially - is 'modest agreement' a higher level of agreement than 'reasonable agreement'? It would be good if more quantitative terminology could be used where ever possible. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. The entire chapter has been carefully scanned to make the language more quantitative.
9-1137	9	31	47	31	51	There remain significant errors in the model simulation of clouds and water vapor in CMIP5, especially in the upper troposphere, where the spread among models and their differences with observations are the largest (Jiang et al. 2012). Since the upper troposphere clouds contributes only ~2-3 W/m2 net cloud forcing (Su et al. 2006a), comparing to the total cloud forcing ~-20W/m2 (Stephens et al. 2012), the large errors in simulations the upper troposphere clouds cloud contribute ~15% of the cloud forcing errors. However, there is clear evidence that cloud simulations have improved in CMIP5, comparing to previous CMIP3 (Balcerak, 2012; Benka, 2012; Jiang et al. 2012). References: Benka, S.G., "Cloud simulations improving in climate models", Physics Today, Volume 65, Issue 10, October 2012.	Taken into account - This is extensively discussed in the relevant section (9.4.1.2) and does not need repeating here beyond the general statement made. The papers provided have been assessed and their findings included as appropriate.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						(http://www.physicstoday.org/daily_edition/physics_update/cloud_simulations_improving_in_climate_models) Jiang, J.H., H. Su, C. Zhai, V.S. Perun, A. Del Genio, L.S. Nazarenko, L.J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, M. , E.M. Volodin, T. Iversen, H. Drange, M.S. Mesquita, W.G. Read, J.W. Waters, B. Tian, J. Teixeira, and G.L. Stephens, "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	
9-1138	9	31	54	31	54	"Warming of recent decades" is not very explicit, and might be confusing given that the trend in observed global mean temperature is not significantly different from zero since 1998. It would be better to be explicit here and give a time frame, e.g., "Warming since 19XX" or "warming over the past XX years". [Thomas Stocker/ WGI TSU, Switzerland]	This discussion is now contained in Box 9.2 that specifically discusses the trends in recent decades.
9-1139	9	31	54			Why do the applied external forcings matter? What are the uncertainties? Scenario uncertainties are one problem, but do you mean how boundary conditions are treated/implemented in models? [Peter Braesicke, United Kingdom]	This discussion is now contained in Box 9.2 which discusses each of these questions more directly. The ability or inability to reproduce the historical record with climate model simulations can be due to a combination of factors that include the model formulation of external forcings as well as the representation of the forcing with a historical scenario. In this sense, model formulation includes how boundary conditions are treated or implemented within the model. The applied external forcings refer to the historical scenario. There are uncertainties in each of these as suggested and this can be clarified further in the text.
9-1140	9	31	59	31	60	I don't share the high confidence, why is prescribed ozone mentioned here? Hemispheric difference? [Peter Braesicke, United Kingdom]	Accepted. The assessment was based on the multi-model mean, but the spread among the individual models is large. Therefore this has been changed to medium agreement.
9-1141	9	31	59	31	60	"There is high confidence that the trends in stratospheric ozone, whether prescribed or calculated interactively, are generally in good agreement with observations, ..." ??? Prescribed ozone should agree per definition with observations!?! I do not understand the meaning of this statement. [Martin Dameris, Germany]	Noted. The Randel and Wu dataset is biased low over Antarctica why the models that use the Cionni et al dataset to prescribe ozone actually underestimate ozone compared to other observations. For details see Cionni et al. (2011) and Randel and Wu (2007). The summary statement has been changed.
9-1142	9	31				Figure 9.12: The Legend of the figure is illegible [Government of Germany]	taken into account. Figure caption revised
9-1143	9	31				Figure 9.11: labels of the panel b are written in a too small font [Government of Germany]	taken into account. Label enlarged
9-1144	9	31				ection 9.4.1.5: "From a global perspective there is high confidence that large scale patterns of surface temperature are well simulated by the CMIP5 models". Of course they are because (a) the surface fields will have been used to tune the models and (b) no major institution would have allowed its models to be included in CMIP5 unless it was pretty certain that was the case. "The CMIP5 models broadly capture the observed historical changes in global surface temperature ...". Of course they do for (a) the reasons given above and (b) because a simple radiation model generates warming with increased CO2. [David Webb, United Kingdom]	Noted - The severity of tests in the light of tuning has been discussed in Box 9.1 and in the new Section 9.2.3. It is worth noting that while global mean temperature might be tuned in some models - as tuning is poorly documented we cannot tell for sure - the pattern certainly is not.
9-1145	9	32	6	32	9	Why no mention of LGM in summary? [Josephine Brown, Australia]	Taken into account. Summary paragraph has been removed due to length limitations.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1146	9	32	6	32	9	Please add a summary for the last glacial maximum as well. [SYLVIE JOUSSAUME, France]	Taken into account. Summary paragraph has been removed due to length limitations.
9-1147	9	32	7	32	9	Does this "medium to poor" performance at regional levels not suggest that using these outputs as forcings for regional climate models (dynamical downscaling methods) may be unreliable? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Summary paragraph has been removed due to length limitations.
9-1148	9	32	11	33	32	I am missing a discussion on tropical SST biases here. I am not an expert in this area, but via my colleagues I know that there are several recent papers discussing these issues, and that they are very relevant for regional climate. [Douglas Maraun, Germany]	Taken into account. Text modified to discuss tropical SST errors
9-1149	9	32	11	38	46	Chapter 9: Section 9.4.2 (Simulation of Recent and Longer Term Records in Global Models. Ocean). The document does not mention any assessment result concerning inland seas, e.g. the Mediterranean. It would be advisable to have it. [Government of Spain]	Rejected. This is an assessment, not a review.
9-1150	9	32	11			Ocean section. There is a very nice study on an evaluation of CMIP3 and CMIP5 ocean surface winds that also touches on biases in sea level distributions because of the biases in the winds. It seems very relevant to this section - but the wind part in itself should probably be addressed above in the atmospheric large-scale circulation section and then the sea level here in this section under sea level. I enclose the reference and abstract here: ----- Evaluation of CMIP3 and CMIP5 Wind Stress Climatology Using Satellite Measurements and Atmospheric Reanalysis Products Tong Lee, Duane E. Waliser, Jui-Lin F. Li, Felix W. Landerer, and Michelle M. Gierach; Journal of Climate, 2012. Wind stress measurements from QuikSCAT satellite and two atmospheric reanalysis products are used to evaluate the annual mean and seasonal cycle of wind stress simulated by CMIP3 and CMIP5 models. The ensemble CMIP3 and CMIP5 wind stresses are very similar to each other. Overall speaking, there is no significant improvement of CMIP5 over CMIP3. CMIP ensemble average zonal wind stress has eastward biases at mid-latitude westerly wind regions (30°-50°N and 30°-50°S, with CMIP being too strong by as much as 55%), westward biases in subtropical-tropical easterly wind regions (15°-25°N and 15°-25°S), and westward biases at high-latitude regions (poleward of 55°S and 55°N). These biases correspond to too strong anticyclonic (cyclonic) wind stress curl over the subtropical (subpolar) ocean gyres, which would strengthen these gyres and influence oceanic meridional heat transport. In the equatorial zone, significant biases of CMIP wind exist in individual basins. In the equatorial Atlantic and Indian Oceans, CMIP ensemble zonal wind stresses are too weak and result in too small east-west gradient of sea level. In the equatorial Pacific, CMIP zonal wind stresses are too weak in the central Pacific and too strong in the west. These biases have important implications to the simulation of various modes of climate variability originating in the tropics. The CMIP models as a whole over-estimate the magnitude of seasonal variability by almost 50% when averaged over the entire global ocean. The biased wind stress climatologies in CMIP models not only have implications to the simulated ocean circulation and climate variability, but other air-sea fluxes as well. [duane waliser, United States of America]	Taken into account. Reference added
9-1151	9	32	12	38	46	9.4 We recommend inclusion of a discussion of AOMIPS in this section; 9.4 - Page 32:12 to 38:46. [Government of United States of America]	Rejected. This MIP it is not used in further chapters and we could not identify a reference on how AOMIP specifically addresses climate change scenario.
9-1152	9	32	12	38	46	Would be appropriate to include a discussion of AOMIPS in this section; 9.4 - Page 32:12 to Page 38:46 [Government of United States of America]	Rejected. This MIP it is not used in further chapters and we could not identify a reference on how AOMIP specifically addresses climate change scenario.
9-1153	9	32	18	32	20	at least a reference for the equator specification should be added. Further, the inclusion of ocean analysis in the evaluation of model performances is per se an added value, even though it contains model biases, it is constrained from observations and it gives a global view of oceanic fields. [Annalisa Cherchi, Italy]	Taken into account. Reference to equator removed
9-1154	9	32	18	32	20	Reference needed for this sentence. [Ronald Stouffer, United States of America]	Taken into account. Reference to equator removed
9-1155	9	32	38	32	44	Concerning "The simulations cover the period 1975 to 2005 from available historical simulations, whereas the observations are from 1874 to 2008."; are biases calculated only for the period 1975 to 2005 or compared to	Taken into account. The model to observational comparison is made between the World Ocean Atlas

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						the full observational record? If the latter is the case, how does warming in the early part of the 20th century affect these results? [Government of Germany]	2009 (WOA09) climatology fields and a 30-year climatological mean from the available CMIP5 models. The data that the WOA09 is comprised of extends from 1874 through to 2008, however the bulk of the data and therefore the median time for each point of the gridded climatology is more representative of the 1980-2010 period (see Figure 2 Durack & Wijffels, 2010). For this reason the ~1980-2010 (obs) and 1975-2005 (model) comparison is considered the most reasonable period in which to directly compare. We have updated the caption to more accurately convey this information. See also Chap. 3.
9-1156	9	32	38			Fig. 9.13: Again here the comparison between model mean and a single obs may be problematic The question is whether this difference is large compared to internal variability. One way to show this would be a second row where only the points are shown as color where the difference is significantly greater than the difference between two simulations of the same model. [Reto Knutti, Switzerland]	Rejected. Defining internal variability for ocean observations with any confidence is almost impossible given the data sparsity. I believe what is already a fairly complicated plot would become even more so with the addition of stippling/masking without any addition of meaningful information. The intention of this graphic is to convey qualitatively, how the CMIP5 model ensemble mean compares to an estimate of the observed mean state, any addition of significance would need considerable supporting text to explain and justify how stippling/masking was undertaken, whilst not really gaining any additional information.
9-1157	9	32	47	32	48	What counts as "serious bias"? Is there a danger of benchmarking against previous models rather than deciding separately what would constitute "serious bias"? Surely in high latitudes even a fraction of a degree is quite important as it can mean the difference between sea ice and no sea ice, and over large areas this is known to influence regional circulation. At lower latitudes perhaps "serious bias" would have a wider tolerance. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text modified to be clearer.
9-1158	9	32	47			I would insist on the very weak (no) improvement in the tropical mean SST (see point #5). [Frédéric HOURDIN, France]	Taken into account. Text modified to discuss tropical SST errors
9-1159	9	32	48	32	48	... fewer individual models... [European Union]	Editorial. Taken into account
9-1160	9	32	48	32	48	delete „s“ in „individuals“ [Douglas Maraun, Germany]	Editorial. Taken into account
9-1161	9	32	55			In regards to the Waliser et al. salinity reference. I submitted that as a "expedited contribution" to J.Climate well before the deadline; it had enough reviewer concerns/comments/suggetions - including suggestions for material that would make it more than an expedited contribution that it was recommended for resubmission. I haven't done that yet and thus I am not sure if this article will be resubmitted soon, what the deadline for this case would be, and based on the above whether it can legitimately be referenced. I would welcome the lead-author's advice on this. Thank you [duane waliser, United States of America]	Taken into account. Reference removed
9-1162	9	32	57	32	57	Nested parenthesis are awkward [Government of United States of America]	Editorial. Taken into account
9-1163	9	32	57	32	57	Pls add " and over north Indian Ocean (Pokhrel et al. 2012)" after "(e.g., double ITCZ in the East Pacific, (Lin, 2007))." The reference is as follows Pokhrel S, Rahaman SH, Parekh A, Saha SK, Dhakate A, Chaudhari HS and Gairola RM (2012), Evaporation-precipitation variability over Indian Ocean and its assessment in NCEP Climate Forecast System (CFSv2), Climate Dynamics, 39(9), 2585-2608, DOI: 10.1007/s00382-012-1542-6 [HASIBUR RAHAMAN, India]	Taken into account. Reference added
9-1164	9	32	57	32	58	Add "and also controls to some extent the model SST error (Pokhrel et al. 2012)" after the last line before fullstop. The reference is as follows Pokhrel S, Rahaman SH, Parekh A, Saha SK, Dhakate A, Chaudhari HS and Gairola RM (2012), Evaporation-precipitation variability over Indian Ocean and its	same as above

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						assessment in NCEP Climate Forecast System (CFSv2), Climate Dynamics, 39(9), 2585-2608, DOI: 10.1007/s00382-012-1542-6 [SAMIR POKHREL, INDIA]	
9-1165	9	32	57			Nested parentheses are awkward. [Government of United States of America]	Editorial. Taken into account
9-1166	9	32				Figure 9.13: The white contours seem unnecessary and distracting, and in any case are not mentioned in the caption. [William Merryfield, Canada]	Taken into account. Figure modified
9-1167	9	33	27	33	32	Figure 9.15 uses very few CMIP5 models. No CMIP5 models from China are used in this figure. Please revise the figure and add more CMIP5 models including FGOALS-g2, FGOALS-s2, BCC-CSM1.1, BNU-ESM and FIO-ESM to it. By the way, there are 18 points in the figure, while the x-axis lists only 16 models. [Bin Wang, China]	Taken into account. Figure removed
9-1168	9	33	36	35	58	See comment above about relevant Lee et al study - illustrates bias in mean sea level, and thus thermocline state, in ALL tropical ocean basins due to biases in surface winds - and thus the atmosphere. [duane waliser, United States of America]	Taking into account. The Lee et al. study is assessed.
9-1169	9	33	45	33	45	Terms such as "modest agreement", "reasonable agreement", "fairly well", "reasonably well" etc, are used throughout Chapter 9 when comparing models with observations. It is not clear how these terms rate sequentially - is 'modest agreement' a better result than 'reasonable agreement'? It would be good if more quantitative terminology could be used where ever possible. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text has been modified with appropriate calibrated language.
9-1170	9	33	47			correct reference to missing section 9.4.1.2 [Barbara Früh, Germany]	Editorial
9-1171	9	33	52	33	53	Please provide more detail as to why SSH is simulated better here. [Government of United States of America]	Taken into account. An explanation of why SSH is more realistically simulated in included.
9-1172	9	33	52	33	53	Give more detail as to why SSH is simulated better here [Government of United States of America]	Taken into account. An explanation of why SSH is more realistically simulated in included.
9-1173	9	33	52	33	53	Specifically, section 13.3.1 for global mean thermal expansion and 13.6.1 about regional patterns. [Jonathan Gregory, United Kingdom]	Taken into account. Appropriate reference is now made to CH13.
9-1174	9	33	55	34	3	Please add more CMIP5 model results to Figure 9.16, e.g., the results of FGOALS-g2, FGOALS-s2, BNU-ESM and FIO-ESM. [Bin Wang, China]	Taken into account. Figure has been updated, but only includes results from models where all data necessary is available.
9-1175	9	33	56	34	3	Please indicate in caption that the individual grey circles indicate the individual CMIP3 models. "[...] CMIP3 models (individual grey filled circles) [...]" [Government of Germany]	Editorial
9-1176	9	33	58	34	1	"root mean square errors": what is shown is the root-mean square discrepancy between observations and models. Climate models are not expected to track the natural variability component of observations, so a discrepancy is not an error. [Martin Juckes, United Kingdom]	Rejected. Figure represent the mean climate (~25 year period) so natural variability plays a very small role.
9-1177	9	33	58	34	1	Figure 9.16 caption: I suppose the root-mean square errors are also normalized by observed standard deviation, so probably better to mention it in the caption by revising the sentence as: "The root-mean square errors are also normalized by the observed standard deviation, and their size is indicated by the dashed grey circles about the observational point. [Koichi Sakaguchi, the United States]	Taken into account. Figure has been updated, and text clarified
9-1178	9	33				Figure 9.16: The caption does not mention that the gray dots are CMIP3 models (presumably) [William Merryfield, Canada]	Taken into account. Figure has been updated, and text clarified
9-1179	9	33				Figure 9.16: I assume the gray filled circles are the CMIP3 results. Are they going to be removed when CMIP5 results are all available? I suggest to keep them for comparison, if the figure is not too busy with all the CMIP5 results. [Koichi Sakaguchi, the United States]	Taken into account. Figure has been updated, and text clarified
9-1180	9	34	1	34	2	Why do you call it global ocean when only the latitudes between 50N and 50S are considered? [Farahnaz Khosrawi, Sweden]	Taken into account. Model calculations are global.
9-1181	9	34	5	34	24	suggest to refer here to Chapter 3 for the assessment of observed ocean heat content anomalies [Thomas	Taken into account. Appropriate reference is now

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Stocker/ WGI TSU, Switzerland]	made to CH03.
9-1182	9	34	13	34	20	A clarification might be needed here: It is suggested that including volcanoes leads to more realistic ocean heat uptake but below, l.18-23, it is stated that ocean stratification leads to an overestimated heat uptake. Would that mean that in the models with volcanic eruptions, given that they have a biased stratification the correct OHC changes result from error compensations? [RYM MSADEK, United States of America]	Noted. This is an interesting hypothesis but has not been addressed in the literature.
9-1183	9	34	14	34	16	Sentence is not clear, please reframe the sentence [HASIBUR RAHAMAN, India]	Taken into account. Figure has been updated, and text clarified
9-1184	9	34	15	34	15	Is it true that Mk3.6 does not include volcanic forcings? It seems to cool in the right places. [Government of Australia]	Taken into account. Figure has been updated, and text clarified
9-1185	9	34	15	34	15	CSIRO-Mk3-6-0 CMIP5 runs do include volcanic forcing; see Rotstayn et al. (2012), in the reference list, and chapter 12. [Leon Rotstayn, Australia]	Taken into account. Figure has been updated, and text clarified
9-1186	9	34	19	34	22	The descriptions of the model errors here: models "less stratified than is observed" "transport heat downwards more efficiently than real ocean" "overestimate oceanic mixing efficiency" all suggest that the majority of models suffer from spurious diapycnal mixing. This underscores my earlier point that it is important to discuss the model discretizations: numerous studies have shown that spurious diapycnal mixing is to be expected from pressure coordinate models. For a recent examination of the spurious mixing in ocean GCMs, see Ilicak, M, Alistair Adcroft, Stephen M Griffies, and Robert W Hallberg, February 2012: Spurious diapycnal mixing and the role of momentum closure. Ocean Modelling, 45-46, doi:10.1016/j.ocemod.2011.10.003. [Sonya Legg, United States of America]	Taken into account. Figure has been updated, and text clarified, including additional assessment of literature addressing vertical mixing.
9-1187	9	34	22	34	22	TCR has not been defined to this point [Government of United States of America]	Editorial
9-1188	9	34	22	34	22	Yes, it would lead to an underestimate of TCR, but Kuhlbrodt and Gregory also find that ocean heat uptake efficiency explains very little of the model spread in TCR, as it also found by Tomassini et al (submitted) and maybe others. So in the end this is probably a small bias. Indeed, you say this on page 72 line 8, in effect. [Jonathan Gregory, United Kingdom]	Taken into account. Impact on TCR has been clarified..
9-1189	9	34	22	34	22	The abbreviation "TCR" has not yet been defined (only defined on p72). [Sonya Legg, United States of America]	Taken into account. Text has been clarified
9-1190	9	34	22			I think the term TCR has not been defined in this chapter yet. The term Transient Climate Response has been used p.32 l.14, so TCR could be added in parenthesis there. [RYM MSADEK, United States of America]	Taken into account. Text has been clarified
9-1191	9	34	34	34	39	Please add more CMIP5 model results to Figure 9.17, e.g., the results of FGOALS-g2, FGOALS-s2, BNU-ESM and FIO-ESM. [Bin Wang, China]	Taken into account. Figure has been updated, but only includes results from models where all data necessary is available.
9-1192	9	34	35	34	39	The y-axis label indicates ocean heat content instead of ocean heat content anomalies. [Government of Germany]	Taken into account. Figure corrected
9-1193	9	34	35			Figure 9.17: suggest to refer in the caption to Chapter 3 for the assessment of observed ocean heat content anomalies; how does the comment about the significant model drift affect the ES statement about "models realistically simulating trends on Ocean Heat Uptake"? It would be good to briefly discuss this here (and perhaps elsewhere if drift correction has to be applied?) [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. All models have been drift corrected in final figure
9-1194	9	34	41			Section 9.4.2.3: suggest to refer to Chapter 3 where appropriate when discussing observed circulation features and their changes. Ideally the same observations as assessed in Ch3 would be referred to here. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Reference to Chap 3 added where appropriate. Observation coherent when possible as some published modelling studies use other reference observations.
9-1195	9	34	47			"In the AR4 ..." Please, be consistent with usage of AR4 models or CMIP3 models when referring to the same set of models. If both terms are indistinctly used, please explain it. [Government of Spain]	Taken into account. CMIP3 used instead of AR4.
9-1196	9	34	51			These values have recently been updated by McCarthy et al. (2012) who included values up to 2009 and	Taken into account. Values and reference updated

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						found a mean AMOC of 18.5 Sv: A reference to this study should be added: McCarthy, G., E. Frajka-Williams, W. E. Johns, M. O. Baringer, C. S. Meinen, H. L. Bryden, D. Rayner, A. Duchez, C. Roberts, and S. A. Cunningham (2012), Observed interannual variability of the Atlantic meridional overturning circulation at 26.5°N, <i>Geophys. Res. Lett.</i> , 39, L19609, doi:10.1029/2012GL052933. [RYM MSADEK, United States of America]	
9-1197	9	34	52			If one wants to be accurate, the direct observations of AMOC measured by the RAPID array are done at 26.5°N not 26N. [RYM MSADEK, United States of America]	Noted. Changed to 26.5; though most CMIP5 models are unlikely to be able to appreciate this difference.
9-1198	9	34	55	34	55	some recent work can be cited here on the impacts of aerosols on the simulation of ocean via atmospheric forcing "the strength of the weakening is correlated with the initial AMOC strength (Gregory et al., 2005). The" can be replaced with "the strength of the weakening is correlated with the initial AMOC strength (Gregory et al., 2005). Numerous other uncertainties and factors are associated with the historical AMOC simulation, findings by Delworth and Dixon (2006) and Collier et al. (2012) have stated that without the influence of anthropogenic aerosols, changes in ocean circulation would have already followed a path much more like one dominated by increasing long-lived greenhouse gases. The". References are Delworth, T.L., Dixon, K.W. (2006). Have anthropogenic aerosols delayed a greenhouse gas-induced weakening of the North Atlantic thermohaline circulation? <i>Geophys. Res. Lett.</i> , 33, L02606, doi:10.1029/2005GL024980 and Collier, M.A., Rotstayn, L.D., Kim, K.Y., Hirst, A.C., Jeffrey, S.J. (2012). Ocean circulation response to anthropogenic-aerosol and greenhouse-gas forcing in the CSIRO-Mk3.6 coupled climate model. <i>Aust. Met. Oceanographic Journal</i> , submitted. [Mark Collier, Australia]	Rejected. In the present section, the mean state of the simulated AMOC is discussed. The suggested references discuss the simulated weakening of the AMOC, and are therefore not added here.
9-1199	9	34	56	34	57	Just because the answer has not changed, does not necessarily mean there has been no improvement. Has the physical representation changed? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. So far, there is no indication for an improvement, but since there has not been a dedicated analysis, the speculation is removed.
9-1200	9	34	57	34	57	the right figure number is 12.35 [Annalisa Cherchi, Italy]	Taken into account. Corrected
9-1201	9	34	57			I suggest replacing "improvement" by "change" since we do not know much about AMOC strength except at few latitudes like 26N. It should also be added that a better representation of Nordic Seas overflows is key for a better simulation of the AMOC. It has been shown in least in one CMIP5 model (NCAR CCSM4) that including a parameterization of overflows leads to a better representation of the NADW and hence of the AMOC (Yeager and Danabasoglu 2012, Msadek et al. 2012) Yeager, S. G. and G. Danabasoglu, 2012: Sensitivity of Atlantic meridional overturning circulation variability to parameterized Nordic Sea overflows in CCSM4. <i>J. Clim.</i> , 25, 2077–2103, doi:10.1175/JCLI-D-11-00149.1. Msadek R., W. E. Johns, S. G. Yeager, G. Danabasoglu, T. Delworth, and T. Rosati, 2012: The Atlantic meridional heat transport at 26.5°N and its relationship with the MOC in the RAPID-array and GFDL and NCAR coupled models, <i>Journal of Climate</i> , accepted [RYM MSADEK, United States of America]	Noted. These references are added to the section on AMOC variability (9.5.3.3.1)
9-1202	9	34	57			"notable improvement"? Can you provide more detail? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Deleted. But added reference to Weaver et al., 2012. Weaver, Andrew J.; Sedláček, Jan; Eby, Michael; Alexander, Kaitlin; Crespin, Elisabeth; Fichet, Thierry; Philippon-Berthier, Gwenaëlle; Joos, Fortunat; Kawamiya, Michio; Matsumoto, Katsumi; Steinacher, Marco; Tachiiri, Kaoru; Tokos, Kathy; Yoshimori, Masakazu; Zickfeld, Kirsten Stability of the Atlantic meridional overturning circulation: A model intercomparison <i>Geophys. Res. Lett.</i> , Vol. 39, No. 20, L20709 http://dx.doi.org/10.1029/2012GL053763
9-1203	9	35	1	35	7	there are no results/analysis for CMIP5? at least it should be mentioned [Annalisa Cherchi, Italy]	Taken into account
9-1204	9	35	2	35	2	"low horizontal resolution if current": insert "most" before current: there is a huge range of model	Taken into account, text modified

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						resolutions in CMIP5. [Martin Juckes, United Kingdom]	
9-1205	9	35	14	35	15	It has also been suggested that ACC transport is determined by the integrated wind stress along the entire path of the ACC: see Allison, LC et al 2010. Where do winds drive the Antarctic Circumpolar Current? GEOPHYSICAL RESEARCH LETTERS, VOL. 37, L12605, doi:10.1029/2010GL043355 [James Christian, Canada]	Taken into account, paper cited
9-1206	9	35	15	35	15	Kuhlbrodt et al. (Ocean Modelling 2012) find that the strongest influence on the simulated ACC strength in CMIP3 is the Gent-McWilliams thickness diffusivity, not the windstress. [Jonathan Gregory, United Kingdom]	Taken into account, paper cited
9-1207	9	35	19	35	19	... Inter-model... [European Union]	Taken into account, text modified
9-1208	9	35	32	35	32	"compared" should be "compare" [Annalisa Cherchi, Italy]	Taken into account, text modified
9-1209	9	35	32	35	32	... to compare model... [European Union]	Taken into account, text modified
9-1210	9	35	33	35	35	It is not clear to us how a reader should reconcile here "relatively well" with "considerable spread"? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account, text clarified
9-1211	9	35	39	35	50	FGOALS in Figure 9.18 should be FGOALS-g or FGOALS-g1.0. If possible, please include more PMIP3 model results (e.g., FGOALS-g2 and FGOALS-s2) to this figure. Please refer to Zheng and Yu (2012) for the details about the PMIP3 results by FGOALS-g2 and FGOALS-s2, and please also cite this paper. By the way, at the end of line 46, "PMIP2/CMIP5" should be "PMIP2/CMIP3". Zheng, W. and Yu. Y., 2012: Paleoclimate simulations of the mid-Holocene and Last Glacial Maximum by FGOALS, Adv. Atmos. Sci., doi:10.1007/s00376-012-2177-6. [Bin Wang, China]	Taken into account. Model name and typo corrected. We do not have enough space to include the reference paper for all the modeling groups. In practice model reference are included in table 9.1.
9-1212	9	35	40			Figure 9.18: need to mention in the caption the isolines for density. Currently only T/S are mentioned [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account, text modified
9-1213	9	35	48	35	50	put the second part of the sentence into the main text (text starting with "and shows....."). [Farahnaz Khosrawi, Sweden]	Editorial. Taken into account
9-1214	9	36	2	36	12	We consider some rather vague language is used in this paragraph without being quantitative: reasonably well, slightly, slightly reduced, fairly close etc. Please try to use more quantitative language. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text modified to be sharper.
9-1215	9	36	2			The subsection is so short that it does not need to be an own subsection. One could start this paragraph with just the section title as headline (thus without section number). [Farahnaz Khosrawi, Sweden]	Taken into account. Text modified.
9-1216	9	36	2			I see now that there is reference made to the above Lee et al study via the figures 9.19, 9.20 but there is not reference given for this in the text, captions or bibliography. [duane waliser, United States of America]	Taken into account. Reference added
9-1217	9	36	3	36	12	Is it possible to say which observations are considered most reliable? (e.g., Quikscat seems to be highlighted). [Government of United States of America]	Taken into account. Text modified.
9-1218	9	36	4	36	4	it could be "The annually and longitudinally averaged zonal wind stress is reasonably...." [Annalisa Cherchi, Italy]	Taken into account. Text modified to be clearer.
9-1219	9	36	14			Same as previous comment. The subsection is so short that it does not need to be an own subsection. One could start this paragraph with just the section title as headline (thus without section number). [Farahnaz Khosrawi, Sweden]	Taken into account. Text modified.
9-1220	9	36	15	36	15	What time or spatial scales do you have to average over so that "steady state" is a good approximation? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account, text clarified
9-1221	9	36	21	37	9	Figure 9.21: suggest to move the explanatory text on page 37, lines 5-9 into the main text [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account, text modified
9-1222	9	36	23	36	23	"The models simulations" should be either "The model simulations" or "The models' simulations" [Sonya Legg, United States of America]	Editorial. Taken into account

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1223	9	36	30	36	31	This also highlights uncertainties even in the observations. [Government of United States of America]	Noted.
9-1224	9	36	34	36	39	Figure 9.19: add estimate of uncertainty observation-based estimate from GCP if available; suggest to move the explanatory text on lines 37-39 into the main text [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text modified.
9-1225	9	36	37	36	39	This sentence should be included in the main text rather than in the figure caption. [Farahnaz Khosrawi, Sweden]	Taken into account. Text modified.
9-1226	9	36	42	36	48	Figure 9.20: suggest to move the explanatory text on lines 46-48 into the main text [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text modified.
9-1227	9	36	46	36	48	Same here. The sentence should be included in the main text rather than in the figure caption. [Farahnaz Khosrawi, Sweden]	Taken into account. Text modified.
9-1228	9	36	50	37	9	Figure 9.21 uses an uncertain model name "LASG". Because LASG has two model versions for CMIP5: FGOALS-g2 and FGOALS-s2, please correctly use one of them to replace the name "LASG" in the figure. [Bin Wang, China]	Taken into account. Text modified.
9-1229	9	36	51	37	9	The text for the figure caption is definitely too long and parts of it would fit much better into the main text. [Farahnaz Khosrawi, Sweden]	Taken into account. Text modified.
9-1230	9	37	5	37	9	Can this note be moved from the caption to the main text? [Government of Germany]	Taken into account. Text modified.
9-1231	9	37	11	38	23	9.4.2.5 Section 3.2 of the Sheffield et al. paper titled "North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology" could be utilized to inform the discussion of simulated SSTs in the CMIP5 models. [Government of United States of America]	Noted. References on SST in CMIP5 added.
9-1232	9	37	14	37	15	"From CMIP1 through CMIP5 Randall et al., 2007)": references are too early to cover CMIP5! [Martin Jukes, United Kingdom]	Taken into account. Text modified.
9-1233	9	37	17	37	17	"and the equatorial cold tongue" vague. What about the cold tongue? Its surface temperature? Its zonal extent? [James Christian, Canada]	Taken into account. Text clarified.
9-1234	9	37	17			Suggest a good summary paper of the tropics as a climate system is Brown, J. N., A. Sen Gupta, et al. (2012). "Implications of CMIP3 model biases and uncertainties for climate projections in the western Tropical Pacific." Climatic Change. DOI 10.1007/s10584-012-0603-5 [Jaclyn Brown, Australia]	Taken into account. Reference added
9-1235	9	37	19	37	19	I would delete "deficient" here. [James Christian, Canada]	Taken into account. Text modified.
9-1236	9	37	20			Should also mention insufficient representation of subgrid-scale convective momentum transport in the atmosphere (Kim et al., Climate Dyn., 2008). [Andrew Wittenberg, United States of America]	Taken into account. Text modified.
9-1237	9	37	25	37	25	Could the equatorial cold bias be more accurately explained here? 'Excessive cold tongue' is unclear. The temperatures in the east Pacific, where the 'tongue' in SST appears, tend to be too high in models. The largest cold bias is in the west-central Pacific (170-200E) -do these relate to an advected tongue? [Government of Australia]	Taken into account. Text modified.
9-1238	9	37	25	37	25	Could the equatorial cold bias be more accurately explained here? 'Excessive cold tongue' is unclear. The temperatures in the east Pacific, where the 'tongue' in SST appears, actually seem to be too high in models. The largest cold bias is in the west-central Pacific (170-200E) -do these relate to an advected tongue? [Ian Watterson, Australia]	Taken into account. Text modified.
9-1239	9	37	32	37	43	This discussion is mainly on atmosphere (precipitation, winds, clouds) so it is not clear that it should be in Ocean section. Perhaps move some material to Atmosphere section and focus here on ocean fields. [Josephine Brown, Australia]	Taken into account. The tropical mean state result from strongly coupled ocean and atmosphere interactions and both sources of model performance need to be assessed together. The section could hence go in either atmosphere or ocean section. We decided to have it in the ocean section.
9-1240	9	37	32			Particularly good reference for potential reasons for the DOUBLE ITCZ -----Hirota, Nagio, Yukari N.	Taken into account. Reference added

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Takayabu, Masahiro Watanabe, Masahide Kimoto, 2011: Precipitation Reproducibility over Tropical Oceans and Its Relationship to the Double ITCZ Problem in CMIP3 and MIROC5 Climate Models. <i>J. Climate</i> , 24, 4859–4873. [duane waliser, United States of America]	
9-1241	9	37	40	37	40	... Insufficient... [European Union]	Editorial. Taken into account
9-1242	9	37	40	37	43	What is the cause for these improvements? [Government of United States of America]	Noted. The causes are multiple and outside the scope of the present section.
9-1243	9	37	40	37	43	What is the cause for these improvements? [Government of United States of America]	same as above
9-1244	9	37	42	37	42	Improvements are being made - where? In CMIP5 models? [Josephine Brown, Australia]	Taken into account. Text clarified.
9-1245	9	37	51	37	51	Remove double negative: "may not under-represent". [Josephine Brown, Australia]	Taken into account. Text modified.
9-1246	9	37	55	38	10	Besides the possible reasons mentioned here, there are some other potential factors may causing the biases in the tropical Atlantic, for example, clouds-radiation-SST interaction. The subtropical low cloud biases are extremely large for almost all current models, particularly in the southwestern Atlantic (Xie 2005; Huang et al. 2007; Wahl et al. 2009; Hu et al. 2011), which may also be a reason for the SST zonal gradient reversion along the equatorial Atlantic Ocean. Diagnostic and model sensitive experiments (Huang et al. 2007; Hu et al. 2011; Wahl et al. 2009) had suggested that cloud–radiation–SST interaction processes play an important role in model bias evolution as well as in anomalous climate events in the southeastern Atlantic Ocean. Xie S-P, 2005: Chapter 4: The shape of continents, air–sea interaction, and the rising branch of the Hadley circulation. In: Diaz HF, Bradley RS (eds) <i>The hadley circulation: present, past and future</i> . Kluwer Academic Publishers, The Netherlands, pp 173–202. Wahl, S., M. Latif, W. Park, and N. Keenlyside, 2009: On the Tropical Atlantic SST warm bias in the Kiel Climate Model. <i>Climate Dynamics</i> , DOI:10.1007/s00382-009-0690-9. Huang, B., Z.-Z. Hu, and B. Jha, 2007: Evolution of model systematic errors in the tropical Atlantic basin from the NCEP coupled hindcasts. <i>Clim. Dyn.</i> , 28 (7/8), 661-682, DOI: 10.1007/s00382-006-0223-8. [Zeng-Zhen Hu, United States of America]	Taken into account. Text modified to take comment into account
9-1247	9	37	57	37	57	Have these issues been solved in the models used in this neww assessment report? If not, are these models still used? If the models are being used, explain why, when there is such big biases. [Juan Blanco, Spain]	Taken into account. Text modified to take comment into account
9-1248	9	37		38		Sections 9.4.5.2.2-3: The Tropical Atlantic and Indian Oceans are treated very minimally. [Government of India]	Taken into account. Space is an issue and this assesement is based on published material.
9-1249	9	38	4	38	4	"A recent study", which study? [Annalisa Cherchi, Italy]	Taken into account. Text clarified.
9-1250	9	38	4	38	7	Doi et al. (2012) showed that the tropical Atlantic biases are significantly improved when resolution is increased (comparing GFDL CM2.1 to GFDL CM2.5) mainly because of a significant reduction of some biases in the seasonal meridional migration of the ITCZ. A reference to this study could be added here: Doi, Takeshi, Gabriel A. Vecchi, Anthony J. Rosati, Thomas L. Delworth, 2012: Biases in the Atlantic ITCZ in Seasonal–Interannual Variations for a Coarse- and a High-Resolution Coupled Climate Model. <i>J. Climate</i> , 25, 5494–5511. doi: http://dx.doi.org/10.1175/JCLI-D-11-00360.1 [RYM MSADEK, United States of America]	Taken into account. Text modified to take comment into account
9-1251	9	38	4			Chang et al. (<i>J. Climate</i> , 2007) and Richter et al. (<i>Climate Dyn.</i> , 2012) showed that a large fraction of the equatorial SST biases are attributable to weaker-than-observed equatorial easterlies during boreal spring, which are in turn due to deficient rainfall over equatorial South America relative to equatorial Africa -- biases that exist even when the atmospheric component is driven by observed SSTs, and which are sensitive to both land surface formulation and atmospheric convective closure. [Andrew Wittenberg, United States of America]	Taken into account. Text modified to take comment into account
9-1252	9	38	12	38	23	chap 14 contains an analysis of the mean state and variability in the Indian Ocean,considering also CMIP5 outputs, and it should be referenced here [Annalisa Cherchi, Italy]	Taken into account. Text modified.
9-1253	9	38	12			In Subsection 9.4.2.5.3 nothing is said about CMIP5 results. [Government of Spain]	Taken into account. Text updated

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1254	9	38	27	38	27	"There is" instead of "It's likely" and remove brackets around "robust evidence and medium agreement" so that the sentence reads: "There is robust evidence and medium agreement that the ocean component.....". [Farahnaz Khosrawi, Sweden]	Editorial. Text modified
9-1255	9	38	27	38	28	Unclear sentence - not clear what the "likely" relates to, because it is not said whether the model components do this simulation well, very well, poor etc. Please correct. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text modified.
9-1256	9	38	27	38	29	I believe the model fidelity in ocean CO2 uptake is not covered in the previous sections and should be removed from this sentence. [Koichi Sakaguchi, the United States]	Taken into account. Text modified.
9-1257	9	38	27			give reference to the definition of "likely" [Barbara Früh, Germany]	Taken into account. Text modified.
9-1258	9	38	28			"at play" colloquial? [Barbara Früh, Germany]	Taken into account. Text modified.
9-1259	9	38	28			"simulate reasonably well"? We posit that "Simulate" alone is too vague. [Government of United States of America]	Taken into account. Text modified.
9-1260	9	38	29	38	29	"There is little evidence ...": this statement can't be justified on the basis of the limited analysis of CMIP5 which has been completed so far – "Analysis completed so far" [Martin Jukes, United Kingdom]	Taken into account. Text modified.
9-1261	9	38	32	38	32	"intermodal" -> "intermodel" [Bart Van den Hurk, Netherlands]	Taken into account. Text modified.
9-1262	9	38	33	38	34	The sentence beginning "New since the AR4.." needs to be rewritten. It is a clumsy sentence. [Government of Australia]	Taken into account. Text modified.
9-1263	9	38	36	38	46	This summary is very confusing. [Josephine Brown, Australia]	Taken into account. Text modified.
9-1264	9	38	36	38	46	this concluding summary for section 9.4.2 Ocean uses the term "skill" quite a bit while all the previous subsection in 9.4.2 do never refer to "skill" -- suggest to consider consistent use of terminology within the subsection. In addition, it is not entirely clear to us how formulations such as "well simulated" or "medium quality" can be compared to the "mixed/high skill" formulations. Consistency in the evaluative language used here is important to avoid confusion. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text modified.
9-1265	9	38	36			I would suggest replacing "SST" by "global mean SST" to avoid misleading interpretations, given that there are significant SST biases if one looks at oceanic basins individually, even on quite large spatial scales. [RYM MSADEK, United States of America]	Taken into account. Text modified.
9-1266	9	38	37	38	37	Write "there is" [Juan Blanco, Spain]	Editorial. Text modified
9-1267	9	38	37	38	37	The use of 'and still significant' at the end of this sentence doesn't flow well. Suggest rephrasing this sentence. [Government of Australia]	Taken into account. Text modified.
9-1268	9	38	37	38	37	... There is robust... [European Union]	Editorial. Text modified
9-1269	9	38	37	38	37	insert "is" after "there" [Bart Van den Hurk, Netherlands]	Editorial. Text modified
9-1270	9	38	38	38	39	limited evidence and medium agreement that western boundary currents are simulated with mixed skill' reads strangely. Limited evidence of mixed skill? [Government of Australia]	Taken into account. Text modified.
9-1271	9	38	38			" There is robust evidence and high agreement that the AMOC is simulated with mixed skill with limited improvements from CMIP3 to CMIP5" I find this sentence ambiguous. Observations of the AMOC are very limited so I do not agree that there is robust evidence. Table 9.3 p.57 states that there is limited evidence for evaluation of the AMOC. Even if that table refers to the variability rather than the mean evaluated here p.38, I do not think the evidence is more robust for the mean and I would suggest being consistent with table 9.3. Further what is meant by "mixed skill" is not clear to me. Does it mean limited skill on average, or good for some models and not as good for others? Also how can there be a high agreement given the large spread among models? This sentence should be clarified. [RYM MSADEK, United States of America]	Taken into account. Text modified.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1272	9	38	41	38	41	Pls remove "zonal" after agreement that [HASIBUR RAHAMAN, India]	Taken into account. Text modified.
9-1273	9	38	45	38	45	"the CMIP3 models" should be CMIP5? [James Christian, Canada]	Taken into account. Corrected
9-1274	9	38	45			Do you mean CMIP5 or CMIP3? Please clarify. [Government of United States of America]	Taken into account. Corrected
9-1275	9	38	48			the same from the comment No 27 [CELSO COPSTEIN WALDEMAR, BRAZIL]	Noted, however this comment would be more helpful if it is more specific.
9-1276	9	38	48			the same from the comment No 30 [CELSO COPSTEIN WALDEMAR, BRAZIL]	Taken into account wherver found relevant, however this comment would be more helpful if it is more specific.
9-1277	9	38	52	38	55	The accuracy of algorithms used to retrieve ice concentration and extent from satellite measurements is currently not addressed in Chapter 4! [Thierry Fichefet, Belgium]	Noted.
9-1278	9	38	53	38	55	This statement is in stark contrast with the discussion of such observations in chapter 4 where there is no discussion of such errors. The question of sea ice extent errors should be settled in Chapter 4 and deserves a clearly stated number. I don't actually believe this statement here is correct. While there are differences in the absolute extent, differences in variability are likely minor. A better target in this context might be ice type (first/multi), thickness either derived from age or measured from freeboard, or volume. Chapter 4 authors should consult with chapter 9 authors and clarify this [Axel Schweiger, United States of America]	Noted
9-1279	9	39	2	39	8	Terms such as "modest agreement", "reasonable agreement", "fairly well", "reasonably well" etc, are used throughout Chapter 9 when comparing models with observations. It is not clear how these terms rate sequentially - is 'modest agreement' a better result than 'reasonable agreement'? It would be good if more quantitative terminology could be used where ever possible. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Language is calibrated.
9-1280	9	39	3	39	3	modest improvement----- intro states "clear improvement"... seems a bit inconsistent [Axel Schweiger, United States of America]	Accepted. Text modified
9-1281	9	39	4	39	4	The statement that there have been no substantial improvements in sea ice parameterisations seems contradicted in line 34, where improvements in sea ice albedo have improved modelled trends [Government of Australia]	Taken into account. Text modified.
9-1282	9	39	4	39	4	"no substantial increase in sophistication of sea-ice": there have been significant improvements in terms of the dynamical equations used and the resolution of vertical structure in the ice. [Martin Jukes, United Kingdom]	Taken into account. Text modified.
9-1283	9	39	12	39	12	skip the text in parantheses, this is already given in the main text. [Farahnaz Khosrawi, Sweden]	Rejected. In the main text, an explanation is given of what the median ice edge is.
9-1284	9	39	12			Figure 9.22: add uncertainty for the observations if available; suggest to visually clarify what the two panels represent by adding more information (e.g., NH, SH, annotate axis [Month], etc.) [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. The figure is modified.
9-1285	9	39	13			In Figure 9.22 legend, number of CMIP3 and CMIP5 models is specified, while in other figures not. [Government of Spain]	Accepted. Figure captions modified
9-1286	9	39	14	39	15	Here it is given for what the abbreviations of observations are standing for. However, in other figure captions this isn't done. This should be consequently done either way in all figure captions. [Farahnaz Khosrawi, Sweden]	Rejected for brevity. References are given which is enough to identify the data sets.
9-1287	9	39	19			Figure 9.23: suggest to better explain how the "normalized number of models" is to be compared to the total number of models. What's the unit used in panel B? %? Does a value of -0.9 mean that the number of models showing 15% concentration at a grid cell is reduced by 90% in CMIP5? If so, please add that information to the caption. If not, please explain what it means. [Thomas Stocker/ WGI TSU, Switzerland]	Noted. Panel removed
9-1288	9	39	28	39	30	I would suggest to stress that the CMIP3 models AS A GROUP underestiamtes the observed September trend subject to impacts of natural variability. A number of ensembels well capture the reality. More information can be found in Zhang (2010). [Xiangdong Zhang, United States of America]	Accepted. Text modified.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1289	9	39	28	39	30	Zhang, X., 2010: Sensitivity of Arctic summer sea ice coverage to global warming forcing: Toward reducing uncertainty in Arctic climate change projections. Tellus A, 62, 220-227, doi:10.1111/j.1600-0870.2010.00441.x. [Xiangdong Zhang, United States of America]	Taken into account. The reference is added.
9-1290	9	39	28	39	34	The underestimation of the observed trend from CMIP3 models is even more spectacular for sea ice thinning (an underestimation by a factor of 4, over the last 3 decades), see Rampal et al., JGR-C, 116, C00D07, 2011. Owing to the importance of this variable, this should be noted. [Jérôme Weiss, France]	Taken into account. The reference is added.
9-1291	9	39	28	39	43	Day et al. 2012 suggest that some of the discrepancy between the CMIP multi-model mean the observed trend could be due to the effect of the AMO on the observed record. This is thought to account for as much as 30% of the decrease since the 1979s. Kay et al. 2012 also discuss the role of internal variability and think this could explain about half of the observed trend. Day, J. J., Hargreaves, J. C., Annan, J. D., Abe-Ouchi, A. (2012) Sources of multi-decadal variability in Arctic sea ice extent. Environmental Research Letters, 7 (3). 034011. ISSN 1748-9326 doi: 10.1088/1748-9326/7/3/034011 [Jonathan Day, United Kingdom of Great Britain & Northern Ireland]	Taken into account. The reference is added.
9-1292	9	39	30	39	30	observational uncertainties? How? We know extent, concentration, and even volume (Schweiger et al. 2011) well enough over the past decades. The issue is that we don't have long enough observations to really assess multi-decadal variability. May need some rephrasing. [Axel Schweiger, United States of America]	Taken into account. Text modified. Observational uncertainties can play a role e.g. for the pre-sat period. But see also Kattsov et al. (2010)
9-1293	9	39	32	39	34	It is true that the CMIP5 models as a group make a better job than the CMIP3 ones in reproducing the observed trend in September Arctic sea ice extent. However, it is worth mentioning that the CMIP5 multi-model mean significantly underestimates the magnitude of this trend from 2005 to 2012. [Thierry Fichefet, Belgium]	Rejected. The multi-model mean is not supposed to reproduce observed trend.
9-1294	9	39	36	39	36	Holland et al should not be in parenthesis [Government of United States of America]	Editorial
9-1295	9	39	36	39	36	Pls change "(Holland et al., 2010)" to "Holland et al., (2010)" [HASIBUR RAHAMAN, India]	Editorial
9-1296	9	39	37	39	39	It may not necessarily true that thin winter ice cover causes rapider melting. More information could be found in Thierry Fichefet's new study. [Xiangdong Zhang, United States of America]	Noted. However the study is not specified
9-1297	9	39	39	39	39	All these 2012 references that aren't published make it hard to assess how much support they provide to the statement [Axel Schweiger, United States of America]	Noted
9-1298	9	39	39	39	43	This is an interesting point about internal climate variability simulating the observed trend. It would be good to have a multi-model perspective, not just one model. Also are there ways of assessing the internal variability in the model which has very large values of it according to the text? [Government of United States of America]	Taken into account. A new figure is added to address this.
9-1299	9	39	45	39	46	Explicitly mention that the Antarctic trends mentioned are for autumn/winter (not all seasons) [Jeff Ridley, United Kingdom]	Taken into account. Text modified.
9-1300	9	39	45	39	49	The possible causes of the inability of models to simulate the current increase in Antarctic sea ice extent need to be addressed here. This is essential for assessing the sea ice projections discussed in Chapter 11 and 12. [Thierry Fichefet, Belgium]	Noted, but this relates to attribution.
9-1301	9	39	45	39	49	Southern Hemisphere sea ice simulated by the CMIP3 and CMIP5 models deserves more attention. The authors should at least hypothesize about the causes for the mismatch between models and observations. Misrepresentation/absence of ozone in these models, strong unforced variability, initialization, ... should be discussed as possible reasons for this mismatch. [François Massonnet, Belgium]	Noted, but this relates to attribution.
9-1302	9	39	46	9	48	A key caveat for me is that the models systematically overestimate the variability of the system compared to satellite data over the last 30 years, at least in winter (see Zunz et al. 2012). Using the models to assess the potential role of internal variability in the recent trend of sea ice extent in the Southern Ocean is thus associated with many uncertainties that should be mentioned at this stage. [Hugues Goosse, Belgium]	Taken into account. Text modified.
9-1303	9	39	46	39	46	This should read 'observed weak increase' [Government of Australia]	Taken into account. Text modified.
9-1304	9	39	46	39	46	weak observed increase' should be written as "small but significantly significant increase" to be consistent	Taken into account. Text modified.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						with Chapter 4. [Thomas Stocker/ WGI TSU, Switzerland]	
9-1305	9	39	46	39	47	The spread is not so much in trends as absolute ice area. Include a comment that the spread in winter ice area is worse than for AR4 with a number of significant outliers (some with next to none winter ice extent. [Jeff Ridley, United Kingdom]	Rejected because of space limits
9-1306	9	39	50			Chapters 4, 10, and 11 of the draft reference the paper by Rampal et al (2011) showing discrepancies between observational estimates and IPCC model simulations of Arctic sea ice thickness decline and speed increase. This could also be mentioned here since the comparison of simulated extent decline to observations is being discussed. Reference: Rampal, P., J. Weiss, C. Dubois, and J.-M. Campin (2011), IPCC climate models do not capture Arctic sea ice drift acceleration: Consequences in terms of projected sea ice thinning and decline, J. Geophys. Res., 116, C00D07, doi:10.1029/2011JC007110 [Government of United States of America]	Rejected. Rampal et al's paper is about CMIP3 models' performance in the Arctic. Not very relevant for the discussion of CMIP5 performance in the Southern Ocean. This paper is referred to earlier in this section.
9-1307	9	39	52			Figure 9.24: please add uncertainty for the NSIDC observations if available [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Figure modified.
9-1308	9	39				9.4.3 Section 4.6 of the Sheffield et al. paper ("North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology") could be integrated into the discussion in this section of Chapter 9, although perhaps merely as a reference as it mostly agrees with the existing discussion. [Government of United States of America]	Noted. The background literature for this paper is already cited in this section.
9-1309	9	39				Figure 9.22, 9.24: Assuming that the gray shading indicates overlap of results from CMIP 3 and CMIP 5, include a label with this color and information, and state so in the figure legends. [Arne Melsom, Norway]	Taken into account. Figures modified.
9-1310	9	40	5	40	5	Pls change "(Koldunov et al., 2010)" to "Koldunov et al., (2010)" [HASIBUR RAHAMAN, India]	Editorial
9-1311	9	40	6	40	6	Pls change "(Melsom et al., 2009)" to "Melsom et al., (2009)" [HASIBUR RAHAMAN, India]	Editorial
9-1312	9	40	9			Here it says that biases in GCM-simulated sea ice may arise due to issues with simulated cloudiness. It bears noting that this possibility was proposed and extensively discussed in Eisenman et al. (2007, doi:10.1029/2007GL029914). [Ian Eisenman, United States of America]	Noted. Such possibility has also been mentioned in earlier assessments - AR4, and ACIA, 2005.
9-1313	9	40	12	40	14	are the "some recent models" different than the "some CMIP5 models"? Please clarify. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text modified.
9-1314	9	40	14	40	14	The Holland et al., 2008 reference points to a paper by P. R. Holland and coauthors: "The response of ice shelf basal melting to variations in ocean temperature". I think there must be an error in the reference since the paragraph is concerned with Arctic sea ice. [François Massonnet, Belgium]	Taken into account. Text modified.
9-1315	9	40	14			Please clarify: how large were the internally generated trends compared to the observed post 1979 trend? [Government of United States of America]	Noted. This text is removed.
9-1316	9	40	16	16	22	That needs to be updated in view of the 2012 observations. [Government of France]	Accepted.
9-1317	9	40	18	40	19	A better tuning might also be at the origin of the performance improvement! [Thierry Fichefet, Belgium]	Noted.
9-1318	9	40	20	40	22	"Most slight increase observed": please say whether the observations and models are consistent (i.e. the difference is within estimated natural variability) or not. [Martin Jukes, United Kingdom]	Taken into account. Text modified.
9-1319	9	40	21	40	22	slight observed increase' should be written as "small but significantly significant increase" to be consistent with Chapter 4. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text modified.
9-1320	9	40	24	42	38	Chapter 9: Section 9.4.4 (Simulation of Recent and Longer Term Records in Global Models. Land Surface, Fluxes and Hydrology). This section does not contain any introduction neither a summary. It would be desirable to have both of them as in previous section [Government of Spain]	Accepted and taken into account - intro and summary added.
9-1321	9	40	28	40	45	In a further comment on Christy 2012 (J Hydr Met) I mentioned the lack of snow declines in the observations that were predicted in models to have occurred by now in the Sierra Nevada of California. It turns out there is no trend at all in the observations through 2011, while models predicted a 6 to 10 percent per decade decline	Noted. However, there is not space here to discuss specific regional evaluations of snow-cover.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						(way off from reality) for the same period. [John Christy, United States of America]	
9-1322	9	40	28	40	45	Is there evidence of a dependence of the seasonal cycle of snow cover, and snow amount, on the vegetation distribution, e.g. do models which include a DGVM show any bias related to errors in the vegetation type? That would be consistent with work by e.g. McCarthy et al (2011; Earth Sys. Dyn. 3, 87-96), Martin and Levine (2012; Earth Sys. Dyn. 3, 245-261) [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Noted. In general we expect the introduction of a new interactive component (such as dynamic vegetation or cloud cover), to degrade the simulation of the current climate compared to models where that component is prescribed from observations. However, we have no specific evidence of DGVMs degrading snowcover simulation in the CMIP5 models.
9-1323	9	40	28	40	45	Terms such as "modest agreement", "reasonable agreement", "fairly well", "reasonably well" etc, are used throughout Chapter 9 when comparing models with observations. It is not clear how these terms rate sequentially - is 'modest agreement' a better result than 'reasonable agreement'? It would be good if more quantitative terminology could be used where ever possible. [Thomas Stocker/ WGI TSU, Switzerland]	Agreed and taken into account - replaced by more quantitative calibrated language.
9-1324	9	40	40	40	40	"mean-model": replace with "multi-model mean". [Martin Juckes, United Kingdom]	Accepted.
9-1325	9	40	41	40	41	how big is "fairly large"? Please use more quantitative language if possible. [Thomas Stocker/ WGI TSU, Switzerland]	Agreed and taken into account - replaced by more quantitative calibrated language.
9-1326	9	40	41	40	43	Are these "bias corrected" for air temperature before working out where the snow is, or not? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	No they are not bias-corrected.
9-1327	9	40	43	40	45	Note Derksen and Brown (2012), doi:10.1029/2012GL053387 [Richard Essery, United Kingdom of Great Britain & Northern Ireland]	Agreed and taken into account - reference added.
9-1328	9	40	43			Does this figure appear anywhere in the AR5 draft? It would be useful to reference it if so. [Government of United States of America]	see response to 3-1327.
9-1329	9	40	44			But what caused the underestimate of boreal land surface warming? Which came first? Any role for natural variability for example? [Government of United States of America]	Agreed and taken into account - we replace "The main reason.." by "..., which is associated with.."
9-1330	9	40	45	40	45	an underestimation of boreal warming can also be a result of improper snow modelling, not only a cause. [Bart Van den Hurk, Netherlands]	Agreed and taken into account - see response to 9-1329.
9-1331	9	40	48	40	54	While visually the simulated and observed permafrost extent agrees reasonably well for Eurasia, there seems to be a systematic poleward displacement over North America. Is it possible to comment on the reasons for this? [Government of Germany]	Noted. The most likely reason for a poleward displacement of permafrost is a cold bias, but we do not have the analysis available to confirm this.
9-1332	9	40	48	40	54	The observed and simulated metrics are related but how can a direct comparison be justified? Can they be directly compared without qualification (Apples vs oranges)? [Government of United States of America]	Agreed and taken into account - we make it clearer that the comparison here between model and observations is qualitative rather than quantitative. We now include the text: "Some CMIP5 models now represent permafrost and frozen soil processes (Koven et al., 2012), but this is not generally the case. Therefore it is difficult to make a direct quantitative evaluation of most CMIP5 models against permafrost observations " .
9-1333	9	40	56	41	6	There is no mention here of the actual permafrost distributions that the CMIP5 models themselves predict, as there ought to, given the chapter focus. The climate models can be evaluated directly by examining the locations where model soil levels stay at or below freezing for extended periods. This was the focus of Koven, et al. (2012), which describes a number of specific metrics to evaluate the quality of the models' physics as relevant to the permafrost problem. The result is a high disagreement in the permafrost distributions predicted in the current climate, resulting from widely diverging treatment of snow, soil hydrology, and freeze/thaw processes, but a more consistent loss of permafrost as a result of warming. [Government of United States of America]	Agreed and taken into account - we now also mention results from Koven et al.(2012).
9-1334	9	40	56	41	12	It would be worth stating here that if simulated deep soil temperatures are directly used to diagnose presence	Agreed and taken into account - see response to 9-

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						or absence of permafrost in the CMIP5 models, the diagnosed permafrost extents differ widely among the models (Koven et al., Analysis of permafrost thermal dynamics and response to climate change in the CMIP5 Earth System Models, J. Climate, in press (currently cited in Ch. 12 as "Koven and Riley, submitted")). [Gerhard Krinner, France]	1334.
9-1335	9	40				9.4.4.2 The paper "North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology" by Sheffield et al. contains a discussion of terrestrial water storage and runoff evaluated using the CMIP5 models, which could be integrated into the discussions in section 9.4.4.2. [Government of United States of America]	Noted, but this paper doesn't seem to have been published and only refers to North America.
9-1336	9	41	3	41	3	Pls change "(Slater and Lawrence, 2012)" to "Slater and Lawrence, (2012)" [HASIBUR RAHAMAN, India]	Editorial
9-1337	9	41	6	41	6	Pls remove extra "." [HASIBUR RAHAMAN, India]	Editorial
9-1338	9	41	11	41	12	There is no published evidence that it is necessary to include the microbial heat release, so we would recommend deleting it from the sentence. If our understanding of the literature is incorrect, please add appropriate reference (s). The much larger uncertainty is due to the fundamental physical differences between the models: the snow insulation effects and treatment of soil water and its latent heat of freezing, in addition to what is already listed in the sentence. [Government of United States of America]	Rejected. There are a number of papers that report the significant effect of microbial heat release on the simulation of permafrost and we already cit one of these (Koven et al., 2011).
9-1339	9	41	18	41	48	Given the importance of soil moisture to climate and impacts, there is a notable lack of information about the basic performance of the CMIP5 models, which include well developed land surface components. There are no results shown, in contrast to other variables. Note also that Fig. 12.23 shows projected changes in (10cm layer) soil moisture. [Government of Australia]	Accepted, but we need to rely on published analyses and there have been regrettably few analyses of the performance of the land-surface components within the CMIP5 models. We do however now discuss the results from Sheffield et al. (see response to 3-1335).
9-1340	9	41	18			Mueller et al. (2011) evaluated AR4 evapotranspiration model output against six remote sensing-based evapotranspiration products. [Mueller, B., S. I. Seneviratne, C. Jiménez, T. Corti, M. Hirschi, G. Balsamo, P. Ciais, P. A. Dirmeyer, J. B. Fisher, Z. Guo, M. Jung, F. Maignan, M. F. McCabe, R. H. Reichle, M. Reichstein, M. Rodell, J. Sheffield, A. J. Teuling, K. Wang, E. F. Wood, and Y. Zhang (2011), Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, Geophysical Research Letters, 38(L06402), doi:10.1029/2010GL046230.] [JOSHUA FISHER, United States of America]	Agreed and taken into account - this study is now referenced.
9-1341	9	41	18			Schwalm et al. (in revision) assessed AR5 evapotranspiration output against six remote sensing-based evapotranspiration products. Schwalm et al. determined that the choices inherent in the process of benchmark evaluation led to any number of climate model output being ranked differently; thus, there was no clear consensus on model skill until these choices are solidified. [Schwalm, C. R., D. N. Huntzinger, A. M. Michalak, J. B. Fisher, J. S. Kimball, B. Mueller, K. Zhang, and Y. Zhang (in revision), Sensitivity of inferred climate model skill to choice of benchmarking datasets and evaluation decisions, Geophysical Research Letters.] [JOSHUA FISHER, United States of America]	Noted.
9-1342	9	41	23			Rooting depth is specifically called out in this sentence. Please include a reference. [Government of United States of America]	Rejected. It is widely-understood (and intuitive) that plant root-depth affects the evapotranspiration flux. We don't believe an additional reference is necessary here.
9-1343	9	41	26	41	29	"but at the cost of increasing the number of poorly known internal model parameters". This is an important point not discussed elsewhere. What is the cost of increasing the number of unknown parameters? This is one of the issues that sooner or later has to be discussed. Maybe the excellent Box 9.1 could add a few lines on this, as well as on the issue of "right for the wrong reasons". [Ramon de Elia, Canada]	Noted.
9-1344	9	41	27	41	29	This is another general point that may be useful to put in the FAQ 9.1 - how in general, during model development, do we distinguish between "useful" extra parameters and those which simply increase the complexity of the model without net benefit (as for example in the case highlighted here, where parameters are introduced but are very poorly constrained and therefore simply increase uncertainty). [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted.
9-1345	9	41	31			"such as high CO2" shouldn't be added to that sentence. [Government of United States of America]	Noted, but these models ARE developed with a view

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							to simulating the response of the climate to increased CO2. So we have replaced "high" with "increased".
9-1346	9	41	35	41	48	The paper titled "Evolving land-atmosphere interactions over North America from CMIP5," prepared by Dirmeyer et al. could inform the discussion in this particular part of section 9.4.4.2, as it examines many of the same fields described in this section for the CMIP5 models, and so is more contemporary. [Government of United States of America]	Noted.
9-1347	9	41	39	41	39	Clarification: does "models that simulated global dimming" refer to models with simulated aerosols, models with prescribed aerosols, or a prescribed decrease in solar forcing? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted. This refers to models that simulated a reduction in surface solar radiation. The text has been modified to clarify this.
9-1348	9	41	39	41	41	"Key role" - does this mean first order importance? If land-surface models are important but are poor in quality, then what implication does this have for regional projections? On what spatial scale does the effect become first order? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted, but these questions are beyond the scope of this chapter.
9-1349	9	41	43	41	48	These few lines attempt to summarise the hugely complex issue of soil moisture - rainfall feedbacks, and the result is not clear. Firstly, the distinction should be made between soil moisture/rainfall coupling and soil moisture/rainfall feedbacks. For example, correlations between precipitation and soil moisture, referred to in lines 46-48, may demonstrate coupling but do not necessarily demonstrate feedback. The detailed discussions in Seneviratne et al. (2010), as referenced, highlight the discrepancies between models and, more importantly, between observational studies on soil moisture - rainfall feedbacks. More information about the complexity, uncertainty and temporal and spatial variability of this process, such as described in Seneviratne et al (2010), should be included. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Accepted and taken into account - these lines have been rewritten but space constraints prevent us from going into great detail on soil moisture/rainfall feedbacks.
9-1350	9	41	46			I recommend additional reference to Taylor et al. (2012); doi:10.1038/nature11377 [Alexander Loew, Germany]	Accepted and taken into account - this study is now referenced in the revised text on rainfall/soil moisture feedbacks.
9-1351	9	41	50	42	20	DGVMs have already been described in this chapter, refer to the earlier section [Government of United States of America]	Accepted - earlier section (9.1.3.2.4) referenced.
9-1352	9	41	50	42	20	DGVMs have already been described in this chapter, refer to the earlier section [Government of United States of America]	Accepted - earlier section (9.1.3.2.4) referenced.
9-1353	9	41	50			Sec 9.4.4.3 on N-cycling – you could also refer back to results in chapter 6, section 6.4.6.2 on N-cycle results from ESMs [Chris Jones, United Kingdom of Great Britain & Northern Ireland]	Accepted - earlier section 6.4.6.2 referenced.
9-1354	9	41	50			Section 9.4.4.3: This section implicitly assumes that the vegetation distribution simulated by CMIP5 models that include DGVMs in the present-day compares well with observations. For example, Martin and Levine (2012; Earth Sys. Dyn. 3, 245-261) show that systematic biases in rainfall are associated with increased present-day bare soil fractions in tropical regions in HadGEM2-ES, and that increased fractions of shrubs rather than trees over northeastern Eurasia are associated with increased snow cover in that region. The ability of the CMIP5 models with DGVMs to represent the present-day vegetation distribution should be evaluated here, and the possible implications of any errors for the projections, via feedbacks e.g. on the carbon cycle, should be discussed. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Noted. We do not assume that vegetation simulations are realistic and there is not space to evaluate the vegetation distributions produced by the few models that include DGVMs. However, we have added a caveat here "... with likely consequences for climate model biases and regional climate projections (Martin and Levine, 2012)".
9-1355	9	42	7	42	8	"associated climate change will tend to reduce it": this is counter-intuitive: temperature and precipitation increase will generally lead to better growing conditions and thus higher CO2 uptake [Bart Van den Hurk, Netherlands]	Rejected. This IS what DGVMs show (mainly due to increased soil and vegetation respiration under warming).
9-1356	9	42	12	42	12	What is about the impact of atmospheric N deposition at regional and global scales on C sequestration by natural ecosystems such as tropical forests? Considered or ignored? (see e.g. Xia and Wan, 2008 Global response patterns of terrestrial plant species to nitrogen addition, New Phytol 179, 428-439) [European Union]	Ignored in this generation of models, most of which ignore N-limitations on plant uptake.
9-1357	9	42	15	42	20	I suggest to replace the term 'DGVMs' with 'land surface models' in order to be more precise, because in the two referenced studies (Thornton et al., 2007; Zaehle et al., 2010a) the models were ran without dynamic vegetation (i.e., prognostic change of natural vegetation cover type). According to these two articles (and the	Accepted and taken into account.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						supplementary materials) the surface land cover types were kept constant, although the biomass and plant structure (e.g., canopy height LAI, and number of individuals for Zaehle et al.) were simulated prognostically. And the CMIP5 models with nitrogen cycle do not necessarily include DGVMs (e.g., NCAR CCSM4 and CESM1). So I recommend to revise the two sentences in this paragraph as: "Most of the land surface models and DGVMs used in the CMIP5 models continue to neglect nutrient-limitations on plant growth, even though these may significantly moderate the response of photosynthesis to CO2 (Wang and Houlton,2009). Recent extensions of two land surface models to include nitrogen limitations to CO2-fertilization improve the fit of these models to "Free-Air CO2 Enrichment Experiments", and suggest that models without these limitations will most likely overestimate the land carbon sink in the nitrogen-limited mid and high latitudes (Thornton et al., 2007; Zaehle et al., 2010a)." [Koichi Sakaguchi, the United States]	
9-1358	9	42	19	42	20	Besides N limitations also K and P limitations are likely to limit C sequestration by vegetation. This also needs to be covered by ESM. See e.g. Wright et al 2011 Potassium, phosphorus, or nitrogen limit root allocation, tree growth, or litter production in a lowland tropical forest, ECOLOGY 92, 1616-1625 [European Union]	Noted, but we already cover this by describing the neglect of "nutrient-limitations".
9-1359	9	42	22		33	The latest ESMs used in CMIP5 attempt to model the CO2 emissions implied by prescribed land-use change and many also simulate the associated changes in the biophysical properties of the land-surface. Do these models also consider the associated N emission and release to inland waters? [European Union]	No this generation of models do not consider these N emissions.
9-1360	9	42	22			Section 9.4.4.4: My comment (14) on section 9.4.4.3 also applies to the discussion of the effects of land-use changes: results may depend on whether these are imposed on observed or simulated vegetation distributions. This should be discussed here. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Noted, but the study referred to is only with prescribed vegetation.
9-1361	9	42	32	43	18	Are there any results for the southern hemisphere to report? [Government of Australia]	Most of the studies referred to are global, so unless otherwise stated, the statements made apply to the SH as well as the NH.
9-1362	9	42	35	42	35	"the modelling impacts" → "the modelling of impacts ..." [Martin Jukes, United Kingdom]	Agreed.
9-1363	9	42	40	45	32	Chapter 9: Section 9.4.5 (Simulation of Recent and Longer Term Records in Global Models. Carbon Cycle). This section does not contain any introduction neither summary. It would be desirable to have both of them as in previous sections (Atmosphere, Ocean) [Government of Spain]	Accepted and taken into account - intro paragraph now added to explain the remit of the section and its relevance to later chapter.
9-1364	9	42	40			Sec 9.4.5: it seems intuitively important to also evaluate carbon STORES as well as fluxes. Anav et al (J Clim submitted) is being revised to add these, and shows that there is a very wide spread in both veg and soil carbon stores across models. Full global observations may be lacking, but we must have some information. Todd-Brown et al., Biogeoscience Discussions (http://www.biogeosciences-discuss.net/9/14437/2012/bgd-9-14437-2012-discussion.html) show CMIP5 performance against two different soil databases. For future carbon changes, it would seem key to get the initial state right as well as the fluxes, and for land-use emissions it is crucial to get the right biomass in order to simulate the emissions from chopping it down [Chris Jones, United Kingdom of Great Britain & Northern Ireland]	Accepted and taken into account - we now briefly mention existing published evaluations of land carbon stocks in the CMIP5 models.
9-1365	9	42	40			Section 9.4.5: suggest to refer to the carbon cycle modelling sections in Chapter 6 (e.g., section 6.4 and Box 6.4) in the relevant sections of Chapter 9, in particular section 9.4.5.3 [Thomas Stocker/ WGI TSU, Switzerland]	Accepted and taken into account.
9-1366	9	42	47	42	47	Here and in other parts of the assessment report, please use the correct spelling: "El Niño" y la "La Niña" [Juan Blanco, Spain]	Editorial
9-1367	9	42	54	42	54	carbon stored in forests and soils [European Union]	Noted, but sub-sections 9.4.5.1 and 9.4.5.2 have been deleted as they relate to evaluations of offline carbon cycle models, and therefore fall under the remit of Chapter 6 rather than Chapter 9.
9-1368	9	42	54	43	3	This subsection is heavily based on the one modeling study (Randerson et al., 2009) while there are more studies focusing on Amazon forest. Therefore I suggest to revise this paragraph about Amazon as: "Accurate simulation of the Amazon is important for representing its buffering of atmospheric CO2 and for projecting the effects of climate change on the amount of carbon stored in forests (Lewis et al., 2011). While two	See response to 9-1367

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						<p>biogeochemical sub-models overestimated the above-ground live biomass by 130 to 190% and underestimated soil carbon by 33 to 40% (Randerson et al., 2009), other DGVMs simulated vegetation carbon within the uncertain range of the observationally-based estimation (Galbraith et al. 2010). These bias and lack of representation of key eco-hydrological mechanisms in the models are recognized (Sakaguchi et al., 2011) and are being improved (e.g., Baker et al. 2009; Bonan et al., 2011)." Galbraith et al. (2010) is already included in the current reference list.</p> <p>(Galbraith, D., P. E. Levy, S. Sitch, C. Huntingford, P. Cox, M. Williams, and P. Meir, 2010: Multiple mechanisms of Amazonian forest biomass losses in three dynamic global vegetation models under climate change. <i>New Phytologist</i>, 187, 647-665.)</p> <p>Baker, I. T., L. Prihodko, A. S. Denning, M. Goulden, S. Miller, and H. R. da Rocha (2008), Seasonal drought stress in the Amazon: Reconciling models and observations, <i>J. Geophys. Res.</i>, 113, G00B01, doi:10.1029/2007JG000644.</p> <p>Bonan, G. B., P. J. Lawrence, K. W. Oleson, S. Levis, M. Jung, M. Reichstein, D. M. Lawrence, and S. C. Swenson (2011), Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, <i>J. Geophys. Res.</i>, 116, G02014, doi:10.1029/2010JG001593.</p> <p>Sakaguchi, K., X. Zeng, B. J. Christoffersen, N. Restrepo-Coupe, S. R. Saleska, and P. M. Brando (2011), Natural and drought scenarios in an east central Amazon forest: Fidelity of the Community Land Model 3.5 with three biogeochemical models, <i>J. Geophys. Res.</i>, 116, G01029, doi:10.1029/2010JG001477. [Koichi Sakaguchi, the United States]</p>	
9-1369	9	42	57	42	57	the cited reference should be Randerson et al., 2009 instead of Lewis et al., 2011. I suggest another revision for the same paragraph below, and this is incorporated in it. [Koichi Sakaguchi, the United States]	See response to 9-1367
9-1370	9	43	1	43	1	underestimated "soil carbon" or "soil carbon and belowground plant biomass"? [European Union]	See response to 9-1367
9-1371	9	43	1	43	3	What is the uncertainty in the observations? [Ronald Stouffer, United States of America]	See response to 9-1367
9-1372	9	43	3	43	3	underestimation of root respiration, plant respiration in total or aboveground biomass respiration? All this points towards wrong plant parameterization and C allocation patterns. Why not directly mention? [European Union]	See response to 9-1367
9-1373	9	43	3	43	3	... Underestimation of... [European Union]	See response to 9-1367
9-1374	9	43	5	43	6	comparing a gross flux (from fire) with the net flux from fossil fuel isn't really a like-for-like comparison and risks overstating the importance of fire (or understating the importance of fossil fuel emissions). Vegetation regrows after burning, so the NET flux from fire is small unless the total fire activity changes over a sustained period. [Chris Jones, United Kingdom of Great Britain & Northern Ireland]	See response to 9-1367
9-1375	9	43	5	43	11	One aspect often missing in fire models used in ESMs are human fire mitigation and prevention strategies. This should be mentioned here at least, as it is a major lack of realism in current models. [Alexander Loew, Germany]	See response to 9-1367
9-1376	9	43	10	43	10	Terms such as "modest agreement", "reasonable agreement", "fairly well", "reasonably good" etc, are used throughout Chapter 9 when comparing models with observations. It is not clear how these terms rate sequentially - is 'modest agreement' a better result than 'reasonable agreement'? It would be good if more quantitative terminology could be used where ever possible. [Thomas Stocker/ WGI TSU, Switzerland]	Agreed and taken into account - replaced by more quantitative calibrated language.
9-1377	9	43	11			Statement too broad. GFDL ESMs (ESM2M and ESM2G) have a simple fire model. Shevliakova et al. 2009 [Ronald Stouffer, United States of America]	See response to 9-1367
9-1378	9	43	22	43	22	"vertical structure near the surface" I would add "density" after "vertical" [James Christian, Canada]	See response to 9-1367
9-1379	9	43	29	43	29	"greater outflow of CO2" outgassing [James Christian, Canada]	See response to 9-1367

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1380	9	43	39	45	32	Section 9.4.5.3 (The Carbon Cycle in Earth System Models) fails to mention that emissions-forced simulations with fully interactive terrestrial and ocean biosphere models are also part of the CMIP5 experimental design. These emissions-driven simulations provide the opportunity to assess the implications of biases resulting from uncertainties associated with ecosystem processes and feedbacks as the effects of those uncertainties propagate through the coupled ESM. Comparison of prognostic atmospheric CO2 mole fraction over the historical period from 15 CMIP5 models indicated that ESMs, on average, have a high bias in their predictions of contemporary atmospheric CO2 and that this bias persists throughout the 21st century in simulations forced with RCP 8.5 equivalent emissions (Hoffman et al., in prep.). [Government of United States of America]	Accepted and taken into account - we now make it clear that emissions-driven runs were also carried-out with the CMIP5 models.
9-1381	9	43	39	45	32	9.4.5.3 Model ocean fluxes should be compared also with those of Sabine et al. and Katiwala et al., which include uncertainties, including in the figure. Land fluxes can be calculated as a difference and also projected onto the figure with propagated uncertainties. [Government of United States of America]	Rejected. We believe that we already have enough fluxes inferred from observations to make a meaningful comparison with the models.
9-1382	9	43	39	45	32	9.4.5.3 Because future fluxes depend upon existing carbon stores, carbon inventories could also usefully be evaluated here. Spatially and temporally-adjusted observationally based estimates of ocean fluxes from Khatiwala et al., with uncertainties, can be integrated to provide an observational range for total anthropogenic carbon accumulation in the ocean and on land. [Government of United States of America]	Accepted. See response to 9-1364.
9-1383	9	43	45			Just in case the copy editor for chapter 9 is not the same as for chapter 6, in chapter 9 C4MIP has the "4" as a superscript, which is not the case in Chapter 6. [Adrian Simmons, United Kingdom]	Noted. The correct nomenclature is however C4MIP - as used in this chapter.
9-1384	9	43	52	43	52	change "(Friedlingstein et al., (2006))" to "(Friedlingstein et al., 2006)" [SAMIR POKHREL, INDIA]	Editorial
9-1385	9	43	52	43	52	Pls change "(Friedlingstein et al., (2006))" to "(Friedlingstein et al., 2006)" [HASIBUR RAHAMAN, India]	Editorial
9-1386	9	44	1	44	5	Using prescribed RCP's exclude biosphere feedbacks on atmospheric CO2 concentrations. How is this handled now, since already the C4MIP runs showed a huge effect of the biosphere on atmospheric CO2 concentrations. For me the argumentation not do so is rather illogical and unrealistic having in mind that anthropogenic CO2 emissions are <6-7% of biosphere-atmosphere C exchange [European Union]	The carbon cycle feedbacks are seen through their impact on the CO2 emissions compatible with the prescribed RCP, as explained in the text.
9-1387	9	44	3	44	3	you could also site Jones et al (J. Clim, submitted) who calculate compatible emissions for CMIP5 models, and also refer back to chapter 6, sec 6.3 for these from CMIP5. [Chris Jones, United Kingdom of Great Britain & Northern Ireland]	Accepted and taken into account.
9-1388	9	44	10	44	10	and elsewhere e.g. Figures 9.27 and 9.45 "fgCO2". I don't see the point of introducing a new jargon term here. The CMIP5 variable name is fgco2. I generally think using these names directly is bad idea, but changing the case is even worse. [James Christian, Canada]	Accepted and taken into account.
9-1389	9	44	10	44	10	year of the reference is missing [Farahnaz Khosrawi, Sweden]	Editorial
9-1390	9	44	10	44	11	skip or update Anav et al., Submitted [European Union]	Editorial
9-1391	9	44	13	44	13	Pls change "(Le Quere et al., (2009))" to "(Le Quere et al., 2009)" [HASIBUR RAHAMAN, India]	Editorial
9-1392	9	44	16	44	16	skip or update Anav et al., Submitted [European Union]	Editorial
9-1393	9	44	16	44	16	year of the reference is missing [Farahnaz Khosrawi, Sweden]	Editorial
9-1394	9	44	16	44	18	Please comment on the key results of the evaluation. [Government of Germany]	Accepted and taken into account - we now include EMIC results in Figure 9.27, and discuss these in the revised text.
9-1395	9	44	16	44	18	please provide a summary of the results from the EMICs model evaluation mentioned, rather than just referring to it mentioning that it has been carried out. In particular as these results are mentioned in the section summary on page 45, line 23/24. [Thomas Stocker/ WGI TSU, Switzerland]	See response to 9-1395.
9-1396	9	44	25	44	25	I don't see "especially after the Pinatubo volcano in 1991" this in the graph. Overinterpretation [European Union]	Rejected. An increase in the land carbon sink is clearly visible at the time of Pinatubo.
9-1397	9	44	25	44	25	Write here "Pinatubo eruption" instead of "Pinatubo volcano". [Farahnaz Khosrawi, Sweden]	Accepted.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1398	9	44	32	44	32	skip or update Anav et al., Submitted [European Union]	Editorial
9-1399	9	44	32	44	32	year of the reference is missing [Farahnaz Khosrawi, Sweden]	Editorial
9-1400	9	44	34	44	34	Insert Figure 9.26 already here instead of later [Farahnaz Khosrawi, Sweden]	Editorial
9-1401	9	44	34	44	50	It is a bit surprising that no text is devoted to the large discrepancy between modeled and observed CO2 fluxes over land for the Northern hemisphere and the Tropics [Bart Van den Hurk, Netherlands]	Rejected. We already point-out these model discrepancies on lines 47 to 50 of page 44.
9-1402	9	44	35	44	35	After "GCP" please consider to insert "(only global estimates)". [Government of Germany]	Accepted.
9-1403	9	44	36	44	36	Could you spell out what "JMA" represents? [Takamitsu Ito, United States of America]	Accepted.
9-1404	9	44	43	44	43	To my knowledge NorESM1-ME and NorESM2-M use the atmosphere and land models of CCSM4 [European Union]	Noted.
9-1405	9	44	43	44	43	NorESM2-M does not exist. Please correct to: NorESM1-M [Trond Iversen, United Kingdom of Great Britain & Northern Ireland]	Accepted and taken into account.
9-1406	9	44	52	44	52	Insert Figure 9.27 already here instead of later [Farahnaz Khosrawi, Sweden]	Editorial
9-1407	9	44	54	44	55	INMCM4 has an extremely simple ocean biogeochemistry model. I don't know what they assume for alkalinity but they do not seem to have staged any data for it. The reference given (Volodin 2007) is incomplete. [James Christian, Canada]	Noted, but this model is no longer included in the analysis presented in Figure 9.27.
9-1408	9	44	56	45	4	uncertainty range for estimates are not provided [European Union]	Noted, but we feel that providing uncertainty estimates on each of the figures quoted here would make this very difficult to read and does not affect the overall messages.
9-1409	9	45	1	45	3	Pls change "(Takahashi et al., 2009)" to "(Takahashi et al., 2009)" [HASIBUR RAHAMAN, India]	Rejected - second bracket is necessary here.
9-1410	9	45	7	45	11	From the figure one might infer that there has been only one El Nino. Since modelled El Nino events and observed El-Nino events do not need to match (as stated in the text), the reader might get a wrong impression. Suggestion: Either indicate all El-Nino years above a certain threshold and comment that this is only used to illustrate the effect of El Nino on the observed fluxes or delete the orange line. [Government of Germany]	Accepted and taken into account - line representing the El Nino of 1998 has been removed.
9-1411	9	45	7	45	11	Please indicate the same number of volcanic eruptions as shown for the graph on global mean temperature in Fig. 9.8. [Government of Germany]	For clarity we have decided to remove the volcanic eruptions completely from this plot.
9-1412	9	45	21	45	21	What does "can simulate" mean? (time scale, spatial scale, consistency bounds?) [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	This means "within the range of observational estimates", as stated at the end of this sentence.
9-1413	9	45	21	45	32	What does this mean for simulated climate change? Are we ok if we get the global number right? [Peter Braesicke, United Kingdom]	Noted. To be frank we don't yet know in detail.
9-1414	9	45	24	45	24	This underestimation is not evident from figures 9.26 and 9.27 [Bart Van den Hurk, Netherlands]	This is because the EMICS were not in this version of the figure. They have been added now though.
9-1415	9	45	26	45	28	"outgassing in the tropics" in line 26 <-> "sink in the tropics" in line 28. Isn't it a contradiction? [Barbara Früh, Germany]	No, because the "outgassing" refers to the ocean carbon cycle while the "sink in the tropics" refers to the land.
9-1416	9	45	30	45	31	N fertilization (direct and indirect by atmospheric N deposition, leaching, erosion etc.) happens as well in the subtropics (e.g. India, while China would mostly count as mid-latitudes). N fertilization may moreover drive C cycle in tropical areas. See e.g. Xia and Wan, 2008 Global response patterns of terrestrial plant species to nitrogen addition, New Phytol 179, 428-439 [European Union]	Noted, but we were hypothesising here about the cause of the underestimation of the carbon sink in the mid-latitude northern hemisphere.
9-1417	9	45	34	46	46	Chapter 9: Section 9.4.6 (Simulation of Recent and Longer Term Records in Global Models. Aerosol Burdens and Effects on Insolation). This section does not contain any introduction neither summary. It would be	Accepted -- introductory statements (on the relationship between trends in tropospheric burdens

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						desirable to have both of them as in previous [Government of Spain]	and why AOD should reflect these trends) and summary statements (on the impact of model biases on the anthropogenic forcing calculated with them) have been added.
9-1418	9	45	38	45	45	The paragraph seems too focused on sulphate. Is it the same problem in the models that they include a too limited set of aerosol types and thus underestimate the AOT? [Gunnar Myhre, Norway]	Noted -- but sulphate dominates the anthropogenic AOD under present-day conditions (according to AR4 forcing chapter) and hence the emphasis here on sulphate is appropriate.
9-1419	9	45	38			Which is the observational uncertainty related to Figure 9.28? [Massimo Bollasina, Italy]	Accepted -- Observational uncertainty has been added to discussion.
9-1420	9	45	43	45	45	Land MODIS AOT retrievals are of inferior quality than over ocean. AOT retrieval is also of questionable quality over persistently cloudy areas like the high latitude southern oceans. The authors should consider providing clarification of this issue in the discussion. [Government of United States of America]	Accepted -- MISR retrievals have been substituted for MODIS over land surfaces in figure 9-28.
9-1421	9	45	44	6	9	Again for inter-chapter consistency, AOT is used here (aerosol optical thickness, though the acronym is not defined, which should be remedied if use of AOT is retained) whereas Chapter 7, devoted to aerosols and clouds, uses AOD (aerosol optical depth). AOD is also what is used in chapter 11. [Adrian Simmons, United Kingdom]	Accepted -- AOD replaced with AOT
9-1422	9	45	44	45	44	... species... [European Union]	Editorial
9-1423	9	45	44	45	44	AOT should be spelled out and specified w.r.t. wavelength: "... Aerosol optical thickness (AOT) at $\lambda = ?$ nm..." [European Union]	Editorial
9-1424	9	45	44	45	44	"AOT" please make sure to spell out acronyms at their first appearance. [Government of Germany]	Taken into account -- combined with comment 9-1423
9-1425	9	45	44	45	45	the period 2001 to 2005, CMIP5 models underestimate the mean AOT due to all tropospheric aerosol species needs to be put into context. My impression was that this underestimate is more systematic than this period and also exists in earlier periods. [Steven Smith, United States of America]	Accepted -- The persistence of the bias from the start of the runs is now discussed, and the ACCMIP figure showing a similar result is cited.
9-1426	9	45	44			explain abbreviation "AOT" [Barbara Früh, Germany]	Taken into account -- combined with comment 9-1423
9-1427	9	45	44			AOT. Quantity should be denoted AOD aerosol optical depth (vertical integral of extinction coeff), not AOT, aerosol optical thickness, which is integral along some path. [Stephen E Schwartz, United States of America]	Taken into account -- combined with comment 9-1421
9-1428	9	45	53	45	53	... Europe and North America... ? [European Union]	Editorial
9-1429	9	45	53	45	54	Very specific, no need to specify an individual model [Massimo Bollasina, Italy]	Rejected -- The statement regarding the fidelity of the specific model in question (HadGEM2) to the transition from global dimming to brightening is not true of all models in the CMIP5 archive.
9-1430	9	45	57	45	57	... to... [European Union]	Editorial
9-1431	9	45	57	45	57	typo oceanic [Ian Watterson, Australia]	Editorial
9-1432	9	46	7	46	8	Which period is used for the models? [Andreas Sterl, Netherlands]	Accepted -- The period used for the figure has been added to the caption.
9-1433	9	46	10	46	14	In Figure 9.28 legend, names of the CMIP5 models are specified while in other Figures not [Government of Spain]	Taken into account -- combined with comment 9-2412
9-1434	9	46	10	46	14	Skip the list of models. [Farahnaz Khosrawi, Sweden]	Taken into account -- combined with comment 9-2412
9-1435	9	46	10			Figure 9.28: We note that "the data version is MODIS Collection 4" and recommend that the figure should be updated with more recent AOT data. Collection 5 has been out for many years and the release of Collection 6 is imminent. In my opinion, Also, the MODIS dataset used for this figure is not included in Table 9.2. If the	Taken into account -- combined with comment 9-2414. Table 9-2 has been updated with the relevant information regarding the MODIS aerosol retrievals.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						figure is not updated with newer data, please change "The data version is MODIS 4" to "The data version is MODIS Collection 4". [Government of United States of America]	
9-1436	9	46	18			This comment refers to the phrase: "corrected MODIS satellite observations". We recommend that the authors provide more specifics besides the journal references. Is the dataset different than the one used for Fig. 9.28? If so, it should be also listed in Table 9.2. [Government of United States of America]	Accepted -- Information regarding these corrected data sets has been added to table 9-2.
9-1437	9	46	34	46	34	Something is wrong in this sentence. [Farahnaz Khosrawi, Sweden]	Editorial
9-1438	9	46	48	47	17	This introductory section should explain the structure of the later sections: "In this section, natural climate variability on timescales from X to Y is evaluated". [Josephine Brown, Australia]	taken into account. This section is not limited to particular time scales, but rather assesses variability in general.
9-1439	9	46	48			Shouldn't the implication of variability on the comparison between model and obs be discussed in more detail here? This is particularly relevant for trends (see recent work by Mahlstein et al GRL/ERL., Hawkins and Sutton BAMS/GRL, Deser et al. Clim, Dyn./Nature Climate Change, etc.). Even with perfect models we would never expect perfect agreement. [Reto Knutti, Switzerland]	taken into account. This section is really about the ability of models to simulate climate variability, but the text there has been revised to include a more general introduction to the topic.
9-1440	9	46	55	46	55	This paragraph leaves the impression that to simulate the unforced variability is just important. I would say that today it is fundamental, and even the core of climate modelling. And not just because it is important for the estimation of the signal-to-noise ratio or to make quantitative estimation of variability changes in the future. It is important to aim to a realistic climate simulation. [Ramon de Elia, Canada]	noted. Variability is an important aspect of evaluating model quality and that is why an entire section of the report is devoted to this topic. The importance relative to simulating mean state of trend depends on the particular application or question, and we refrain from making comparative statements of importance.
9-1441	9	47	2	47	5	Which aspects of "the specific evolution" are being referred to here? If it means "weather" then we know the predictability horizon already. This is an important point and could be expanded into an FAQ? Which aspects of the "specific evolution" do we need to be able to simulate, in order to have confidence in the projection? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	taken into account. Text revised to make it clearer.
9-1442	9	47	2			We recommend that the authors change "extreme rainfall" to "extreme precipitation" to include the consideration of snow. [Government of United States of America]	accepted.
9-1443	9	47	7	47	12	Good point, should come earlier! [Peter Braesicke, United Kingdom]	noted. In the interests of space and readability, we have tried not to overload the introductory sections, and prefer to leave specifics like this to the sections in which they seem 'natural'.
9-1444	9	47	9	47	12	for the first part of the sentence I can follow; however, to conclude the "extension..." is not comprehensible [Barbara Früh, Germany]	taken into account. Text modified.
9-1445	9	47	15	47	15	... evolution are... [European Union]	editorial
9-1446	9	47	15	47	15	Add a space between "evolution" and "are" [Hai Lin, Canada]	editorial
9-1447	9	47	19			9.5.2 The authors may want to augment the discussion in this section with a discussion of seasonal variability that goes beyond the discussion of seasonal features such as monsoons, the MJO, and blocking patterns. Such a discussion would parallel section 9.5.2.1 nicely. A paper that could be consulted for this discussion is Fu et al. ("Assessing Future Changes of Climate and Drought over the South-Central United States Projected by the CMIP5 Models"), which, in spite of its title, includes evaluations of the historical simulations in the CMIP5 models, and statements on their skill in simulating various seasonal fields. [Government of United States of America]	rejected. The discussion of seasonal and climatological aspects can be found in section 9.4
9-1448	9	47	21	47	21	I suggest to change the section title from "Diurnal Cycles of Physical Climate Variables" to "Diurnal cycles of surface air temperature and precipitation" since we only briefly mention surface air temperature and mainly focus on precipitation. Unfortunately I am not aware of other studies focused on the diurnal cycles of essential climate variables. [Koichi Sakaguchi, the United States]	Accepted Done
9-1449	9	47	21			Section 9.5.2.1: The discussion of the models' ability to represent the diurnal cycle of rainfall should, like	Taken into account

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Figure 9.30, be divided into land and sea, since the behaviour and errors are quite different. The tendency to start moist convection prematurely is largely limited to land, while the tendency to rain too frequently at reduced intensity (on a daily mean basis) is mostly seen over ocean. Similarly, the possible parametrisation deficiencies relating to the diurnal cycle errors will also differ between land and sea. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	
9-1450	9	47	23	47	26	I would explicitly mention, among the “many other geophysical parameters” for which there are “easily observable diurnal variations”, certain types of clouds in certain regions. [Robert Kandel, France]	Taken into account. However we only refer here to publication for which model-data comparison are available for the same model version as the one used for future climate projection (mainly CMIP3 and CMIP5 simulations)
9-1451	9	47	23	47	26	Sentence does not read well [Bart Van den Hurk, Netherlands]	Accepted, revised
9-1452	9	47	28	47	28	Over which continent? All of them? Similar comment for line 35. [Government of Australia]	Noted. The first statement is a generic one (taken from the AR4). The second is more specific to tropical lands
9-1453	9	47	28	47	28	magnitude of what? Diurnal cycle? [Bart Van den Hurk, Netherlands]	Taken into account. yes of diurnal cycle. Precision included
9-1454	9	47	29			The underestimation of the amplitude is even stronger over ocean areas [Barbara Früh, Germany]	Taken into account.regions here refer to both land and ocean
9-1455	9	47	29			The underestimation of the amplitude is even stronger over the oceans [Government of Germany]	Taken into account.regions here refer to both land and ocean
9-1456	9	47	30	47	30	What is meant by 'spatial distribution of the diurnal amplitude'? The variation shown is in time, not space. [Government of Australia]	Noted. the spatial pattern is referenced in Wang et al. 2011, not in the figure.
9-1457	9	47	31	47	32	The two processes (early onset of moist convection in the diurnal cycle) and the tendency for too frequent rain at reduced intensity (on a daily mean basis) are not necessarily related (one is largely restricted to land and the other is over land and ocean) and should not be discussed together (lines 31and 32). While the early and short diurnal cycle of rainfall over land, and its tendency to occur every day, may indeed result in too persistent low amounts of daily mean rainfall over land, other issues such as convective intermittency (where parametrised convective rain is actually very intense in a model timestep but is short-lived and with gaps) can contribute to the mid-named "drizzle bias" as much as actual over-persistence of light stratiform rain. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	taken into account. Results for land and ocean are now treated separately.
9-1458	9	47	34			I think it would be interesting to talk more about the coupling of the submodels; reading this I started sketching out a coupling diagram, something like this might be useful to include, illustrating the point that our submodels talk to each other and how they talk to each other (and when). [Peter Braesicke, United Kingdom]	Rejected. Although this is interesting, the length limit of the chapter do not allow to go too far in the description of couplings in this section
9-1459	9	47	36	47	36	"In addition" - should this be "In contrast"? [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Taken into account.
9-1460	9	47	39			Additional relevant references: Pastor MA,Casado MJ (2012) Use of Circulation Types to evaluate AR4 climate models over the Euro-Atlantic region. Clim Dyn 39:2059–2077. doi:10.1007/s00382-012-1499-2 [Government of Spain]	Taken into account. Note however that this is an assessment and that due to space limit it is not possible to include all references on the subject
9-1461	9	47	48	47	48	"These conclusions" - which conclusions? [Josephine Brown, Australia]	Taken into account. The section has been slightly reorganised
9-1462	9	48	5	48	8	To avoid any ambiguity, I suggest adding "oceanic" before "convection" l.5-6 and "atmospheric" before "convection" l.7-8 [RYM MSADEK, United States of America]	rejected, the process discussed here concern the model parameterization and are not specific to land or ocean convection.
9-1463	9	48	10	48	10	Chapter 9: Section 9.5.2 (Simulation of Variability and Extremes. Diurnal to Seasonal Variability). It seems that a Subsection 9.5.2.2 comprising Blocking and MJO is lacking. Please revise it. [Government of Spain]	Accepted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1464	9	48	10	48	10	9.5.2.2.1 Blocking should be: 9.5.2.2 Blocking [Koichi Sakaguchi, the United States]	Accepted
9-1465	9	48	10	48	13	In Chapter 2, it says Dole (2011) was referenced here, but I could not find this. The comment for Ch 2 page 67 line 23 above would also be fitting here. [Anthony Lupo, United States of America]	Noted. The Dole paper doesn't really address evaluation of model blocking so it not relevant here.
9-1466	9	48	10			Section 9.5.2.2.1 Blocking: Please refer to and check with Ch14, Section 14.6.3, regarding their assessment of Blocking, with a statement elevated to the Executive Summary. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account
9-1467	9	48	11	48	13	In this context you can also cite Pfahl, S. and H. Wernli (2012), Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the Northern Hemisphere on (sub-)daily time scales, Geophys. Res. Lett., 39, L12807, doi:10.1029/2012GL052261. They underline the importance of blocking for summer temperature extremes, whereas Sillmann et al. 2011 and Buehler et al. 2011 focus on winter temperature extremes. [Jana Sillmann, Canada]	Taken into account. Reference included
9-1468	9	48	15	48	16	Please provide a reference for this statement? [Government of United States of America]	Taken into account referring to Scaife et al (2010) which is already cited.
9-1469	9	48	16			I believe this text on models' ability to simulate blocking is too dismissive, and the impression is given that only high resolution climate models can produce blocking. However, at least one standard resolution model has simulated a reasonable representation of blocking; possible text to insert here could be: "...occurrence of blocking. However, recent work has shown that a standard resolution global coupled climate model can simulate many aspects of blocks in the Northern Hemisphere related to number and location (Sillmann and Croci-Maspoli, 2009). Recent work has shown..." Reference: Sillmann, J., and M. Croci-Maspoli, 2009: Present and future atmospheric blocking and its impact on European mean and extreme climate, Geophys. Res. Lett., 36, L10702, doi:10.1029/2009GL038259. Reference: Sillmann, J., and M. Croci-Maspoli, 2009: Present and future atmospheric blocking and its impact on European mean and extreme climate, Geophys. Res. Lett., 36, L10702, doi:10.1029/2009GL038259. [Gerald Meehl, United States of America]	Taken into account. The text has been clarified. In particular, the blocking index used by Sillmann and Croci-Maspoli is different to many others used in that it takes an anomaly rather than flow reversal approach. Woollings (2010, Phil Trans) showed that this model indeed suffers from the conventional bias when a conventional blocking index is used. However, it's certainly not clear which blocking index is better. This issue is discussed in the second paragraph here, which already mentions this paper. Text has been revised to make things clearer.
9-1470	9	48	18	48	18	Add a period after "2010b)" [Hai Lin, Canada]	editorial, corrected
9-1471	9	48	18	48	18	Pls add "." after "Matsueda et al., 2010b) " [HASIBUR RAHAMAN, India]	editorial. Corrected
9-1472	9	48	18			... 2010b). (typo) [Government of France]	editorial. Corrected
9-1473	9	48	31	48	32	(e.g., Sillmann and Croci-Maspoli, 2009). [Hai Lin, Canada]	editorial. corrected
9-1474	9	48	31	48	32	change "(eg (Sillmann and Croci-Maspoli, 2009))" to "(eg Sillmann and Croci-Maspoli, 2009)" [SAMIR POKHREL, INDIA]	editorial corrected
9-1475	9	48	31	48	32	Pls change "(eg (Sillmann and Croci-Maspoli, 2009))" to "(eg.,Sillmann and Croci-Maspoli, 2009)" [HASIBUR RAHAMAN, India]	editorial corrected
9-1476	9	48	31	48	32	there are extra parenthesis in the citation, which should be removed. [Jana Sillmann, Canada]	editorial corrected
9-1477	9	48	31	48	32	In this context you should also cite Scaife et al. 2010 as they actually show the differences between the different blocking index definitions in their study. [Jana Sillmann, Canada]	Taken into account
9-1478	9	48	34	48	34	add "e.g.", Martius et al., 2009 is not the only publication showing this. [Farahnaz Khosrawi, Sweden]	editorial. Taken into account
9-1479	9	48	40	48	41	There is NWP experience in improving the forecasting of blocking through model improvement could perhaps be drawn on here also. It may help raise the confidence level above "medium". [Adrian Simmons, United Kingdom]	Taken into account with reference to Matsueda, M. (2009). Blocking predictability in operational medium-range ensemble forecasts. SOLA, 5(0), 113-116.
9-1480	9	48	40	48	41	The term "high-top models" has not been used in this section before. To avoid confusion this term should be introduced in the foregoing paragraph (lines 34-38), or line 41 should be reformulated as "... that models with a good representation of the stratosphere better represent blocking". [Andreas Sterl, Netherlands]	Taken into account. The last paragraph has been rewritten
9-1481	9	48	43	48	43	9.5.2.2.2 Madden Julian Oscillation should be: 9.5.2.3 Madden Julian Oscillation [Koichi Sakaguchi, the United States]	Accepted.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1482	9	48	43			9.5.2.2.2 The authors should include a discussion of the paper titled “MJO and Convectively Coupled Equatorial Waves Simulated by CMIP5 Climate Models” by Hung et al., which details an examination of MJO fidelity in the CMIP5 models. The paper concludes that the CMIP5 models are better than the CMIP3 models at capturing tropical intraseasonal variability, and that the associated Equatorial-region waves are also better captured in the CMIP5-era models. The authors might also consider Jiang et al. (“Simulations of the Eastern North Pacific Intraseasonal Variability in CMIP5 GCMs”), who evaluate intraseasonal variability over the Eastern North Pacific, including expressions of the MJO on this region. In addition, the authors may wish to consider examining and integrating the discussion in section 3.1 of the Sheffield et al. paper (“North American Climate in CMIP5 Experiments. Part II: Evaluation of 20th Century Intra-Seasonal to Decadal Variability”) into this section of the chapter. [Government of United States of America]	Taken into account. However this section does not intend to evaluate MJO in all regions, but only to provide an global view of model ability to reproduce it. Hung et al referenc included
9-1483	9	48	43			Section 9.5.2.2.2: Hung et al. (2012) has evaluated the MJO and convectively coupled equatorial waves in 20 CMIP5 models, and compared the results with CMIP3 models evaluated by Lin et al. (2006). The citation is: Hung, M.-H., J. L. Lin, W. Wang, D. Kim, T. Shinoda, and S. Weaver, 2012: MJO and Convectively Coupled Equatorial Waves Simulated by CMIP5 Climate Models. J. Climate, revised. The results show that the CMIP5 models have an overall improvement over the CMIP3 models in the simulation of tropical intraseasonal variability, especially the MJO and several CCEWs. The CMIP5 models generally produce larger total intraseasonal (2-128 day) variance of precipitation than the CMIP3 models, as well as larger variances of Kelvin, equatorial Rossby (ER), and eastward inertio-gravity (EIG) waves. Almost all models have signals of the CCEWs, with Kelvin and MRG-EIG waves especially prominent. The phase speeds, as scaled to equivalent depths, are close to the observed value in 10 of the 20 models, suggesting that these models produce sufficient reduction in their "effective static stability" by diabatic heating. The CMIP5 models generally produce larger MJO variance than the CMIP3 models, and a more realistic ratio between the variance of the eastward MJO and that of its westward counterpart. About one third of the CMIP5 models generate the spectral peak of MJO precipitation between 30-70 days, but the model MJO period tends to be longer than observations as part of an over-reddened spectrum, which in turn is associated with too strong persistence of equatorial precipitation. Only one of the 20 models is able to simulate a realistic eastward propagation of the MJO. [Jialin Lin, United States of America]	Taken into account. However this section does not intend to evaluate MJO in all regions, but only to provide an global view of model ability to reproduce it. Hung et al referenc included
9-1484	9	48	46	48	46	remove the "(" at the beginning of this line. [Hai Lin, Canada]	Editorial
9-1485	9	48	49			A useful review reference spanning a range of issues surrounding the mJO and its interactions with monsoons, enso, midlatitudes, etc is: Lau, W. K. M., and D. E. Waliser (Eds.) (2011), Intraseasonal Variability of the Atmosphere–Ocean Climate System, 2nd Edition, 613 pp., Springer, Heidelberg, Germany. [duane waliser, United States of America]	Taken into account. However, given the space limite of the section we are not able to include all references provided a priority to peer reviewed journals literature and multi model analyses. Additional aspects of MJO (at regional scale in particular) is provided in chapter 14
9-1486	9	48	54			Lin et al. (2008) also evaluated the boreal summer intraseasonal oscillation (BSIO) in CMIP3 models and the results were different from those of Sperber and Annamalai (2008). The citation is: Lin, J. L., K. M. Weickmann, G. N. Kiladis, B. E. Mapes, S. D. Schubert, M. J. Suarez, J. T. Bacmeister, and M.-I. Lee, 2008: Subseasonal variability associated with Asian summer monsoon simulated by 14 IPCC AR4 coupled GCMs. J. Climate, 21, 4541-4567. They found that the models are able to simulate the northward propagation of BSIO but can not simulate well the eastward propagation. This is consistent with the models difficulty in simulating eastward propagation of winter time MJO and also MJO in North American monsoon (Lin, J. L., B. E. Mapes, K. M. Weickmann, G. N. Kiladis, S. D. Schubert, M. J. Suarez, J. T. Bacmeister, and M.-I. Lee, 2008: North American monsoon and convectively coupled equatorial waves simulated by IPCC AR4 coupled GCMs. J. Climate, 21, 2919-2937). Currently, eastward propagation of MJO and BSIO is still a challenge for global climate models. [Jialin Lin, United States of America]	taken into account. However it is not possible given the length limit of the section to include all the references, There is no fondamental contradiction between the two studies.
9-1487	9	48	56			"...and Jiang et al. (2011) have shown that a few models represent the MJO expression in the eastern Pacific quite well". ----- Jiang, X., D. E. Waliser, D. Kim, M. Zhao, M. Khairoutdinov, W. Stern, S. D. Schubert, K. R. Sperber, G. J. Zhang, W. Wang, R. Neale, and M.-I. Lee (2011), Simulation of the	taken into account, reference included

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Intraseasonal Variability over the Eastern Pacific ITCZ in Climate Models, Climate Dynamics,, DOI 10.1007/s00382-011-1098-x. [duane waliser, United States of America]	
9-1488	9	48	57	49	58	...and Australia. Fan et al. (2010) also... [Hai Lin, Canada]	editorial not at the right place. It is 49 57 and 58. Done
9-1489	9	49	1	49	1	Write "to simulate". [Juan Blanco, Spain]	Taken into account.sentence rewritten
9-1490	9	49	4			diagnostics? [Peter Braesicke, United Kingdom]	Noted. Comment unclear.
9-1491	9	49	5			The studies by Zhang et al. and by Kim et al. would be paritcularly appropriate here: Kim, D., K. Sperber, W. Stern, D. Waliser, I. S. Kang, E. Maloney, W. Wang, K. Weickmann, J. Benedict, M. Khairoutdinov, M. I. Lee, R. Neale, M. Suarez, K. Thayer-Calder, and G. Zhang (2009), Application of MJO Simulation Diagnostics to Climate Models, Journal of Climate, 22(23), 6413-6436. ----- Zhang, C., M. Dong, H. H. Hendon, E. D. Maloney, A. Marshall, K. R. Sperber, and W. Wang (2006), Simulations of the Madden-Julian Oscillation in Four Pairs of Coupled and Uncoupled Global Models, Climate Dynamics, DOI: 10.1007/s00382-006-0148-2. [duane waliser, United States of America]	Taken into account, However given the space limit of the section we are not able to mention all the literature on the subject. Kim et al. Reference included
9-1492	9	49	10	49	10	skip or update Sperber et al., Submitted [European Union]	Accepted.
9-1493	9	49	10	49	10	Pls change "However, (Sperber et al., submitted)" to "However, Sperber et al., (submitted)" [HASIBUR RAHAMAN, India]	editorial. Text revised
9-1494	9	49	14	49	14	An overview reference for this paragraph would be: Sperber, K., J. Slingo, and P. Inness (2011), Modeling Intraseasonal Variability, in Intraseasonal Variability of the Atmosphere-Ocean Climate System, 2nd Edition, edited by W. K. M. Lau and D. E. Waliser, p. 613, Springer, Heidelberg, Germany. [duane waliser, United States of America]	see 9-1485
9-1495	9	49	14	49	20	Also mention of the importance of the mean state: e.g. Inness, P. M., J. M. Slingo, E. Guilyardi, and J. Cole (2003), Simulation of the Madden-Julian oscillation in a coupled general circulation model. Part II: The role of the basic state, Journal of Climate, 16(3), 365-382. [duane waliser, United States of America]	Taken into account. Reference included
9-1496	9	49	14			We recommend that the authors qualify the statement by adding "possibly". The statement, as it reads now seems too definitive. Better simulation can result for the wrong reason in some cases if there are large cancelling errors, etc. [Government of United States of America]	Accepted. statement revised
9-1497	9	49	15	49	15	... atmosphere... [European Union]	Editorial, corrected
9-1498	9	49	15	49	15	Correct the spelling of "atmosphere" [Hai Lin, Canada]	Editorial, corrected
9-1499	9	49	15			Typo "atmospherer" [Adrian Simmons, United Kingdom]	Editorial, corrected
9-1500	9	49	22	49	24	We suggest you replace "northeastward" with "northward" and add the following to the end of the sentence: ", which coincides with the eastward propagation seen year round in the region." [Government of United States of America]	rejected, the proposition will add confusion to the statement; However the text has been revisited.
9-1501	9	49	27	49	37	The contours showing the continental coasts are hard to distinguish from the eof-contours. [Government of Germany]	Taken into account, Figure updated
9-1502	9	49	32	49	37	Starting from "The maximum"; please consider to move this to the section text instead of the caption. [Government of Germany]	Editorial, done
9-1503	9	49	39	49	39	9.5.2.3 Large Scale Monsoon Rainfall and Circulation should be : 9.5.2.4 Large Scale Monsoon Rainfall and Circulation [Koichi Sakaguchi, the United States]	Accepted
9-1504	9	49	39			Section 9.5.2.3: suggest to ensure consistency with assessment provided in Ch14, Section 14.2 [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. The point of view is different in the two chapters.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1505	9	49	41	49	41	Need to clarify which monsoon this refers to, ie. replace "The monsoon is ..." with "Monsoons are ..." (or possibly "The Asian monsoon is ...", but I think it's the former that's meant). [Dave Rowell, United Kingdom of Great Britain & Northern Ireland]	Accepted.
9-1506	9	49	41	49	47	chap 14 contains an extensive analysis of the CMIP5 models representation of the monsoons (global and individuals) amd their main aspects. It should be referenced here. [Annalisa Cherchi, Italy]	Accepted, done
9-1507	9	49	49	49	58	Could you also mention the fact that if models do not correctly include absorbing aerosols such as dust and carbonaceous, the possible effect of these aerosols on the monsoon circulation could lead to an error in the representation of the intensity of precipitation? (Here I refer at the Lau et al works about the Elevated Heat Pump mechanism. [Chiara Cagnazzo, Italy]	Taken into account , but the results are specific to the Asian monsoon. Reference added
9-1508	9	49	54			As written, the definition of the threat score is incorrect. Yes, it's a categorical metric, but it is not only applicable for monsoon precipitation, or only for precipitation, for that matter. It is a categorical comparison of a field to a threshold value, e.g., precipitation exceeding 20 mm or temperature exceeding 30 C. Thus, we recommend that the parenthetical definition in the text should be corrected. [Government of United States of America]	taken into account. Text modified accordingly
9-1509	9	49	55			What is a "monsoon precipitation domain"? We recommend that the authors clarify this concept. [Government of United States of America]	Accepted
9-1510	9	50	1			What about a longer record than the past half century? Are data available and if so what does the comparison show? [Government of United States of America]	Taken into account. we added a few words on the monsoon trends and past climate
9-1511	9	50	4	50	4	Since HadGEM3 is not among the CMIP5 models, it would be better to say that this this bias is common in CMIP5 models, as demonstrated in Levine et al. (2012; Climate Dynamics, revised). It is also visible in Fig. 9.2 (though small in the multi-model mean). [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Accepted. Text revised accordingly
9-1512	9	50	4	50	5	Very specific, I suggest to include the study in a new broader sentence pointing toward the role of air-sea interactions over the Indian Ocean and their poor representations in models (see also: Bollasina and Ming 2012). Bollasina, M., and Y. Ming, 2012: The general circulation model precipitation bias over the southwestern equatorial Indian Ocean and its implications for simulating the South Asian monsoon. Climate Dynamics, 10.1007/s00382-012-1347-7 [Massimo Bollasina, Italy]	Accepted Text rewritten to be less specific
9-1513	9	50	4	50	13	"some models" -- suggest to be more specific regarding what "some models" refers to here [Thomas Stocker/ WGI TSU, Switzerland]	Accepted. text revised.
9-1514	9	50	4			For the first time a particular CMIP5 model is mentioned (HadGEM3). This needs to be better justified [Government of Spain]	Taken into account. The reference to a specific model has been supress to make the point more generic.
9-1515	9	50	7	50	7	... contribute... [European Union]	Editorial
9-1516	9	50	13	50	13	Provide reference for this statement. [Josephine Brown, Australia]	Taken into account
9-1517	9	50	13	50	13	Add a period at the end of the sentence. [Hai Lin, Canada]	Editorial, done
9-1518	9	50	17			Please see comment 260. "observations from GPCP" could be changed to "observation-based estimate from GPCP" or "GPCP analysis of observations". [Adrian Simmons, United Kingdom]	taken into account
9-1519	9	50	23	50	25	"biases that limit model credibility" -> Can this summary statement make use of the uncertainty language to assign a confidence level? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account.uncertainty language now used in the final assessment
9-1520	9	50	23			General conclusions for CMIP5 models behaviour, taking into account only one specific CMIP5 model? [Government of Spain]	Taken into account. Assesment revised
9-1521	9	50	24	50	24	List or summarise the systematic biases. [Josephine Brown, Australia]	Taken into account. However the space is limited and everything cannot be listed. We only consider biases that affect the "global monsoon". Specif assesment at the regional scale is part of chapter 14

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1522	9	50	24	50	24	... but they still... [European Union]	Taken into account.final statement on this section has been revised
9-1523	9	50	24	50	24	Change "but the still suffer" to "but they still suffer" [Hai Lin, Canada]	Taken into account.final statement on this section has been revised
9-1524	9	50	27	50	27	Should this section be titled "Interannual to multi-decadal variability"? The first subsection deals with Last Millennium simulations, but the introductory paragraph specifies interannual to multi-decadal time scales. [Josephine Brown, Australia]	Taken into account. Introductory text modified
9-1525	9	50	27	57	43	Section 9.5.3 - Recommend that the teleconnection subtopics be discussed along with the individual modes of variability, rather than talking about teleconnections separately. The current form of the section detracts from highlighting the global impacts of the modes of variability. [Government of United States of America]	Taken into account. Text modified to re-organise section
9-1526	9	50	27			Section 9.5.3: please ensure consistency in the assessment of modes with Chapters 2, 3 and 14 [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Coordination with Chap 14 ensured.
9-1527	9	50	29	50	29	In addition to the diurnal-to-seasonal variability described above, ... [Hai Lin, Canada]	Taken into account. Text modified
9-1528	9	50	33	50	33	This sentence perpetuates a false and misleading dichotomy: human-induced climate change is not additive with the "regional manifestation of modes of variability"; instead, it may cause changes in those modes. This is an important communication point for the IPCC to stress; the effects of climate change are not small-magnitude smooth fields added on top of natural variability, but small perturbations to a complex system with many (known and unknown) nonlinearities and feedbacks. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text clarified to stress that point
9-1529	9	50	33	50	33	Alternatively, make it clear that sentence refers only to historical "human-induced climate change" on large scales, and not to possible future changes. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text clarified
9-1530	9	50	33			typo: "usu[jally]" [Andrew Wittenberg, United States of America]	Editorial. Taken into account
9-1531	9	50	34	50	34	Only "sometimes" too short? Surely it is always too short, unless you have good evidence or very optimistic assumptions about a complete lack of long timescale variability? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text clarified
9-1532	9	50	39	50	39	What time scale(s) does this subsection deal with? It is not clear. [Josephine Brown, Australia]	taken into account. Text clarified to better discuss the different time considered with these spectra
9-1533	9	50	41	50	42	The same issue of the importance of simulating climate variability. Perhaps it could be rewritten as "Accurately simulating climate variability on various time scales HAS CONSEQUENCES both for detection and attribution and for predictions and projections of future climate". [Ramon de Elia, Canada]	taken into account. Text revised
9-1534	9	50	47	50	47	... internal variability of annual mean surface temperature from.... (?) [European Union]	Editorial . Text revised
9-1535	9	50	47			insert "(top panel)" after Figure 9-33 [Barbara Früh, Germany]	Editorial . Text revised
9-1536	9	50	47			insert "(bottom panel)" after Figure 9-33 [Barbara Früh, Germany]	Editorial. Text revised
9-1537	9	50	48	50	54	Hartmann, B., and Wendler, G. 2005. The significance of the 1976 Pacific climate shift on the climatology of Alaska. Journal of Climate 18: 4824–4839. http://climate.gi.alaska.edu/researchprojects/hartmann%20and%20wendler%202005.pdf Wendler, G., and Shulski, M. 2009. A Century of Climate Change for Fairbanks, Alaska. Arctic 62 (3): 295-300. http://climate.gi.alaska.edu/papers/Arctic62-3-295.pdf [Richard Keen, United States of America]	taken into account, however not that given space limit it is not possible to include all references
9-1538	9	50	50	50	54	If control run drift is an issue, why not remove it from the historical run series and then do the analysis? [Government of United States of America]	Rejected. This would raise other issues. Simulations have not been detrended to compute the spectra (except the last millennium for MIROC, which is now listed in the figure caption)

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1539	9	50	50			The "power spectral density" for what fields? What kind of variability, spatial or temporal? [Government of United States of America]	Edditorial. Text clarifies
9-1540	9	50	51	50	52	...datasets. Sen Gupta et al. (2012) found... [Hai Lin, Canada]	Editorial. Taken into account
9-1541	9	50	52	50	52	It may be helpful to clarify that the "local drift" refers to spurious, unforced model drift. [Josephine Brown, Australia]	Taken into account
9-1542	9	50	52	50	53	Most of the time it is not stated whether or not "model drift" is removed prior to stating results. Could this be clarified in the introduction? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Editorial. Text clarifies
9-1543	9	50	56	50	56	Similar conclusions - similar to what? [Josephine Brown, Australia]	Editorial. Text modified
9-1544	9	50	57	50	57	Northern Hemisphere [Hai Lin, Canada]	Editorial. Considered
9-1545	9	51	3	51	4	The MPI-ESM model is of vintage 2012, but the Jungclaus et al. paper of 2010. This paper refers to older millennium simulations based on a predecessor of the MPI-ESM-P model, for which spectra are shown in Fig. 9.33 [European Union]	Taken into account. The paragraph has been rewritten
9-1546	9	51	4			link to CO2 cycle discussed earlier [Peter Braesicke, United Kingdom]	taken into account. Text revised
9-1547	9	51	5	51	5	Change "suggest" to "suggests" [Hai Lin, Canada]	typo. Corrected
9-1548	9	51	5	51	6	Please clarify what is meant by "low-frequency" here? [Government of United States of America]	Taken into account. Text has been clarified to highlight the low frequency part (> 50 years of the spectrum)
9-1549	9	51	7			"adequate representation"? -- please specify [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Take has been rewritten
9-1550	9	51	8			How serious is the underestimation discrepancy at long time scales? What are those time scales? [Government of United States of America]	Taken into account. However this last par has been removed because the reference was not available in time.
9-1551	9	51	10	51	21	1) Figure 9.33 (top) incorrectly presents the results of FGOALS-g2. We calculate the standard deviation of zonal-mean surface temperature of the CMIP5 pre-industrial control simulation by FGOALS-g2 (200-900ys), the results are quite different from those in the figure. The maximum standard deviation is about 0.7. Please check and correct the top figure. 2) The model name "FGOALS gl/s2" in the bottom figure is incorrect. It should be FGOALS-s2 or FGOALS-g2. Please check and revise it. [Bin Wang, China]	Taken into account. CMIP5 standard name are used on the revised version of the figure.
9-1552	9	51	11	51	14	put "top", "middle" and "bottom" at the end of the sentence instead of at the begin of the sentence. [Farahnaz Khosrawi, Sweden]	Editorial. Considered
9-1553	9	51	13	51	13	it's a bit unclear that the orange lines labeled "historical" refer to the CMIP5 simulations [Bart Van den Hurk, Netherlands]	Editorial. Figure revised
9-1554	9	51	23	57	16	subsections from 9.5.3.2 to 9.5.3.5: the topics are analyzed also in chap14. It should be worth explain here what is said more/differently. [Annalisa Cherchi, Italy]	Taken into account. Text modified in introduction
9-1555	9	51	27	51	27	Please explain, what is meant by "these modes". [Government of Germany]	Taken into account. Text clarified
9-1556	9	51	27	51	27	Specify which modes are meant here. Since the name of the subsection includes now 'Extra-tropical Circulation', it is better to write explicitly which modes are meant here. [Alexey Karpechko, Finland]	Taken into account. Text clarified
9-1557	9	51	29			Additional relevant references: Casado MJ, Pastor MA (2012) Use of variability modes to evaluate AR4 climate models over the Euro-Atlantic region. Clim Dyn 38(1–2):225–237. doi:10.1007/s00382-011-1077-2 [Government of Spain]	Taken into account. Reference cited
9-1558	9	51	40	51	42	It might be useful to define the word "trend" in a way that distinguishes it from long timescale natural variability. Then, is the change in NAO a "trend" or not? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text clarified
9-1559	9	51	42	51	43	I suppose it also can have consequences on detection studies. [Ramon de Elia, Canada]	Taken into account. Text modified to reflect comment

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1560	9	51	42	51	43	We recommend a change to the following: "Underestimation of NAO trends can contribute substantially to the underestimation of late 20th century observed warming trends over Asia in historical forcing runs (Knutson et al. 2006)." [Government of United States of America]	Taken into account. Text modified to reflect comment
9-1561	9	51	42	51	43	Depending on whether it is a "trend" to be extrapolated or "natural variability" to revert to the mean, then mis-characterisation may result in either over- or under-prediction of future change. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text modified
9-1562	9	51	50			The justification for reduced emphasis on NAO should be substantiated with a reference. [Government of United States of America]	Taken into account. Text modified to reflect comment
9-1563	9	51				Figure 9.33: The top panel could (and should) include an indication of observed values [William Merryfield, Canada]	Taken into account. Figure modified
9-1564	9	52	3	52	4	Please, add references to Sigmond and Scinocca (2010) and Karpechko and Manzini (2012) here. Both studies showed that the SLP response to doubling CO ₂ , and projecting on NAO/NAM, is sensitive to representation of the stratosphere in climate models, consistent with Scaife et al. (2011b). Missing references: (1) Sigmond, M., and J. F. Scinocca (2010), The influence of the basic state on the Northern Hemisphere circulation response to climate change, J. Clim., 23, 1434–1446, doi:10.1175/2009JCLI3167.1. (2) Karpechko A. Yu., and E. Manzini, Stratospheric influence on tropospheric climate change in the Northern Hemisphere (2012), J. Geophys. Res., 117, D05133, doi:10.1029/2011JD017036. [Alexey Karpechko, Finland]	Taken into account. References cited
9-1565	9	52	3	52	4	So what does this imply for confidence in regional (north Atlantic/western Europe/Greenland) projections? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted. This implies additional model uncertainty in regional projections, for example in the precipitation response over Europe. Given the pressure on space this cannot be added in the text.
9-1566	9	52	4	52	5	Actually, Morgenstern et al. (2010a) looked at the NAM response in 850-hPa geopotential heights, not in SLP. This makes large difference. At 850hPa and above in the troposphere, the geopotential height response is strongly influenced by the polar amplification, as is evident from Figure 4 by Woollings (2008). Figure 4 by Woollings (2008) shows that the NAM index response in many CMIP3 models changes sign from positive at 1000 hPa to negative at 850 hPa and above in the troposphere. Morgenstern et al. (2010a) do not show the response of NAM below 850hPa, therefore it is not clear whether it is any different from the response in the CMIP3 and CMIP5 models which simulate a positive response at SLP (and 1000hPa). Therefore I strongly suggest to remove the reference to negative NAM response in stratosphere-resolving models by Morgenstern et al. (2010a) because there is no evidence that the response is due to the stratosphere. On the contrary, it is very likely that Morgenstern et al. (2010a) only demonstrated the effect of polar amplification. I should credit Dr. Elisa Manzini who pointed me to this simple explanation of the negative NAM response in Morgenstern et al. (2010a). Reference: Woollings, T. (2008), Vertical structure of anthropogenic zonal-mean atmospheric circulation change, Geophys. Res. Lett., 35, L19702, doi:10.1029/2008GL034883. [Alexey Karpechko, Finland]	Noted. This is an interesting and useful comment .The reference is removed as suggested
9-1567	9	52	6	52	7	I don't think this statement is quite right, storm tracks and NAO are not independent [Peter Braesicke, United Kingdom]	Taken into account. Phrase removed
9-1568	9	52	24			maybe motivate why certain patterns are important for regional climate [Peter Braesicke, United Kingdom]	Taken into account. Text clarified
9-1569	9	52	31	52	32	Add reference to recent to Msadek et al. (2012) who compared the AMOC in two climate models (GFDL CM2.1 and NCAR CCSM4) to RAPID observations. This reference is also relevant for I.35-36 as Msadek et al. (2012) showed that the Ekman mass transport that defines the wind-driven part of the AMOC at 26N is too strong in the NCAR and GFDL models because of too strong easterlies. A reference to that study can also be added I. 39-40 as this paper evaluated the geostrophic variability associated with the AMOC at 26.5°N and showed that in the two CMIP5 models evaluated, the baroclinic geostrophic transport variability is underestimated, yielding a different than observed overturning meridional heat transport. Msadek R., W. E. Johns, S. G. Yeager, G. Danabasoglu, T. Delworth, and T. Rosati, 2012: The Atlantic	Taken into account. Reference cited

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						meridional heat transport at 26.5°N and its relationship with the MOC in the RAPID-array and GFDL and NCAR coupled models, Journal of Climate, accepted [RYM MSADEK, United States of America]	
9-1570	9	52	34	52	40	"appears well presented", "generally well represented", "might improve" -- avoid vague, non committing language [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Text clarified
9-1571	9	52	42	53	11	9.5.3.3.2 The discussion of the AMO is focused on CMIP3-era models. The paper titled "North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology" by Sheffield et al. contains a discussion of PDO fidelity in the CMIP5 models (section 6.2), which should be integrated here. [Government of United States of America]	Reject. This section is about the AMO and not the PDO, which is discussed elsewhere in the chapter.
9-1572	9	52	47	52	48	Please add: Recently AMOC fingerprints have been developed using observed subsurface temperature (Zhang 2007, 2008), and these observed AMOC fingerprints bring new evidence that the observed AMO is linked to AMOC variability. [Rong Zhang, United States of America]	Noted. Text modified to take comment into account
9-1573	9	52	47	52	48	Reference: Zhang, 2007: Anticorrelated multidecadal variations between surface and subsurface tropical North Atlantic, [Rong Zhang, United States of America]	Noted. Text modified to take comment into account
9-1574	9	52	47	52	48	Geophys. Res. Lett., 34, L12713, doi:10.1029/2007GL030225. [Rong Zhang, United States of America]	Noted. Text modified to take comment into account
9-1575	9	52	47	52	48	Zhang, 2008: Coherent surface-subsurface fingerprint of the Atlantic meridional overturning circulation, [Rong Zhang, United States of America]	Noted. Text modified to take comment into account
9-1576	9	52	47	52	48	Geophys. Res. Lett., 35, L20705, doi:10.1029/2008GL035463. [Rong Zhang, United States of America]	Noted. Text modified to take comment into account
9-1577	9	52	48	52	50	This paragraph is incomplete. It should be added than a number of models have an even shorter variability associated with the AMO and the AMOC, with time scales around 20-30 years (e.g. Dong and Sutton 2005, Msadek and Frankignoul 2009, Frankcombe et al. 2010, Escudier et al. 2012). Dong B, Sutton RT (2005) Mechanism of interdecadal thermohaline circulation variability in a coupled ocean–atmosphere gcm. J Clim 18:1117–1135 Msadek, R., and C. Frankignoul, 2009: Atlantic multidecadal oceanic variability and its influence on the atmosphere in a climate model. Climate Dyn., 33, 45–62. Frankcombe, L. M., A. von der Heydt, and H. A. Dijkstra (2010), North Atlantic multidecadal climate variability: An Investigation of dominant time scales and processes, J. Clim., 23, 3626–3638 Escudier R., Mignot J. and Swingedouw D., 2012: A 20-yr coupled ocean-sea ice-atmosphere variability mode in the North Atlantic in an AOGCM, doi: 10.1007/s00382-012-1402-4 [RYM MSADEK, United States of America]	Noted. This section describes model representations of the Atlantic Multidecadal Oscillation, "multidecadal" meaning the variability has phases that last for multiple decades, as is seen in observations. Variability at 20 year timescales does indeed occur in models - an "Atlantic Interdecadal Oscillation" perhaps - but this is not multidecadal in the sense it is understood from the observed AMO phenomenon. This is clarified in the revised text
9-1578	9	52	48			"despite this" - To what does the "this" refer? It is unclear. Furthermore if the this refers to the time scale (70 years), what is the uncertainty in the estimate of that time scale in the observations? The instrumental record is too short for much certainty...if any. [Ronald Stouffer, United States of America]	Taken into account. Text clarified
9-1579	9	52	50			The reference to Msadek and Frankignoul (2009) for timescales of a century or more is incorrect. The time scale associated with the AMOC and the AMO in the model analyzed by Msadek and Frankignoul (2009) has a much shorter variability, around 20yr. [RYM MSADEK, United States of America]	Noted. The simulation described in this paper may have interdecadal as well as longer-period variability, but the paper shows that the largest contribution to the variance comes from the approximately 100 year or longer variations (see fig. 2c of the paper and its discussion in the text).
9-1580	9	52	52	52	52	Is the term "convergence" used in this sense elsewhere? If not, I would avoid it and go stright to the explanation. [Ramon de Elia, Canada]	Reject. This usage is in the normal mathematical sense. This is the definition from a common online dictionary: "Convergence: Mathematics The property or manner of approaching a limit, such as a point, line,

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							function, or value".
9-1581	9	52	53	52	53	"Recent modelling", which one? [Annalisa Cherchi, Italy]	Taken into account. Text clarified
9-1582	9	52	54			"Internal damped oceanic mode" should also be added to the mechanisms list 1.54-56 [RYM MSADEK, United States of America]	Taken into account. Text modified to reflect comment
9-1583	9	53	3			I suggest replacing "not a result" by "not only a result" [RYM MSADEK, United States of America]	Taken into account. The text has been modified to clarify to respective sizes of the forced and internal components found in these studies.
9-1584	9	53	4	53	6	How do we know Booth et al provide a better reproduction of the historical AMO fluctuations. A criticism of the Booth et al. study (Zhang et al. 2012, submitted) suggests that for many important AMO-related metrics, the Booth et al. simulation is not very realistic, casting doubt on the Booth et al. claim that radiative forcing explains most of the AMO variability over the 20th century. [Government of United States of America]	Taken into account. Text modified and reference cited
9-1585	9	53	4	53	7	Please add: Recent study (Zhang et al. 2012) shows that the simulations used in Booth et al. 2012 have important [Rong Zhang, United States of America]	Taken into account. Text modified and reference cited
9-1586	9	53	4	53	7	discrepancies with many observed changes in the North Atlantic. The discrepancies are strongly influenced by [Rong Zhang, United States of America]	Taken into account. Text modified and reference cited
9-1587	9	53	4	53	7	aerosols, and cast considerable doubt on the claim that aerosols drive the bulk of the AMO. [Rong Zhang, United States of America]	Taken into account. Text modified and reference cited
9-1588	9	53	4	53	7	Reference: Zhang et al. 2012, Have Aerosols Caused the Observed Atlantic Multidecadal Variability?, Journal of A [Rong Zhang, United States of America]	Taken into account. Text modified and reference cited
9-1589	9	53	4	53	7	Atmospheric Sciences, Submitted. [Rong Zhang, United States of America]	Taken into account. Text modified and reference cited
9-1590	9	53	4			Add a reference to Terray (2012) who investigated the drivers of the AMO in the CMIP5 models: Terray, L., 2012: Evidence for multiple drivers of North Atlantic multi-decadal climate variability. Geophys. Res. Lett., 39, L19712. [RYM MSADEK, United States of America]	Taken into account. Reference cited
9-1591	9	53	5	53	7	The fact that a better representation of the AMO could result from a more realistic simulation of aerosols is controversial. The influence of aerosols relative to other processes on decadal variability in the North Atlantic remains uncertain and considerable doubt has been cast on the main conclusion of Booth et al. (2012) because of important discrepancies between the model they used and observations (Zhang et al. 2012). This paragraph should account for the lack of consensus when discussing these results: Zhang R., T. Delworth, R. Sutton, D. L. R. Hodson, K. W. Dixon, I. M. Held, Y. Kushnir, J. Marshall, Y. Ming, R. Msadek, J. Robson, A. Rosati, M. Ting, G. Vecchi, 2012: Have aerosols caused the observed multidecadal variability?,submitted [RYM MSADEK, United States of America]	Taken into account. Text modified and reference cited
9-1592	9	53	6			The fact that aerosols may play a role does not necessarily mean the AMO is forced. The aerosol fluctuations could be natural, linked to atmospheric circulation changes that in turn are linked to the AMO. But the last sentence of the paragraph does refer to volcanic aerosols, so does suggest forcing. Perhaps all that is needed is to move the statement that "this suggests that at least part of the AMO may in fact be forced" to the end of the paragraph. [Adrian Simmons, United Kingdom]	Taken into account. Text modified to reflect comment
9-1593	9	53	7			Why use "the fact" here when you are talking about something that "may also be a strong driver"? It is a fact that something may be true? We recommend that the authors strike "the fact" in order to clarify the statement. [Government of United States of America]	Taken into account. Text clarified
9-1594	9	53	41			But what caused the mean biases? Perhaps both mean bias and variability problems are ultimately related to model resolution for example. [Government of United States of America]	Noted. Certain model biases, such as the severe warm SST bias off the west coast of southern Africa,

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							are related to oceanic meso- and frontal-scale features. In the case of the warm SST bias, recent studies show that it is likely related to inability of global climate models to resolve the Angola-Benguela front in the southeast Atlantic. Enhancing model resolution, in this case, will help to improve the SST bias problem.
9-1595	9	53	45	55	25	Section 9.5.3.4: ENSO and Indian Ocean dipole mode were discussed separately. However, recent studies have shown intrinsic interactions between the tropical Pacific and Indian Ocean. This inter-basin interactions affect both ENSO and IOD and their predictability (Wu and Kirtman 2004; Behera et al. 2005; Luo et al. 2010; Izumo et al. 2010). References: 1) Wu, R., and B. P. Kirtman, 2004: Understanding the impacts of the Indian Ocean on ENSO variability in a coupled GCM. <i>J. Climate</i> , 17, 4019-4031. 2) Behera, K. S., J.-J. Luo, S. Masson, S. A. Rao, H. Sakuma, and T. Yamagata, 2006: A CGCM study on the interaction between IOD and ENSO. <i>J. Climate</i> , 19, 1688-1705. 3) Luo, J.-J., R. Zhang, S. Behera, Y. Masumoto, F.-F. Jin, R. Lukas, and T. Yamagata, 2010: Interaction between El Nino and extreme Indian Ocean Dipole. <i>J. Climate</i> , 23(3), 726-742. 4) Izumo, T., J. Vialard, M. Lengaigne, C. Boyer Montegut, S. K. Behera, J.-J. Luo, S. Cravatte, S. Masson, and T. Yamagata, 2010: Influence of the state of the Indian Ocean Dipole on the following year's El Niño. <i>Nature Geoscience</i> , 3, 168-172. [Government of Australia]	Taken into account. Establishing boundaries in the tropics is indeed very difficult. Text modified to stress the strong link between the IOD and ENSO
9-1596	9	53	45	55	25	Section 9.5.3.4: ENSO and Indian Ocean dipole mode were discussed separately in this section. However, recent studies have shown intrinsic interactions between the tropical Pacific and Indian Ocean. This inter-basin interactions affect both ENSO and IOD and their predictability (Wu and Kirtman 2004; Behera et al. 2005; Luo et al. 2010; Izumo et al. 2010, among others). The inter-basin interactions may also play a considerable role in multi-decadal and centennial climate changes in the tropics (Luo et al. 2012). [Jing-Jia Luo, Australia]	Taken into account. Establishing boundaries in the tropics is indeed very difficult. Text modified to stress the strong link between the IOD and ENSO
9-1597	9	53	45	55	25	References: 1) Wu, R., and B. P. Kirtman, 2004: Understanding the impacts of the Indian Ocean on ENSO variability in a coupled GCM. <i>J. Climate</i> , 17, 4019-4031. 2) Behera, K. S., J.-J. Luo, S. Masson, S. A. Rao, H. Sakuma, and T. Yamagata, 2006: A CGCM study on the interaction between IOD and ENSO. <i>J. Climate</i> , 19, 1688-1705. 3) Luo, J.-J., R. Zhang, S. Behera, Y. Masumoto, F.-F. Jin, R. Lukas, and T. Yamagata, 2010: Interaction between El Nino and extreme Indian Ocean Dipole. <i>J. Climate</i> , 23, 726-742. 4) Izumo, T., et al., 2010: Influence of the state of the Indian Ocean Dipole on the following year's El Niño. <i>Nature Geoscience</i> , 3, 168-172. 5) Luo, J.-J., W. Sasaki, and Y. Masumoto, 2012: Indian Ocean warming modulates Pacific climate change. <i>PNAS</i> , www.pnas.org/cgi/doi/10.1073/pnas.1210239109 . [Jing-Jia Luo, Australia]	Noted. This is an assessment and not a review. Due to space constraint, we are limited by the number of studies we can cite.
9-1598	9	53	48	53	49	The ENSO is also the dominant driver of global average temperature because 50% of the Earth's surface lies between 30N and 30S. See also 4AR chapter 3 Executive Summary which said "The dominant mode of global-scale variability on interannual time scales is ENSO...", and 4AR 3.6.2.1 "The [ENSO] has global impacts..." and "[ENSO] links to global mean temperature perturbations are given by Trenberth et al. (2006)." [John McLean, Australia]	Taken into account. Text modified in introduction
9-1599	9	53	48	53	58	Please add "While the characteristics of the ENSO are relatively well recognised, the level of scientific understanding of its drivers remains low and this limits forecasting." [John McLean, Australia]	Taken into account. Text modified to reflect comment
9-1600	9	53	55	53	55	Add "in" after "traced" [Hai Lin, Canada]	Editorial. Taken into account
9-1601	9	53	55			Newer references beyond Watanabe et al. 2011: Watanabe and Wittenberg (GRL, 2012), and Watanabe et al. (GRL, 2012). [Andrew Wittenberg, United States of America]	Taken into account. Reference cited
9-1602	9	53	55			"traced [to] the" [Andrew Wittenberg, United States of America]	Editorial. Taken into account
9-1603	9	53	56			Include additional references and ocean processes: "feedbacks (... , [Kim et al., <i>Climate Dyn.</i> 2008], as well as the representation of mixing in the upper ocean (Anderson et al., <i>Ocean Sci.</i> 2009; Dunne et al., <i>J. Climate</i> 2012a).]" [Andrew Wittenberg, United States of America]	Taken into account. Reference cited
9-1604	9	53	57	53	58	Models perform particularly well - for mean state or ENSO? [Josephine Brown, Australia]	Taken into account. Text clarified

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1605	9	53	57			We recommend that the authors append the sentence with: "although ultimate causation is difficult to establish." [Government of United States of America]	Taken into account. Text modified to reflect comment
9-1606	9	54	4	54	4	Replace biannual (2 times per year) with biennial (every 2 years) [Josephine Brown, Australia]	Taken into account. Text modified
9-1607	9	54	4	54	4	... biennial... (biannual is twice per year) [European Union]	Taken into account. Text modified
9-1608	9	54	14	54	14	Guilyardi (2009) and Leloup (2008) are CMIP3 studies not CMIP5 studies. [Josephine Brown, Australia]	Taken into account. Text updated with CMIP5 references
9-1609	9	54	15	54	15	Use El Nino and La Nina rather than warm and cold phase ENSO for consistency. [Josephine Brown, Australia]	Taken into account. Text modified to reflect comment
9-1610	9	54	27	54	27	This figure could be omitted, it is only shortly mentioned in the text. Thus, the reference where the figure was published before would be enough. [Farahnaz Khosrawi, Sweden]	Rejected. This is a requirement from IPCC AR5 scoping meeting.
9-1611	9	54	28	54	33	In Fig 9.35 bring observed lines 'to front', as they currently cannot be seen under mass of model lines. [Government of Australia]	Editorial. Taken into account
9-1612	9	54	29	54	30	The upper panel of Figure 9.35 lists "ERA40", while the caption explains that "ECMWF" mean ERA-15. ? [European Union]	Taken into account. Text clarified
9-1613	9	54	35	54	35	Same here. The figure could be omitted. The given reference is sufficient. [Farahnaz Khosrawi, Sweden]	Rejected. This figure provide a quantitative estimate in support of the text.
9-1614	9	54	35	54	41	In Figure 9.36, what does IAP (number 7) mean? It seems the FGOALS-g1.0, one of CMIP3/AR4 models. Among CMIP5 models, IAP has two versions: FGOALS-g2 and FGOALS-s2. Please revise the figure with a correct model name for number 7. Please also include the results of both FGOALS-g2 and FGOALS-s2 (CMIP5 models) like the way of presenting the results of GISS models. [Bin Wang, China]	Taken into account. Figure updated with latest CMIP5 results
9-1615	9	54	36	54	42	The Figure misses panels (e) and (f), and (c) and (d) are swapped. [Andreas Sterl, Netherlands]	Editorial. Taken into account
9-1616	9	54	42			model standard deviation -> inter model standard deviation? [Andreas Sterl, Netherlands]	Taken into account. Text clarified
9-1617	9	54	48	54	48	Pls change "(Yu and Kim, 2011)" to "Yu and Kim, (2011)" [HASIBUR RAHAMAN, India]	Editorial. Taken into account
9-1618	9	54	52	54	54	You cannot sensibly state "short observational record" without providing the reader with the context. Is it since 1937 or since 1870 (to give two limiting dates for ENSO data) or is it some other period? Proxy records of ENSO conditions go back at least 125,000 years (see Tudhope). [John McLean, Australia]	Taken into account. Text clarified
9-1619	9	54	52			"...evaluation of ENSO performance is hampered by the short observational record": We recommend that the following recent paper by Su and Jiang (2012) should be cited: Su, H., and J.H. Jiang, "Tropical Clouds and Circulation Changes During the 2006-07 and 2009-10 El Niños," J. Climate, doi:10.1175/JCLI-D-12-00152.1, in press, 2012. [Government of United States of America]	Taken into account. Reference cited
9-1620	9	54	54	54	56	Van Oldenborgh (2005), Guilyardi (2009) and Leloup (2008) are CMIP3 studies not CMIP5 studies. [Josephine Brown, Australia]	Taken into account. Text updated with CMIP5 references
9-1621	9	54	56			The result of reduced number of poor-performing models on the model ensemble in CMIP5 is demonstrated for Tropical Pacific in Grose, M. R., J. N. Brown, et al. (2012). "Assessment of the CMIP5 global climate model simulations of the western tropical Pacific climate system and comparison to CMIP3." submitted to International Journal of Climatology. [Jaclyn Brown, Australia]	Taken into account. Reference cited
9-1622	9	54				Figure 9.36: the labels of the horizontal axis overlap. [Government of Germany]	Editorial. Taken into account
9-1623	9	54				Figure 9.35: Several issues here: 1) There is no reference to the ECMWF analysis in the upper panel, just ERA40 in contradiction to the caption; 2) The caption should indicate what scenario and years (or length of time series if preindustrial, to indicate sample size) are used for the models; 3) The data are a mixed bag of SST and surface air temperature--does that matter? [William Merryfield, Canada]	Taken into account. Caption clarified
9-1624	9	54				Fig. 9.36: This figure has evidently been updated from one that appeared in an article by Guilyardi et al. in	Taken into account. Caption clarified

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Clivar Exchanges (v. 58, p. 29, 2012). In particular, 8 columns have been added but the legend has not changed, so that not only are some models not labeled, but most are misidentified. In addition the letters have not changed position even though the accompanying data points have. [William Merryfield, Canada]	
9-1625	9	55	19	55	25	move paragraph up to other ENSO discussion [Peter Braesicke, United Kingdom]	Taken into account. Text re-organised
9-1626	9	55	26			<p>In cross-chapter meetings in Marrakech, it was agreed that ch 14 would describe possible future changes to the tropospheric biennial oscillation (TBO), and therefore there must be a comparable description in ch. 9. This material was supplied as agreed but not included. I include it here again: "A dynamical framework for the IOD as well as ENSO is the tropospheric biennial oscillation (TBO, Ch. 14, section 14.6.4). The TBO is a biennial tendency of many phenomena in the Indo-Pacific region that affects droughts and floods over large areas of south Asia and Australia (e.g. Meehl, 1997; Chang and Li, 2000; Li et al., 2001; Meehl et al., 2003). The IOD involves regional patterns of SST anomalies in the TBO in the Indian Ocean during the northern fall season following the south Asian monsoon (Loschnigg et al., 2003). The transition processes in the TBO have been used to produce El Nino forecast skill by monitoring the state of the IOD in northern fall (Izumo et al., 2010). The TBO has been successfully simulated in a number of global coupled climate model simulations (e.g. Meehl, 1997; Ogasawara et al., 1999; Loschnigg et al., 2003; Nanjundiah et al. 2005; Turner et al., 2007; Meehl and Arblaster, 2011)."</p> <p>References Chang, C.P., and T. Li, 2000: A theory for the tropical tropospheric biennial oscillation. <i>J. Atmos. Sci.</i>, 57, 2209–2224. Izumo, T., J. Vialard, M. Lengaigne, C. de Boyer Montegut, S.K. Behera, J.-J. Luo, S. Cravatte, S. Masson, and T. Yamagata, 2010: Influence of the state of the Indian Ocean Dipole on the following year's El Nino. <i>Nature Geoscience</i>, 3, 168—172. Li, T., C.W. Tham and C.P. Chang, 2001: A coupled air-sea-monsoon oscillator for the tropospheric biennial oscillation. <i>J. Climate</i>, 14, 752-764. Loschnigg, J., G. A. Meehl, P. J. Webster, J. M. Arblaster, and G. P. Compo, 2003: The Asian monsoon, the tropospheric biennial oscillation and the Indian Ocean Dipole in the NCAR CSM. <i>J. Climate</i>, 16, 2138–2158. Meehl, G. A., 1997: The south Asian monsoon and the tropospheric biennial oscillation. <i>J. Climate</i>, 10, 1921–1943. Meehl, G. A., J. M. Arblaster, and J. Loschnigg, 2003: Coupled ocean-atmosphere dynamical processes in the tropical Indian and Pacific Ocean regions and the TBO. <i>J. Climate</i>, 16, 2138–2158. Meehl, G. A., and J. M. Arblaster, 2011: Decadal variability of Asian-Australian monsoon-ENSO-TBO relationships. <i>J. Climate</i>, 24, 4925-4940. Nanjundiah, R.S., V. Vidyumala, and J. Srinivasan, 2005: The impact of increase in CO2 on the simulation of tropical biennial oscillations (TBO) in 12 coupled general circulation models. <i>Atm. Sci. Lett.</i>, 6, 183—191. Ogasawara, N., A. Kitoh, T. Yasunari, and A. Noda, 1999: Tropospheric Biennial Oscillation of ENSO-monsoon system in the MRI coupled GCM. <i>J. Meteorol. Soc. Japan</i>, 77, 1247—1270. Turner, A.G., P.M. Inness, and J.M. Slingo, 2007: The effect of doubled CO2 and model basic state biases on the monsoon-ENSO system. II: Changing ENSO regimes. <i>Q.J.Roy. Meteorol. Soc.</i>, 133, 1159—1173. [Gerald Meehl, United States of America]</p>	Taken into account. Edited text added
9-1627	9	55	27	55	41	How can models simulate PDO with limited skill but simulate IPO reasonably well? These are effectively the same or closely related modes of variability. [Josephine Brown, Australia]	Taken into account. Text clarified
9-1628	9	55	27	55	41	9.5.3.4.3 The discussion of the PDO is focused on CMIP3-era models. The paper titled "North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology" by Sheffield et al. contains a discussion of PDO fidelity in the CMIP5 models (section 6.1), which we recommend be considered here. [Government of United States of America]	Taken into account. Reference cited (Part II is the relevant study here)
9-1629	9	55	27	55	41	What is the relationship between the PDO and the IPO? Are they just two different views of the same phenomenon? [Government of United States of America]	Yes, the PDO and IPO are basically two different views of the same phenomenon, but they are defined differently. The PDO is defined from North Pacific SSTs (as stated in the first sentence of this section), whereas the IPO is defined from SSTs over the entire Pacific basin (as

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							stated in the parentheses in the last sentence of this Section).
9-1630	9	55	27	55	41	Annex I, Figure AI.16 The historical ensemble models in the figure show no evidence of the actual winter (DJF) PDO signal in Alaska and western North America over the past century, and show no climate variation of the proper amplitude (2-4C) and time scale (~60 year periodicity) for the region. The PDO is the dominant contributor to climate variations of Alaska over the past century, and the models fail to capture this process. This should be noted so the reader can assess the validity of regional reconstructions and projections. Refs. below. [Richard Keen, United States of America]	Taken into account. Text modified to reflect comment
9-1631	9	55	40			the section number in ch. 14 has changed--this should now be section 14.6.6 [Gerald Meehl, United States of America]	Taken into account. Text modified
9-1632	9	55	43	56	8	Section 9.5.3.4.4: Ocean decadal variability in the tropical Pacific can be caused by various processes; STC variation is only one of the possible reasons, which may mostly provide a positive feedback to tropical SST changes. A possible mechanism for the tropical ocean decadal variability is the subsurface ocean signals originated from extra tropical oceans, particularly the South Pacific, which provide a long memory for the tropical decadal variability (Gu and Philander 1997; Luo and Yamagata 2001; Luo 2010). References: 1) Gu, D., and S. G. H. Philander, 1997: Interdecadal climate fluctuations that depend on exchanges between the tropics and extra tropics. Science, 275, 805-807. 2) Luo, J.-J., and T. Yamagata, 2001: Long-term El Niño-Southern Oscillation (ENSO)-like variation with special emphasis on the South Pacific. J. Geophys. Res., 106, 22211-22227. 3) Luo, J.-J., 2011: Ocean dynamics not required? Nature, 477, 544-546. [Government of Australia]	Section removed
9-1633	9	55	43	56	8	Section 9.5.3.4.4: Ocean decadal variability in the tropical Pacific can be caused by various processes; STC variation is only one of the possible reasons, which may mostly provide a positive feedback to tropical SST changes. A possible mechanism for the tropical ocean decadal variability is the subsurface ocean signals originated from extratropical oceans, particularly the South Pacific, which provide a long memory for the tropical decadal variability (Gu and Philander 1997; Luo and Yamagata 2001; Luo 2010; among others). References: 1) Gu, D., and S. G. H. Philander, 1997: Interdecadal climate fluctuations that depend on exchanges between the tropics and extratropics. Science, 275, 805-807. 2) Luo, J.-J., and T. Yamagata, 2001: Long-term El Niño-Southern Oscillation (ENSO)-like variation with special emphasis on the South Pacific. J. Geophys. Res., 106, 22211-22227. 3) Luo, J.-J., 2011: Ocean dynamics not required? Nature, 477, 544-546. [Jing-Jia Luo, Australia]	Section removed as material not used by subsequent chapters
9-1634	9	55	45	55	56	Are there any results on how the CMIP5 models perform in simulating STCs compared to the CMIP3 models? [Government of Germany]	Section removed as material not used by subsequent chapters
9-1635	9	55	52	55	53	The claimed weakening of the Walker Circulation has a simple explanation. Since 1976 ENSO conditions on the El Nino side of absolutely neutral (ie. SOI = 0) have dominated, and the increase in Hadley Circulation and decrease in Walker Circulation is in accordance with numerous papers that describe ENSO circulation (refer Oort and Yienger (1996), Bhaskaran and Mullan (2003, Trenberth (2002).) [John McLean, Australia]	Section removed as material not used by subsequent chapters
9-1636	9	55	56			the following material was submitted as a consequence of cross-chapter meetings in Marrakech and was not incorporated, I submit it here to be inserted again: "However, one modeling study (Meehl and Hu, 2006) demonstrated that the STCs played a crucial role for simulating decadal timescale variability associated with the IPO (section 9.5.3.4.3), while another study with a different model showed that decades with little global warming, such as the recent hiatus decade, are likely related to a spin-up of the STCs in the Pacific producing cooler SSTs in the tropical Pacific (Meehl et al., 2011). Therefore, some aspects of the STCs can be simulated by present-day AOGCMs. References: Meehl, G. A., and A. Hu, 2006: Megadroughts in the Indian monsoon region and southwest North America and a mechanism for associated multi-decadal Pacific sea surface temperature anomalies. Journal of Climate, 19, 1605–1623.	Section removed as material not used by subsequent chapters

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Meehl, G.A., J.M. Arblaster, J. Fasullo, A. Hu, and K.E. Trenberth, 2011: Model-based evidence of deep ocean heat uptake during surface temperature hiatus periods. <i>Nature Climate Change</i> , 1, 360–364, doi:10.1038/NCLIMATE1229. [Gerald Meehl, United States of America]	
9-1637	9	56	1			Section 9.5.3.5: In my review of FOD I observed that 'the section 9.5.3.5 'Teleconnections' misses an important teleconnection: ENSO-NAO.' The section 9.5.3.5 still misses this teleconnection. Since there is no feedback from the author team I do not know why my comment was ignored. Therefore I repeat it here. The ENSO/NAO teleconnetcion has important consequences for European winter climate. A comprehensive review of this phenomenon, including observational evidences, is provided by Brönnimann (2007). There is a number of model studies showing importance of a good stratospheric representation for simulation of the ENSO-NAO teleconnection. Relevant publications: (1) Brönnimann, S., 2007: Impact of El Nino–Southern Oscillation on European climate. <i>Rev. Geophys.</i> , 45, RG3003, doi:10.1029/2006RG000199. (2) Bell, C. J., L. J. Gray, A. J. Charlton-Perez, M. M. Joshi, A. A. Scaife, 2009: Stratospheric Communication of El Niño Teleconnections to European Winter. <i>J. Climate</i> , 22, 4083–4096. (3) Cagnazzo, C., and E. Manzini, 2009: Impact of the stratosphere on the winter tropospheric teleconnections between ENSO and the North Atlantic and European Region. <i>J. Climate</i> , 22, 1223–1238. (4) Ineson, S., and A. A. Scaife, 2009: The role of the stratosphere in the European climate response to El Nino. <i>Nature Geosci.</i> , 2, 32–36, doi:10.1038/ngeo381. [Alexey Karpechko, Finland]	Taken into account. Text re-organised
9-1638	9	56	3	56	8	An important methodological point to make in the introduction sect.9.5.3.5 may be that systematic analyses of the ability of GCMs to replicate teleconnections have been very limited to date. I'm aware of Yang and DelSole (2012) (for ENSO) and Rowell (2013) (for Africa), though there may well be others for other regions. Yang, X. and DelSole, T., 2012: Systematic Comparison of ENSO Teleconnection Patterns between Models and Observations. <i>J. Climate</i> , 25, 425–446. Rowell, D.P., 2013: Simulating Large-Scale Teleconnections to Africa: What is the State of the Art? <i>J. Climate</i> , submitted [Dave Rowell, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text modified to reflect comment
9-1639	9	56	10			Section 9.5.3.5: (a) I suggest that tropical teleconnections should be the first sub-section since this is where the strongest teleconnections are found. (b) I also think that the section 'Tropical ENSO teleconnections' needs to be extended to include all tropical teleconnections, not only ENSO; there are very important teleconnections from the Atlantic and Indian Oceans, and Mediterranean, that affect Africa and South America which cannot really be ignored. (c) The section 'Teleconnections affecting North America' needs to be extended to include all extra-tropical teleconnections, ie. Europe, Asia, and perhaps southern South America, so that readers do not perceive a US-centric review. [Dave Rowell, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text re-organised
9-1640	9	56	13			"in winter time (Stonger et al., 2009)[, and the PNA pattern has improved in some models in tandem with improved spatial resolution (Delworth et al., <i>J. Climate</i> 2012).]" [Andrew Wittenberg, United States of America]	Taken into account. Text modified
9-1641	9	56	25	46	26	The statement "In contrast to CMIP3, the CMIP5 ensemble is driven by a single internally consistent set of SO4 concentrations and SO2 emissions." is not correct from my understanding. 1) Yes, the models were provided consistent emissions, but not consistent SO4 concentrations. SO4 loading differ between models. 2) Also, it might be useful to mention that historical emissions were not implemented the same in all models (I had thought there was some evidence of this, but I could be mistaken)-- so some inconstancies may still remain. [Steven Smith, United States of America]	Accepted -- the statement that the SO4 concentrations are internally consistent has been removed.
9-1642	9	56	26			Section 9.5.3.5.2: The teleconnection between ENSO and the South Asian summer monsoon rainfall should be mentioned here. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Rejected. No specific references where either provided or found on this topic. This is an assessment, not a review.
9-1643	9	56	27	56	30	This sentence is a little unclear. I would suggest to simplify it, for example as: The tropical ENSO teleconnections, in which a baroclinic tropospheric warming signal propagates within convection zones, involve mechanisms related to the precipitation response to global warming (Chiang and Sobel, 2002; Neelin et al., 2003), and provide challenging test statistics for model precipitation response. [Koichi Sakaguchi, the United States]	Taken into account. Text modified to reflect comment

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1644	9	56	31	56	31	Replace ")"(" with "," [Hai Lin, Canada]	Editorial. Taken into account
9-1645	9	56	31	56	31	Pls change "Langenbrunner and Neelin, 2012) (Coelho and Goddard, 2009" to "Langenbrunner and Neelin, 2012;Coelho and Goddard, 2009" [HASIBUR RAHAMAN, India]	Editorial. Taken into account
9-1646	9	56	31	56	32	change "(Cai et al., 2009; Langenbrunner and Neelin, 2012) (Coelho and Goddard, 2009; Neelin,2007)" to "(Cai et al., 2009; Langenbrunner and Neelin, 2012; Coelho and Goddard, 2009; Neelin, 2007)" [SAMIR POKHREL, INDIA]	Editorial. Taken into account
9-1647	9	56	34			In the western Pacific, the sign of the precipitation associated with ENSO can be wrong depending on the location of the cold tongue bias in each particular model (ref Brown et al. Climatic Change) Brown, J. N., A. Sen Gupta, et al. (2012). "Implications of CMIP3 model biases and uncertainties for climate projections in the western Tropical Pacific." Climatic Change. DOI 10.1007/s10584-012-0603-5 [Jaclyn Brown, Australia]	Taken into account. Text modified and reference cited
9-1648	9	56	37	56	37	Repeated citation. [Dave Rowell, United Kingdom of Great Britain & Northern Ireland]	Editorial. Taken into account
9-1649	9	56	38	56	50	Sentence "Teleconnection patterns from both ENSO and the Indian Ocean Dipole..." should be moved to paragraph beginning Line 50 on Australian teleconnections [Josephine Brown, Australia]	Taken into account. Text clarified
9-1650	9	56	40			"multi-model [mean]" [Andrew Wittenberg, United States of America]	Editorial. Taken into account
9-1651	9	56	42	56	48	Other, equally strong, SST influences on the WAM that should be mentioned are the Indian Ocean (eg. Palmer 1986, Rowell 2001) and Mediterranean (eg. Rowell 2003, Jung et al. 2006). Also, Rowell (2013) has carried out a systematic assessment of the ability of CMIP3 and CMIP5 models to represent SST teleconnections to 6 large-scale regions across Africa, including information that could be added to Fig.9.45, and the result that no improvement is seen from CMIP3 to CMIP5. Palmer, T. N., 1986: Influence of Atlantic, Pacific and Indian Oceans on Sahel rainfall. Nature, 322, 251–253. Jung, T; Ferranti L; Tompkins AM, 2006: Response to the summer of 2003 Mediterranean SST anomalies over Europe and Africa. J. Climate, 19, 5439-5454. Rowell, D.P., 2001: Teleconnections between the Tropical Pacific and the Sahel. Q. J. R. Meteorol. Soc., 127, 1683-1706. Rowell, D.P., 2003: The Impact of Mediterranean SSTs on the Sahelian Rainfall Season. J. Climate, 16, 849-862. Rowell, D.P., 2013: Simulating Large-Scale Teleconnections to Africa: What is the State of the Art? J. Climate, submitted [Dave Rowell, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Text modified and reference cited
9-1652	9	56	51			What is "the Nino 3.4"? [Government of United States of America]	Taken into account. Text clarified
9-1653	9	57	1	57	16	What does this mean and do we believe parameterised waves under climate change? (see 9-28 / 4 as well) [Peter Braesicke, United Kingdom]	Taken into account. Text clarified
9-1654	9	57	1	57	16	QBO should be briefly described as is done for the other phenomena. [Government of United States of America]	Taken into account. Text clarified
9-1655	9	57	4	57	4	Are you sure MPI-ESM-LR reproduces the QBO? Is it not the MPI-ESM-MR? You could also include here the CMCC-CMS model. [Chiara Cagnazzo, Italy]	Taken into account. Text clarified
9-1656	9	57	4	57	4	... MPI-ESM-MR... (MPI-ESM-LR does not simulate the QBO as it has a coarser vertical grid) [European Union]	Taken into account. Text clarified
9-1657	9	57	4			This should be MPI-ESM-MR, not LR. The latter has no QBO. [Thorsten Mauritsen, Germany]	Taken into account. Text clarified
9-1658	9	57	8	57	8	"Krismser et al., 2012" would fit here; Krismser, T. R., M. A. Giorgetta, and M. Esch, 2012, Seasonal Aspects of the Quasi-Biennial Oscillation in MPI-ESM and ERA-40, J. Geophys. Res., submitted. [European Union]	Taken into account. Reference cited
9-1659	9	57	18			How does this link to CMIP5 and is it relevant for projections? [Peter Braesicke, United Kingdom]	Noted, but this text has been deleted. Incidentally, the link to projections is explained, for the case of tropical land carbon feedbacks, in section 9.8.3.
9-1660	9	57	20	57	27	Are the "coupled" models driven by observed SSTs? If not, how then can the comparison to observed carbon flux variability be made? [Bart Van den Hurk, Netherlands]	Noted, but this text has been deleted. Incidentally, the first paragraph refers to standalone land models

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							driven by observed climate, not coupled AOGCMs.
9-1661	9	57	29	57	30	"When these components are linked to fully coupled Earth system models, these models tend to overestimate the long-term trend in global-mean atmospheric CO2 concentrations (Cadule et al., 2010)" I think what is meant by "long-term trend" is not completely clear here. Also this passage overstates the results of Cadule et al who examined only 3 models, two of which overestimated the observed trend while one underestimated it. [James Christian, Canada]	Noted, but this text has been deleted.
9-1662	9	57	32	57	32	skip or update Cox et al., Submitted [European Union]	Noted, but this text has been deleted. Incidentally, this paper has now been published in Nature.
9-1663	9	57	37	57	41	Chapter 9: Section 9.5.3 (Simulation of Variability and Extremes. Interannual to Centennial Variability. Summary). The paragraph summarizing the whole section 9.5.3 should emphasize that most of variability modes presented have not been yet assessed in CMI [Government of Spain]	Taken into account. Text modified
9-1664	9	57	37	57	41	Point out for which processes you derive high confidence and for which low confidence (exemplarily) [Farahnaz Khosrawi, Sweden]	Taken into account. Text clarified
9-1665	9	57	38	57	38	From the table it looks like that most CMIP3/CMIP5 comparisons weren't done. How can you then judge that there is a modest improvement between CMIP3 and CMIP5? [Farahnaz Khosrawi, Sweden]	Taken into account. Text clarified
9-1666	9	57	44	57	44	Column headings "EG" and "Panel" are not clear. [Josephine Brown, Australia]	Taken into account. Table updated
9-1667	9	57	44	57	44	Since the Panel is always "B", please consider to delete this column. [Government of Germany]	Taken into account. Table updated
9-1668	9	57	44	57	44	It is ludicrous to assign a high level of confidence to ENSO models because the ultimate driver of ENSO has not been identified. Without the knowledge of this driver and being able to predict its likely future states, the predictability of ENSO is very limited. Predictive models for ENSO are rarely accurate 12 months ahead and, based on monthly reports of model prediction for ENSO this year (2012), many are of questionable accuracy after 3 or 4 months. [John McLean, Australia]	Taken into account. There are no "ENSO models". The high level of confidence is in the assessment itself an dteh agreement among studies that the models quality is medium. Reviewer is referred to refer to IPCC uncertainty language for guidance.
9-1669	9	57	44			Table 9.3, Column: Difference with AR4 (including CMIP5 vs. CMIP3), "No assessment" is the main conclusion? [Government of Spain]	Taken into account. Table updated with latest CMIP5 studies
9-1670	9	57	44			tbl. 9.3 What does "Panel" refer to in the Table 9.3? [Government of United States of America]	Taken into account. Table updated
9-1671	9	57	44			table 9.3 should include an entry for the TBO to be consistent with ch. 14, box 14.1, table 1; such an entry could have the following elements: Tropospheric biennial oscillation TBO medium medium medium medium no assessment B 9.5.3.4.2 [Gerald Meehl, United States of America]	Taken into account. Table updated
9-1672	9	57	44			Something is wrong with the column "Panel": All entries read "B" (which makes this column redundant), but panel B of Fig. 9.45 is about trends, not variability, which makes this column also wrong. [Andreas Sterl, Netherlands]	Taken into account. Table updated
9-1673	9	57	44			Table 9.3: change "level of evidence" to "amount of evidence" & "Level of agreement" to "degree of agreement" as used in the Uncertainty Guidance Note; is "Model Quality" the same as "Model skill" used in some sections of Chapter 9; The assessment is part of panel C (not B) in Figure 9.45; BTW, what is provided in panel C of Figure 9.45 is inconsistent with what is given in this table (e.g., SAM, PDO, AMOC are all given as low confidence in the table, but are in the medium confidence sector in the Figure) [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account. Table updated and consistency with Fig. 9.45 ensured
9-1674	9	57		57		table 9.3 - why doesn't this table start/include intraseasonal variability - it is the lead material/subsection in this section 9.5, especially since it is included in figure 9.45? [duane waliser, United States of America]	Taken into account. Table updated
9-1675	9	58	3	58	17	9.5.4. Regarding extreme events, shouldn't extratropical cyclones, just as tropical cyclones, be touched upon in this subsection? [Government of United States of America]	Noted - already touched upon in section 9.4.1.
9-1676	9	58	3	58	32	There is a very nice and interesting new paper by James Hansen et al. (2012) published recently in PNAS. This should be mentioned here, e.g. in line 23. [Martin Dameris, Germany]	Noted - the paper not cited as it is on observed changes in extreme events and thus not quite relevant here.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1677	9	58	5			rephrase the sentence; extreme events are characterised by state variables; those have pdfs and states that are associated with extreme events lie in the tails of a stat variable pdf ... [Peter Braesicke, United Kingdom]	Rejected - text unchanged as the current phrasing makes sense and is more concise.
9-1678	9	58	6			One could add "and evaluate" after "represent". [Adrian Simmons, United Kingdom]	Rejected - text unchanged as it doesn't seem necessary.
9-1679	9	58	21	59	9	Brief evaluation of simulation of temperature extremes by models. Needs emphasising for policy makers as demonstrates models do a good job. [European Union]	Noted - they are highlighted in ES, TS and SPM, though this section can't be expanded due to space limit.
9-1680	9	58	23			there are sampling and resolution challenges in reanalysis data as well, please mention [Peter Braesicke, United Kingdom]	Noted - text unchanged as the problem is not critical for temperature extremes.
9-1681	9	58	28	58	29	(and Figure 9.37a) I find this quite counter-intuitive and would expect that the mean model should do worse (in an RMS sense) for extremes, based on the smoothing effect of averaging the fields. How is the mean model defined? Is it actually the mean of the number of extremes in each model rather than the extremes in the mean model? Can you clarify the procedure? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted - text unchanged as it is rather obvious that it is the mean of extreme indices in each model.
9-1682	9	58	31	58	31	Write "Figure 9.37" instead of just "figure". [Farahnaz Khosrawi, Sweden]	Accepted.
9-1683	9	58	31	58	31	"CMIP ensemble performs slightly better" -> Can more quantitative language be used? [Thomas Stocker/ WGI TSU, Switzerland]	Noted - text revised according to the revised paper and 'slightly better' is no longer used.
9-1684	9	58	32			suggest inserting here the following to make the important point that patterns of extremes must be diagnosed with single realizations: "An analysis of a collection of single ensemble members from the same model found that one of the ensemble members that more closely simulated the so-called warming hole (little warming over the second half of the 20th century over the southeastern U.S., Kunkel et al., 2006) also produced a ratio of record high temperature to record low temperature that was closer to the observations (Meehl et al., 2012). This points to the importance of distinguishing between ensemble average multi-model results and single realizations in studying spatial patterns of extremes compared to observations. References: Kunkel, K.E., X.-Z. Liang, J. Zhu, Y. Lin, 2006: Can CGCMs simulate the Twentieth-Century "Warming Hole" in the Central United States? J. Climate, 19, 4137-4153. Meehl, G.A., J.M. Arblaster, and G. Branstator, 2012: Mechanisms contributing to the warming hole and the consequent U.S. east-west differential of heat extremes. J. Climate, doi: http://dx.doi.org/10.1175/JCLI-D-11-00655.1 . [Gerald Meehl, United States of America]	Noted - we gave up adding it due to space limitation, though we agree with the importance of the point.
9-1685	9	58				Table 9-3: I do not understand how for the Atlantic Meridional Mode (AMM) we can have high confidence about the variability even if we acknowledge that the model quality is low to represent this phenomenon. [RYM MSADEK, United States of America]	Taken into account. The entry should be understood as "high confidence that model quality is low". The caption is modified to clarify the reading of the table.
9-1686	9	59	2	59	9	suggest to elaborate on the inconsistency highlighted on line 9 between the results in Figure 9.37 with what is presented in Ch10 and in the CMIP3 models. Just stating that they are inconsistent without providing more details leaves the reader wondering why it is being highlighted in the first place [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - text revised.
9-1687	9	59	5	59	9	Does this statement imply that the models simulated larger variability than the observation? As the speculation is that climate variability in a warmer climate can be larger and extremely low temperatures may occur more frequently even in a warmer climate, this statement requires clarification to address its clear implications. [Government of Canada]	Noted - it doesn't imply such a speculation. Hopefully the expression is clear enough as it is adopted from IPCC SREX.
9-1688	9	59	5	59	9	"Interestingly, [...]" this sentence and particularly the final statement " – a result at odds with [...]" needs either to be explained in more detail or has to be phrased differently to avoid that readers who are not familiar with the methodologies applied in the cited studies get confused or draw wrong conclusions from this statement. [Jana Sillmann, Canada]	Taken into account - text revised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1689	9	59	5	59	9	There are several substantial differences in the methodologies and underlying datasets used in the studies this statement refers to: (1) Christidis et al. 2011 and Zwiers et al. 2011 use a much smaller sample of models than are used to produce Figure 9.37 (Sillmann et al 2012), also Sillmann et al 2012 use CMIP5 models instead of CMIP3 models. (2) Christidis et al 2011 and Zwiers et al. 2011 use detection and attribution methods based on extreme value statistics (GEV) to identify the human contribution to the observed trends, whereas Sillmann et al. 2012 show a global average of the temporal evolution of percentile threshold indices (I am assuming that you only refer to panels b-e in figure 9.37 with this statement as the portrait diagram does not explicitly show trends for TXx and TNn). (3) If my assumption is correct and you are referring to panels b-e (please also indicate that in the text), then also the variables used in the cited studies differ, as Christidis et al. 2011 and Zwiers et al. 2011 use the annual maxima(minima) of the maximum(minimum) temperature, whereas the panels show trends for the threshold exceedance rates of days with maximum(minimum) temperatures exceeding the 10th or 90th percentile of the base period. The mean biases in the model climatologies are thus not considered in these indices. [Jana Sillmann, Canada]	Taken into account - text revised.
9-1690	9	59	5	59	9	Currently we are working (paper in preparation) on an quantitative assessment of trends for the last 3 decades as simulated in the CMIP5 models and reanalyses compared to trends represented in various observational datasets (e.g., HadEX2, GHCNDEX) based on the ETCCDI indices. Preliminary results are displayed in our AGU 2012 poster (http://fallmeeting.agu.org/2012/eposters/eposter/gc43a-1017/). In the figures showing decadal trend estimates for the different datasets on a regional and global scale, one can clearly see that the median trend simulated in the CMIP5 ensemble for TXx and TNn is agreeing well with the observed range of trends on a global scale. On a regional scale, however, CMIP5 models tend to overestimate the recent trend in observed TXx and underestimate it for TNn particularly on the North American continent. Thus, it seems that the misrepresentation of trends is only a regional and not a general problem of the models and needs to be further investigated. [Jana Sillmann, Canada]	Taken into account - text revised.
9-1691	9	59	5	59	9	To conclude, since the studies you are referring to in this paragraph are not directly comparable, I don't think that the last statement "a results at odds..." is justified, and needs further clarification. It could also be further discussed in Chapter 10. [Jana Sillmann, Canada]	Taken into account - text revised.
9-1692	9	59	7	59	7	Can this be formulated more clearly? Does that mean that cold extremes get warmer? I thought these extremes get colder. Does that depend on which atmosphere regions is considered (troposphere or stratosphere)? [Farahnaz Khosrawi, Sweden]	Noted - text unchanged. It is correct that cold extremes get warmer. Hopefully the expression is clear enough as it is adopted from IPCC SREX.
9-1693	9	59	8			The sentence beginning "See also..." could be better explained. We recommend that the authors clarify this sentence. [Government of United States of America]	Taken into account - text revised.
9-1694	9	59	11	59	30	Figure 9.37 should include as many as possible other CMIP5 model results, including those of FGOALS-g2, FGOALS-s2, BNU-ESM and FIO-ESM. [Bin Wang, China]	Taken into account - as many CMIP5 models that are available on 15 March are included.
9-1695	9	59	12	59	14	I do not understand the first two columns of Fig. 9.37. As far as I understand the legend, colors indicate whether a model performs better or worse than the median model. How then can the median (mean) models perform better than the median model? [Andreas Sterl, Netherlands]	Noted - text unchanged. ENSmedian (or mean) indicates it takes median (or mean) of ensemble at each grid box, instead of being the model with median score.
9-1696	9	59	12	59	30	The figure caption is too long. Parts of the text should be included in the main text. [Farahnaz Khosrawi, Sweden]	Noted - caption unchanged as it should be self-contained.
9-1697	9	59	12			Figure 9.37: please add uncertainty to the observation- and reanalysis-based estimates [Thomas Stocker/WGI TSU, Switzerland]	Noted - we considered how to do it but came to a conclusion that there is no obvious and meaningful way of giving uncertainty range there.
9-1698	9	59	32	60	2	Discussion of precipitation extremes and models' ability to simulate them. Needs to be emphasised for policy makers who are concerned with flooding. [European Union]	Noted - they are highlighted in ES, TS and SPM, though this section can't be expanded due to space limit.
9-1699	9	59	32	60	2	9.5.4.2 The discussion in this section could be augmented by considering the results in "North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology" by Sheffield et al., specifically section 3.6 of that paper, which describes extreme precipitation and soil moisture	Noted - unfortunately the paper was not accepted in time.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						conditions in the context of the overall simulation of the hydroclimate in CMIP5 models. [Government of United States of America]	
9-1700	9	59	34			We are not sure that "uncertainty" is the right word here. Perhaps it is part of the message, but we think there are more representativeness issues with precipitation over land because precipitation is more heterogeneous in spatial coverage (and not subjected to the predictable diurnal cycle that temperature has). [Government of United States of America]	Noted - text unchanged. Our intention is that 'uncertainty' covers those issues and spatial representativeness is specifically discussed later.
9-1701	9	59	45	59	45	it should be "[...] fewer consecutive WET days [...]", not dry days. Please correct this. [Jana Sillmann, Canada]	Accepted.
9-1702	9	59	51	59	51	Terms such as "modest agreement", "reasonable agreement", "fairly well", "fairly realistically" etc, are used throughout Chapter 9 when comparing models with observations. It is not clear how these terms rate sequentially - is 'modest agreement' a better result than 'reasonable agreement'? It would be good if more quantitative terminology could be used where ever possible. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - text revised.
9-1703	9	59	51	59	52	(Kawazoe and Gutowski, 2012; Sakamoto et al., 2012; Wehner et al., 2010) --> (Kawazoe and Gutowski, 2012; Sakamoto et al., 2012; Wehner et al., 2010; Endo et al. 2012) Reference: Endo, H., A. Kitoh, T. Ose, R. Mizuta, and S. Kusunoki, 2012: Future changes and uncertainties in Asian precipitation simulated by multiphysics and multi-sea surface temperature ensemble experiments with high-resolution Meteorological Research Institute atmospheric general circulation models (MRI-AGCMs). J. Geophys. Res., doi:10.1029/2012JD017874. http://www.mri-jma.go.jp/Dep/cl/endo/paper/paper_12.pdf [Shoji Kusunoki, Japan]	Accepted.
9-1704	9	59	52	59	53	add a sentence along the lines "Two UK case studies suggest that the dependence of extreme precipitation on the large scale atmospheric circulation is in general captured (Maraun et al., 2012), and the phase of the annual cycle of extreme precipitation is in general well represented, whereas the amplitude of the annual cycle may be strongly misrepresented (Schindler et al., 2012)" Citations: Maraun, D.; Osborn, T. J.; Rust, H. W. (2012): The influence of synoptic airflow on UK daily precipitation extremes. Part II: regional climate model and E-OBS data validation. CLIMATE DYNAMICS 39(1-2): 287-301, DOI: 10.1007/s00382-011-1176-0 Schindler, A.; Maraun, D.; Luterbacher, J. (2012): Validation of the present day annual cycle in heavy precipitation over the British Islands simulated by 14 RCMs. JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES 117: D18107, DOI: 10.1029/2012JD017828 [Douglas Maraun, Germany]	Noted - we can't add it due to space limitation, though it is of some use.
9-1705	9	59	55	59	55	Expand "D & A" [HASIBUR RAHAMAN, India]	Spelled out to be 'detection and attribution'.
9-1706	9	59	55		58	I would add that the data available for the "observed" result were very scanty, so the large apparent trend may not be representative [William Ingram, United Kingdom]	Noted - I agree with the point, but text is unchanged as it only mentions the signs and relative magnitude relation between modeled and observed trends, and not the absolute magnitude of observed trends.
9-1707	9	59	55			"D & A study"??? [Barbara Früh, Germany]	Spelled out to be 'detection and attribution'.
9-1708	9	59	55			Most readers will not know that "D&A" means. Consider spelling out this acronym. [Government of Canada]	Spelled out to be 'detection and attribution'.
9-1709	9	59	55			Please explain what "D & A" is? [Government of United States of America]	Spelled out to be 'detection and attribution'.
9-1710	9	59	55			Jargon "D & A" should be spelt out. [Adrian Simmons, United Kingdom]	Spelled out to be 'detection and attribution'.
9-1711	9	59	55			explain "D&A" [Andreas Sterl, Netherlands]	Spelled out to be 'detection and attribution'.
9-1712	9	60	1	60	1	There is additional evidence that the climate models underestimate precipitation extremes. Please see Shiu, C.-J., S. C. Liu, C. Fu, A. Dai, and Y. Sun (2012), How much do precipitation extremes change in a warming climate?, Geophys. Res. Lett., 39, L17707, doi:10.1029/2012GL052762. [Shaw Liu, China]	Taken into account - the paper cited.
9-1713	9	60	1	60	1	Allan et al. 2010 doi:10.1088/1748-9326/5/2/025205 is more nuanced about discrepancies than Allan Soden (2008). [Timothy Merlis, United States of America]	Taken into account - the paper cited.
9-1714	9	60	1			The Allan & Soden (2008) reference should be removed in my opinion as that result has effectively been	Taken into account - the paper cited but Allan and

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						withdrawn by Allan & al (2010, ERL), who show it could well be an artifact of increased data availability to the reanalysis [William Ingram, United Kingdom]	Soden (2008) is kept as it is still relevant. In Allan et al. (2010), models' underestimation of heavy precipitation increase is still found, though its magnitude is revised.
9-1715	9	60	1			I have not read the Liu & al paper properly, but I see its abstract claims a rate of increase of mean temperature with warming an order of magnitude greater than is generally accepted (& by this SOD: bottom of 7-50) as energetically possible. So I don't think IPCC should be taking its other claims too seriously [William Ingram, United Kingdom]	Noted - although the large number itself might be some kind of artifact, the obs-model discrepancy in their claim is still meaningful and some methodological questions are addressed in their new paper (Shiu et al., 2012).
9-1716	9	60	2			Consider appending "...same tendency although O'Gorman (2012) [Nature Geosciences] find that observed interannual relationships offer a powerful constraint upon future projections of precipitation intensity." [Richard Allan, United Kingdom]	Taken into account - the paper cited and text revised considering its claim.
9-1717	9	60	4	60	23	Section 9.5.4.3: We suggest separate sections here for tropical cyclones and drought, since they are the only things in this "Other Extremes" section [Government of United States of America]	Accepted.
9-1718	9	60	4			9.5.4.3, Title "Other extremes". Very general title when only two extremes are considered [Government of Spain]	Noted - section now split into those for tropical cyclones and drought.
9-1719	9	60	6	60	12	The text is written as if the tropical cyclone intensity problem has been solved with this new 20 km model. In fact, models can be tuned to have higher intensity storms, but this does not mean they are getting the right answer for the right reasons. For example the Mizuta et al model lack ocean coupling, a known strong negative feedback on TC intensity. Other tests besides a simple measure of the models maximum intensity TC could be done (but have not) that could demonstrate skill with intensity. For example, does the model have operational skill in predicting intensities for real world storms? Does the model in climate mode have skill at reproducing the interannual or decadal variations in the pdf of intensity (years with more storms in the Atlantic also tend to be years with stronger storms on average). Does the model have a correlation of intensity with local SST like storms in the real world? Comparison of the full pdf of intensity vs observations and comparisons for individual basins? And so forth. TC modelers tend to push for much higher resolution than 20 km on the basis that storm structure (eye wall structure, precip banding etc) is much better simulated at higher resolution. [Government of United States of America]	Taken into account - text deleted.
9-1720	9	60	6	60	14	Some discussion of CMIP5-class climate model performance in the area of tropical cyclones should be considered, instead of limiting the discussion only to high-resolution atmospheric models. Suzana Camargo's paper titled "Global and regional aspects of tropical cyclone activity in the CMIP5 models," could be used as a resource for such a discussion. In addition, section 4 of the paper "North American Climate in CMIP5 Experiments. Part I: Evaluation of 20th Century Continental and Regional Climatology" by Sheffield et al. could further inform discussions of CMIP5 model performance in the area of direct simulation of tropical cyclones as well as model fidelity in the area of properly simulating environmental conditions that influence tropical cyclone development and propagation (e.g., wind shear). [Government of United States of America]	Noted - unfortunately the suggested two papers were not accepted in time.
9-1721	9	60	6	60	23	To us, this sub-section "other extremes" seems a bit thin. If a more comprehensive assessment regarding droughts, tropical and extra-tropical cyclones etc is not possible here, please refer the reader to the main conclusions coming out of the SREX on these topics. In addition, cross-referencing to the tropical and extra-tropical cyclone boxes of Chapter 14 would be useful. [Thomas Stocker/ WGI TSU, Switzerland]	Noted - text unchanged. Please note that there is no comprehensive assessment of model evaluation in SREX. Extra-tropical cyclones are assessed in 9.4.1 (a pointer added earlier). Cross-referencing to Chapter 14 doesn't seem very relevant as they are discussing projections. In addition, please understand that some discussion of other minor extremes has dropped from earlier draft due to space limitation.
9-1722	9	60	6			Please clarify what "high-resolution AGCMs" meant for AR4. High-resolution in AR4 is coarser than for AR5. We suggest adding a parenthetical definition. [Government of United States of America]	Noted - text unchanged as a 20 km resolution model (the highest resolution AGCM in CMIP5) was already used at the time of AR4. The following text mentions the fact.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1723	9	60	10	60	12	One could, again, also refer to NWP experience here. The current 16-km operational high-resolution global forecast model of ECMWF also does a reasonable job on the track and intensity of many tropical cyclones, the more so at shorter forecast ranges, improving on earlier versions that had lower horizontal resolution and poorer parametrizations, particularly of convection. [Adrian Simmons, United Kingdom]	Noted - text unchanged due to space limitation and only partial relevance of the suggested point.
9-1724	9	60	12	60	14	Stachan et al, 2012: Investigating Global Tropical Cyclone Activity with a Hierarchy of AGCMs: The Role of Model Resolution, J Clim, in press, show that the annual number and geographical distribution of tropical cyclones is well represented at resolutions of 135 km and higher, while simulating the interannual variability of storm occurrence requires resolutions of approximately 60 km. [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Taken into account - reference added.
9-1725	9	60	12			Regarding "driven only by observed sea surface temperature": Do you mean SST is the only driver, or that observed SST is required? Please consider rephrasing to clarify intended meaning. [Government of United States of America]	Taken into account - text revised. It means observed year-to-year variation of SST is required but no other observed year-to-year information is needed.
9-1726	9	60	14			Mention that Model performance in other basins for this metric was not as skillful (Zhao et al. 2009). [Government of United States of America]	Taken into account - text revised.
9-1727	9	60	16	60	23	The paper titled "Assessing Future Changes of Climate and Drought over the South-Central United States Projected by the CMIP5 Models" by Fu et al. could inform this discussion; it also informs the stated claim in those lines that "A comprehensive evaluation of CMIP5 models for drought is currently not available except..." In addition, Long et al. ("Drought and Persistent Wet Events Projected in the CMIP5 Experiments") could inform the discussion in this session, as well as broader discussions in the chapter on the topic of extreme precipitation events. [Government of United States of America]	Noted - unfortunately the suggested two papers were not accepted in time.
9-1728	9	60	18	60	18	Terms such as "modest agreement", "reasonable agreement", "fairly well", "reasonably well" etc, are used throughout Chapter 9 when comparing models with observations. It is not clear how these terms rate sequentially - is 'modest agreement' a better result than 'reasonable agreement'? It would be good if more quantitative terminology could be used where ever possible. [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - text revised.
9-1729	9	60	19			suggest to refer to Chapter 3 of SREX (Seneviratne et al. 2011) in addition to Ch2, WGI AR5 [Thomas Stocker/ WGI TSU, Switzerland]	Accepted.
9-1730	9	60	23			It is important to note that extremes associated with ENSO have been reasonably simulated, to be inserted here as follows: " Interannual variability associated with ENSO has also been shown to be associated with extremes in temperature and precipitation over the U.S., and these patterns of ENSO-related extremes are well-simulated in at least one climate model (Meehl et al., 2007)." Meehl, G.A., C. Tebaldi, H. Teng, and T. Peterson, 2007: Current and future U.S. weather extremes and El Niño. Geophys. Res. Lett., 34, L20704, doi:10.1029/2007GL031027. [Gerald Meehl, United States of America]	Rejected - though the point has some importance, there is no space to cite a single-model paper due to the length limit.
9-1731	9	60	27	60	44	This paragraph read like a table in text form. Just make a table out of it! [Andreas Sterl, Netherlands]	Rejected - we think this form works as it is.
9-1732	9	60	28			We recommend inserting "global" between "observed" and "warming trend" in order to clarify the statement. [Government of United States of America]	Accepted.
9-1733	9	60	35			"medium but different lines of evidence"? -- is "different" referring to "independent"? [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - it is 'independent'. Text revised.
9-1734	9	60	38	60	40	There is limited evidence that 20km grid AGCMs can simulate a realistic distribution of TC intensity. There is medium evidence that medium resolution (50-120 km grid spacing) global models forced with observed SSTs can simulate the year to year count variability of Atlantic hurricanes remarkably well, though with less skill in other basins. [Government of United States of America]	Taken into account - TC intensity in AGCMs is no longer highlighted in the summary also due to space limitation.
9-1735	9	60	39	60	39	Pls change "~20km" to "~20 km" [HASIBUR RAHAMAN, India]	Editorial
9-1736	9	60	42	60	44	Please, make explicit that the summary of performance of CMIP5 models is limited due to the scarce number of CMIP5 results [Government of Spain]	Rejected - we think there are substantial number of models in the final version of Figure 9.38.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1737	9	60	46			not sure the order is right, downscaling first and a subsequent discussion of extreme events (even though I appreciate that SREX uses global models) would work better for me [Peter Braesicke, United Kingdom]	Noted. Indeed, the suggested order of the topics would be an alternative, but there is also merit and motivation for the structure as in the Second-Order-Draft.
9-1738	9	60	55			Consider changing "may be too low" to "is often too low". That is why downscaling is so common. [Government of United States of America]	Accepted.
9-1739	9	60	57	60	57	Consider revising "region-specific climate information" to read "region-specific climate projections", given that this section is discussing climate model simulations and not information derived from physical observations. [Government of Canada]	Rejected. "Projections" usually refer to climate change simulations, not climate model simulations in general. Climate modelling is one way of generating information on climate, and can complement information gleaned from measurements.
9-1740	9	60				9.5.4.4 summary - Suggest adding a statement similar to: In some cases, extremes are attributable to features such as the MJO or TCs. A model's ability to simulate these features will determine the confidence in future trends of extremes. [Jaclyn Brown, Australia]	Rejected - though it is true, the statement is not suitable for the summary where assessment results of this section is summarized.
9-1741	9	61	4	63	50	9.6.1 The pattern of cooling over the South Eastern United States over the 20th century presents a challenging case to evaluate the model's ability to reconcile the general warming influence of the greenhouse gas forcing with regional influences of natural variability on decadal time scales. This feature was not well simulated by the CMIP3 models, and is not authoritatively better in the CMIP5 models, but it perhaps merits discussion. There are two papers the authors may wish to consult to inform a discussion, should they decide to include one: "Multi-decadal Climate Variability and the 'Warming Hole' in North America – results from CMIP5 20th and 21st Century Climate Simulations" by Kumar et al. and "Inter-model variability and mechanism attribution of central and southeastern U.S. anomalous cooling in the 20th century as simulated by CMIP5 models" by Pan et al. [Government of United States of America]	Rejected. The comment concerns trends, rather than climatological means as is the focus here.
9-1742	9	61	6	61	17	References are required. [IGOR SHKOLNIK, RUSSIAN FEDERATION]	Noted. The relevant figure and discussion are based on the information in the CMIP database, accessed by the Chapter 9 authors.
9-1743	9	61	8	61	7	is contained really the right expression in this context? [Farahnaz Khosrawi, Sweden]	Accepted. The meaning is more as "addressed"
9-1744	9	61	8	61	8	Which regions? I guess you mean latitudinal regions. You should add that. [Farahnaz Khosrawi, Sweden]	Noted. The reference is to the regions in Fig. 9.38.
9-1745	9	61	14	61	17	Why here? Please give citation! [Peter Braesicke, United Kingdom]	Noted. The discussion is based on the figure considered in the paragraph (9.38 in the SOD).
9-1746	9	61	20			Figure 9.38: suggest to add uncertainty for the observational data in the figure; duplication of map with regional outlines is not needed. Suggest to add one separate map and rearrange the seasonal panels [Thomas Stocker/ WGI TSU, Switzerland]	Accepted. Figure was re-done including with additional observational data.
9-1747	9	61	21	61	21	I would suggest to write "model mean" instead of "average". [Farahnaz Khosrawi, Sweden]	Rejected. "Average" and "mean" are the same.
9-1748	9	61	22	61	22	I would suggest to write "model mean" instead of "average". [Farahnaz Khosrawi, Sweden]	Rejected. "Average" and "mean" are the same.
9-1749	9	61	22	61	22	Write "each model is given by a thin red line", otherwise it is somewhat confusing. [Farahnaz Khosrawi, Sweden]	Editorial
9-1750	9	61	22	61	22	Same as above: Write "each model is given by a thin green line", otherwise it is somewhat confusing.r [Farahnaz Khosrawi, Sweden]	Editorial
9-1751	9	61	28	61	30	"Substantial cold biases over the Northern Hemisphere regions are more prevalent in winter than summer". Are such biases related to the underestimation of the rising trend of extremely low temperatures? [Government of Canada]	Noted. Not every aspect can be assessed here.
9-1752	9	61	28			Consider clarifying "temperature" as "mean temperature" (assuming that is what the data refer to). Presumably the ranges of impacts on daily/monthly maximum/minimum temperatures would be different.	Accepted.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						[Government of United States of America]	
9-1753	9	61	40	61	40	... shown in the... [European Union]	Editorial.
9-1754	9	61	40	61	40	typo: take out second 'in' [CLAUDIA STUBENRAUCH, France]	Editorial.
9-1755	9	61	40			Consider clarifying "precipitation" as "seasonal total precipitation" (if that is what is intended). Presumably the impacts on daily totals have not been evaluated in the supporting text. [Government of United States of America]	Accepted. Clarified.
9-1756	9	61	54			Presumably "uncertainty" is only one factor that affects evaluating precipitation. Perhaps "scarcity" and/or "representativeness" would add depth to that statement. [Government of United States of America]	Accepted. Clarified.
9-1757	9	61				The Fig. 9.38 underpins a high degree of similarity between CMIP3 and CMIP5 simulations rather than increased accuracy in representing annual cycles across regions in CMIP5 models. [IGOR SHKOLNIK, RUSSIAN FEDERATION]	Noted. We find that this is expressed in the text.
9-1758	9	62	3	62	4	"In contrast, Maris et al. (2012) found a general overestimation of GCM-simulated Antarctic land precipitation, especially inland." This sentence is confusing. If it means that the GCMs overestimated the observations, suggest rewording as: "In contrast, Maris et al. (2012) found GCM-simulated Antarctic land precipitation overestimated the observed data, especially inland." If this is not the correct interpretation, please clarify. [Government of Canada]	Accepted. Clarified.
9-1759	9	62	10	62	11	The sentence about the regions' (not region's!) should be placed at the end of this legend, together with the description of the regions. [Andreas Sterl, Netherlands]	Editorial
9-1760	9	62	25	62	31	Please consider to illustrate these regions by a map as done in Figure 9.38. [Government of Germany]	Rejected. While the suggestion has merit, the number of regions is so large as to lead into an overly complicated figure.
9-1761	9	62	36	62	36	It is useful to have these magnitudes. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted.
9-1762	9	62	39	62	40	So what spatial scale of averaging is required to make the projected change greater than the present day bias? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted. The appropriate spatial scale depends on the magnitude of both the bias and the signal, and thus varies with the model, region, variable, its natural variability and the amount of the climate change.
9-1763	9	62	41			The statement about the weak relation between present-day climate and projected changes needs some modification. There are several indications of a relation between present-day performance and projections. Some of them are even mentioned in this chapter (section 9.8.3.2 Emergent Constraints): snow-albedo feedback (p 9-81, l 8-10), land carbon sensitivity to warming (p 9-91, l 25-29). Fischer and Schär (Fischer, E.M., J. Rajczak, and C. Schär (2012) changes in European summer temperature variability revisited, Geophys. Res. Let., 39, L19702, doi: 10.1029/2012GL052730) show that the simulation of the latent-to-sensible heat flux ratio in the present-day climate strongly influences projections of temperature variability. [Andreas Sterl, Netherlands]	Taken into account. Noted on in section 9.6.6 with further reference to the discussion contained in 9.8.
9-1764	9	62	49	62	51	"Masson and Knutti (2011a) found optimal scales of smoothing to vary from the grid-point scale to around 2000 km depending on the variable and the region in question." It is not clear what is meant by "optimal scales of smoothing", as this would depend on how the smoothed data are to be used. Consider clarifying. [Government of Canada]	Taken into account. Clarified.
9-1765	9	62	52	62	52	Obtaining or imposing? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted. "Obtaining" is correct.
9-1766	9	63	1	63	2	References are required. [IGOR SHKOLNIK, RUSSIAN FEDERATION]	Rejected. The discussion pertains to analysis of CMIP archives, performed by the Chapter 9 writing team.
9-1767	9	63	1			This sentence should be supported by references. [Government of United States of America]	Rejected. The discussion pertains to analysis of CMIP archives, performed by the Chapter 9 writing team.
9-1768	9	63	4	63	5	Pls change "Cattiaux et al. (Cattiaux et al., 2012)" to "Cattiaux et al., (2012)" [HASIBUR RAHAMAN, India]	Noted. Editorial.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1769	9	63	8	63	27	Discussion of model simulations of net precipitation over Antarctica. Important points for decision makers. [European Union]	Noted. Note: This topic is in the final draft addressed in Chapter 13 and thus removed from Chapter 9.
9-1770	9	63	8			Section 9.6.1.1.1. It is useful that you have this section. We have a similar discussion in 13.3.3.2, page 19, lines 24-45. We should consider merging these discussions, and containing the material either in ch9 or in ch13, not both, each with a reference to the other. [Jonathan Gregory, United Kingdom]	Noted. Note: This topic is in the final draft addressed in Chapter 13 and thus removed from Chapter 9.
9-1771	9	63	9			"Net precipitation" is a poor term for P-E, especially as "net" is used in a different sense near the end of this same line. P-E is the "net excess of precipitation over evaporation". [Adrian Simmons, United Kingdom]	Noted. Note: This topic is in the final draft addressed in Chapter 13 and thus removed from Chapter 9.
9-1772	9	63	18	63	18	Pls change "1989-2009" to "1989-2009" [HASIBUR RAHAMAN, India]	Editorial. Note: This topic is in the final draft addressed in Chapter 13 and thus removed from Chapter 9.
9-1773	9	63	29	63	41	include reference: Nguyen et al., 2011 (Global 60 km simulations with CCAM: evaluation over the tropics) [Government of Australia]	Rejected. The suggested paper has partial relevance, but it mainly focuses on large-scale fields rather than regional scale. AR5 is an assessment and not a review.
9-1774	9	63	39	63	41	We recommend that the authors change the sentence to: "Also the fact that AGCMs do not simulate coupled interactions with the ocean may limit their ability to represent some important phenomena such as the cold wake in surface ocean waters generated by a tropical cyclone (e.g., Knutson et al. 2001; Hasagawa and Emori 2007)." Note that the Zhou et al 2009 paper cited here uses SST forced models, not ocean coupling. New Reference: Knutson, T R., R E Tuleya, W Shen, and I Ginis, 2001: Impact of CO2-induced warming on hurricane intensities simulated in a hurricane model with ocean coupling. Journal of Climate, 14(11), 2458-2468. [Government of United States of America]	Accepted. Text modified in line with the suggestion. The suggested additional reference is not cited as it is from 2001.
9-1775	9	63	45	63	50	Grose et al. (2012) showed improvement in jet stream position over Tasmania using a stretched grid model. In Bennett et al. (2012) simulations of hydrological flows in Tasmania was simulated using a stretched grid. [Government of Australia]	Taken into account. We cited White et al. (2013) as an example for studies for the region. White et al. discusses the impact on extremes, which complements the rest of the section.
9-1776	9	63	45	63	50	The literature review here is scant. A useful addition would be: Nguyen et al. (2012) Global 60 km simulations with CCAM: evaluation over the tropics, Clim. Dyn., 39, 637–654, doi: 10.1007/s00382-011-1197-8. [Government of Australia]	Rejected. The study quoted is an AGCM-one (cf. comment 9-1773), not one with a variable-resolution global model.
9-1777	9	63	45	63	50	Zou et al. (2010) showed that the variable-resolution AGCM of Laboratoire de Meteorologie Dynamique-zoom can be a useful tool for the dynamical downscaling of rainfall variability over eastern China. [Reference: Zou, L., T. Zhou, L. Li, and J. Zhang, 2010: East China summer rainfall variability of 1958-2000: Dynamical downscaling with a variable-resolution AGCM, J.Climate, 23, 6394-6408,doi: 10.1175/2010JCLI3689.1.] [Tianjun Zhou, China]	Noted. The suggestion of additional references was considered in the assessment point of view and while there is merit, the reference is to a specific finding.
9-1778	9	63	48	63	48	Rabinovitz et al (2008) is not included in the references [Fabrice Chauvin, France]	Editorial.
9-1779	9	63	51	63	51	Mention should be done to Caron et al (2011): Caron, L.-P., C. G. Jones, and K. Winger, 2011: Impact of resolution and downscaling technique in simulating recent Atlantic tropical cyclone activity. Clim. Dyn., 37, 869-892. [Fabrice Chauvin, France]	Noted. The suggestion of additional references was considered in the assessment point of view.
9-1780	9	63	51	63	51	Idem for Daloz et al (2012): Daloz, A.-S., F. Chauvin, and F. Roux, 2012: Impact of the configuration of stretching and ocean-atmosphere coupling on tropical cyclone activity in the variable-resolution GCM ARPEGE. Clim. Dyn., 39, 2343-2359. [Fabrice Chauvin, France]	Noted. The suggestion of additional references was considered in the assessment point of view.
9-1781	9	63	52	66	28	Section 9.6.2 discusses Regional Climate Downscaling and 9.6.3 discusses the Skill of Downscaling Methods. A substantial amount of research has been undertaken in Australia on downscaling but there are limited references to this work. Work up until ~2011 is well summarised in Frost et al (2011) J Hydrol 408 1-18 A comparison of multi-site daily rainfall downscaling techniques under Australian conditions. [Government of Australia]	Taken into account and included. The paper is cited in Section 9.6.5 (see also comment 9-1848).
9-1782	9	63				FAQ 10.2, Fig. 1: The y-axis in x-y-plots cannot be read, please improve. The color bar should go from blue	Noted. Wrong chapter.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						(cold) to red (warm9, not vice versa as in the current version. [Government of Germany]	
9-1783	9	64	6			Statistical [Frank Kreienkamp, Germany]	Editorial.
9-1784	9	64	10	64	11	Can it be said that RCMs also rely on the stationarity hypothesis to some extent since it is assumed that the tuning and/or bias correction used to produce realistic simulations of present-day conditions applies under changed climatic conditions? [Government of Australia]	Noted. (Cf 9-1863) Climate model parameterisations (both for global and regional climate models) are based on physical understanding of the underlying processes and phenomena, rather than statistical fits between different sets of data. Model tuning is an aspect that has bearing as to the review-comment, and is discussed elsewhere in Chapter 9.
9-1785	9	64	10	64	11	This is a physical assumption, not really a statistical one. It would be useful to describe some of the physical implications. For example, it makes sense to assume that small scale physical relationships (water holding capacity of air varying with temperature, for instance) hold in the future, but this may not be the case for large scale emergent relationships (such as the effect of sea surface temperature on regional circulation patterns). [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted. Behind the statistical relation there is certainly a physical relation but what is assumed in the stationarity hypothesis is that the STATISTICAL relation between the predictors (generally larger scale fields) and the predictand (smaller-scale field) remains the same in the future, which can not necessarily be verified due to for example the fact that this statistical relation was constructed using base data that does not need to resemble the one obtained under different radiative forcing. I.e., the future conditions may be "outside" the base data.
9-1786	9	64	11			"Present-day relationships" between what? [Government of United States of America]	Accepted. Revised.
9-1787	9	64	17			Please add a reference after "different methods". [Government of United States of America]	Noted. Two references had been provided. (van Vliet et al., 2011; Vrac and Naveau, 2008)
9-1788	9	64	20			In addition to Emanuel et al. 2008, the authors may add Villarini and Vecchi (2012). Reference: Villarini, G., and G.A Vecchi (2012): Projected Increases in North Atlantic Tropical Cyclone Intensity from CMIP5 Models. J. Climate (submitted). [Government of United States of America]	Noted. The reference was considered from the assessment point of view,
9-1789	9	64	21	64	23	There are a lot techniques that preserve the physical consistency. Not only Zhang ... 2011. Please add a few lines stating that this is the minimum requirement for most impact studies. If a SD does not preserve the physical (between elements) and regional (stations or gridpoints) consistency how can it be used for hydrological studies and so on. (cites for a SD from Germany being in use since 1999 having both consistencies: Enke et al 2005:Results of five regional climate studies applying a weather pattern based downscaling method to ECHAM4 climate simulations. Meteorologische Zeitschrift, Vol. 14, No. 2, 247-257; Huebener et al. 2010 Regional Climate Change and Impact Assessment for the Federal State Hesse, Germany, and Implications of the Global 2 °C Climate Target. ImTech; Kreienkamp et al 2011: Climate Signals on the Regional Scale Derived with a Statistical Method: Relevance of the Driving Model's Resolution. Atmosphere 2011, 2, 129-145; doi:10.3390/atmos2020129) [Frank Kreienkamp, Germany]	Taken into account. The discussion was focused on evaluation of climate models and in the case of 9.6, downscaling in a broader sense. Climate impact research aspects are of the remit of Working Group II.
9-1790	9	64	32	64	34	'... SD ... provide a valuable addition..' --- even if you are not able to provide an overall evaluation why are SD just a valuable addition? And to what? Please rewrite this 3 lines. Separating between there is no overall evaluation (may add a link the the COST action VALUE) and the value (not additional) of SD for impact studies. [Frank Kreienkamp, Germany]	Accepted. The sentence has been revised. The value of SD for impact studies is not of explicit relevancy to the topic of this chapter (rather to WGII) and is not discussed in more detail here.
9-1791	9	64	33	64	33	The text refers to the variety of statistical downscaling schemes, but this statement is not supported by a comprehensive list of the schemes available together with key references illustrating each approach. Also, targeted intercomparisons do exist (see, e.g., Chiew et al. (2010) Comparison of runoff modelled using rainfall from different downscaling methods for historical and future climates, J. Hydrol., 387, 10-23, doi:10.1016/j.jhydrol.2010.03.025) and studies such as these should be cited. [Government of Australia]	Taken into account. The aim here is not to list all the SD schemes, but to briefly assess new developments.
9-1792	9	64	36			This subsection consists roughly of reports on four groups of topics. (1) Following an introductory sentence (lines 38-40), only two examples of people of groups addressing benefits RCMs might or have achieved compared to driver GCMs. (2) Very numerous reports, and a figure, on what happened when people or	Noted.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						groups increased resolution of their RCMs. Vautard et al. (2012), lines 47-48, belonging to this group, perhaps by mistake is found in the paragraph above, where (1) is dealt with. (3) Several reports on groups coupling their atmospheric RCMs to other components, or of adding parameterizations of some non-standard processes (e.g., drifting snow!). (4) A very comprehensive report on regions for which RCMs have been used, and experiments organized, along the way mixed with some results achieved. [Fedor Mesinger, United States, and Serbia]	
9-1793	9	64	43			Another supporting reference for high-resolution is Otte et al., J. Climate, 2012, who used historical data and showed improvement in mesoscale variability and extremes with a RCM relative to high-resolution reanalysis. [Government of United States of America]	Taken into account. (We that the reference is to Otte et al. 2012, J Climate 25, 7046-7066, which we consider in the context of nudging, Section 9.6.5.)
9-1794	9	64	48	64	48	Consider to add: Furthermore, it is reported that over some regions like East Asia, use of high resolution RCM does not only show more detailed geographic features but it also provides a much better reproduction of the large scale monsoon precipitation pattern over the region (Gao et al., 2008; 2012). References: Gao XJ, Shi Y, Song RY, Giorgi F, Wang YG, Zhang DF, 2008. Reduction of future monsoon precipitation over China: Comparison between a high resolution RCM simulation and the driving GCM. Meteorology and Atmospheric Physics, 100, 73–86. doi: 10.1007/s00703-008-0296-5 Gao XJ, Shi Y, Zhang DF, Wu J, Giorgi F, Ji ZM, Wang YG, 2012. Uncertainties of monsoon precipitation projection over China: Results from two high resolution RCM simulations. Climate Research. 52, 213–226. doi: 10.3354/cr01084 [Xuejie Gao, China]	Taken into account. The more recent reference has been cited in Section 9.6.4.
9-1795	9	64	48	64	48	Include "resolution" so that it reads "50 km resolution runs". [Farahnaz Khosrawi, Sweden]	Editorial
9-1796	9	64	48	64	48	Pls give a gap after km [HASIBUR RAHAMAN, India]	Editorial
9-1797	9	64	50	65	12	Effect of resolution on performance of regional climate model simulations. Important points for decision makers who will use model projections for forming policy. [European Union]	Noted
9-1798	9	64	53	64	54	A logical statement; spatial and temporal scales are not independent and prescribing boundary conditions too often constraints the model quite severely. [Peter Braesicke, United Kingdom]	Noted.
9-1799	9	64	54	64	54	Concerning "higher temporal resolution led to some degradation"; please briefly comment on the reason. In Woollings et al. (2010b) it reads "This may be because, as described in Sect. 2.3, the sub-monthly SST variance in the R02 dataset has an unrealistically large spatial scale, and that this in fact worsens the simulation.". Therefore, spatial and temporal resolution are closely connected. [Government of Germany]	Taken into account. The discussion of the finding is, however, moved to Section 9.6.5 as is more concerns boundary conditions (SST) rather than the model itself.
9-1800	9	64	54			degradation of what [Frank Kreienkamp, Germany]	Taken into account. Cf the response to the comment 9-1799.
9-1801	9	64	56	64	56	... that when... [European Union]	Editorial
9-1802	9	64	57			Unclear what "highly variable regional orography" means? [Government of Australia]	Noted. What is expressed is mountain regions or combination of mountains, valleys, plains, i.e. significant local-to-regional scale variability in terrain heights.
9-1803	9	64				Presence of some of these topics in the subsection "Developments of Dynamical Models" I find puzzling. It requires no development to run an RCM at higher resolution. All these references and summaries of some results achieved running RCMs at higher resolution belong more appropriately to the Value Added section. But I also wonder are all these references really useful to anyone interested in climate change? If a group had difficulties achieving improved results running at higher resolution this has to be a specific problem of a particular model and I feel the SOD authors should consider if it is really useful to have the group's problem included in the AR5. A general comment on difficulties encountered should do. [Fedor Mesinger, United States, and Serbia]	Noted. Chapter 9 is about evaluation of climate models, which underlines the use of these models in climate projections. While different models do have different improvement needs, the pursuant of higher resolution and the evaluation of these development s has relevant general aspects.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1804	9	64				I have found it somewhat strange to have a list of regions covered by RCMs studies since AR4 here with hardly any mention of what were the results. This being a report on climate change I suggest this list to be expanded by pointing out some of the results. Or, are the lines 42-49 meant to be a brief summary of the results? They sound more like a summary of what one could do (Christensen et al.), or how have the RCMs performed in testing mode (e.g., "RCMs run for Africa ... did well for ... but ..."). [Fedor Mesinger, United States, and Serbia]	Noted. The overall 9.6 considers RCM (and other methods) developments since the AR4. Although the application of RCMs has certainly become ever more widespread, the coordinated CORDEX studies is a major new (very recent) development, involving several RCMs and many regions. Many CORDEX results are underway, but not available in published literature in time for the AR5.
9-1805	9	64				What however I find missing in this subsection is more of an effort to identify actual developments, as the title of the subsection announces; new features in the design of RCMs, since AR4, aimed to advance or deeper explore the ability of RCMs to improve on what can be done by GCMs or AOGCMs. Some of this is included in the Value Added section, such as three references to papers in which two-way nesting was used, as opposed to simply forcing on the RCM boundary conditions of the global model, relaxation style. These are the references to papers of Inatsu and Kimoto, Inatsu et al., and Lorenz and Jacob. In these three papers two-way interacting lateral boundary conditions have been used, which do not take the global model results as given but have the RCM impact the driver global fields. [Fedor Mesinger, United States, and Serbia]	Noted. The literature is at present very scarce on two-way nesting. Identified papers are cited.
9-1806	9	64				I suggest that this feature be emphasized; it empowers the RCM to do better than if it were simply forced by the given global fields. The downside of course is that with the two-way interaction the global model and the RCM need to be run at the same time, which not many groups can afford. [Fedor Mesinger, United States, and Serbia]	Noted. Belongs together with the comment 9-1805.
9-1807	9	64				May I suggest the SOD authors to also note a study of the use of boundary conditions less forceful than relaxation, in use in the Eta model, with variables prescribed only along the nested model's outermost row of grid points, and one less variable prescribed at the outflow than at the inflow points (Mesinger and Veljovic, 2012). Such boundary conditions are more consistent with the mathematical nature of the problem, and allow the nested model freedom to depart to a greater degree from the driver data. In tests of Mesinger and Veljovic they resulted in forecasts in most cases in better agreement with analyses than the driver ECMWF 32-day ensemble forecasts, in spite of the downside of the nested model having to overcome the boundary conditions error. [Fedor Mesinger, United States, and Serbia]	Noted. The paper considers NWP/monthly-scale forecasting experiments, and as such is out of scope in Chapter 9.
9-1808	9	64				Reference not in the SOD: [Fedor Mesinger, United States, and Serbia]	Noted. Belongs together with the comments 9-1805--1807 and 9-1809.
9-1809	9	64				Mesinger, F., and K. Veljovic, 2012: Limited area NWP and regional climate modeling: A test of the relaxation vs Eta lateral boundary conditions. Meteor. Atmos. Phys., in press [Available as Online First.] [Fedor Mesinger, United States, and Serbia]	Noted. Belongs together with the comment 9-1807!
9-1810	9	65	1	65	12	The authors may note here that other model features besides resolution (i.e., physical parameterizations) can affect some of these same model features. For example, models of a given resolution may simulate tropical cyclones with quite different upper limit intensities for the present climate due to different physical parameterizations (e.g. two versions of MRI AGCM: AGCM20_3.1 vs 3.2 analyzed in Murakami et al. 2012). [Government of United States of America]	Noted. The aspect of physical parameterisations was very briefly mentioned in Section 9.6.5. The issue is one that is shared with global climate models, and there is more information in the rest of Chapter 9.
9-1811	9	65	3	65	3	Include "resolution" so that it reads "to 12 km resolution". [Farahnaz Khosrawi, Sweden]	Editorial
9-1812	9	65	6	65	7	"Pryor et al. (2012) noted that the impact of higher RCM resolution (6 km compared to 50 km) was much larger in extreme wind speeds than in mean wind speed." Consider specifying as to what the "impact" was: did it make the simulated values of extreme wind speeds more accurate or less? Or does it mean that the sensitivity of estimated extreme windspeeds to increased model resolution was greater (without any apparent improvement in accuracy)? [Government of Canada]	Accepted. Clarified better that the Pryor et al. example was one of how resolution affects the simulation of extremes, in this case of winds.
9-1813	9	65	21	65	21	Seems something is missing in this sentence. Do you mean "ocean parameterization" or "ocean simulation"? [Farahnaz Khosrawi, Sweden]	Rejected. "Interactive ocean" is rather common expression in climate modelling context.
9-1814	9	65	21	65	29	In recent years, coupled regional climate models with interactive ocean are also developed with focus on western North Pacific (WNP) summer monsoon (Zou and Zhou, 2011, 2012a, 2012b). Zou and Zhou (2012a)	Taken into account. Zou and Zhou 2012(b) have been cited as an example of these developments.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						investigated the impacts of two regional atmospheric models on the performance of regional ocean-atmosphere coupled model (ROAM) in simulating the WNPSM of 1998. Results suggest that the biases of vertical profile of temperature and specific humidity in stand-alone simulations may be responsible for the different SST biases in regionally coupled simulations. The cold biases of simulated SST by the ROAM over WNP are significantly reduced by applying the convection suppression criterion (Zou and Zhou, 2011). The interannual variability of WNPSM is better simulated in the ROAM than that in the uncoupled simulation by reducing the unrealistic and too strong SST forcing (Zou and Zhou, 2012b). [References: Zou, L., and T. Zhou, 2011: Sensitivity of a regional ocean-atmosphere coupled model to convection parameterization over western North Pacific. <i>J. Geophys. Res.</i> , 116, D18106, doi: 10.1029/2011JD015844. Zou, L., and T. Zhou, 2012a: Development and evaluation of a regional ocean-atmosphere coupled model with focus on the western North Pacific summer monsoon simulation: Impacts of different atmospheric components. <i>Sci. China Earth Sci.</i> , 55, 802-815. Zou, L., and T. Zhou, 2012b: Can a Regional Ocean-Atmosphere Coupled Model Improve the Simulation of the Interannual Variability of the Western North Pacific Summer Monsoon? <i>J. Climate</i> , doi:10.1175/JCLI-D-11-00722.1, In press [Tianjun Zhou, China]	
9-1815	9	65	22	65	22	Same here. Seems there is a word missing after "ice". [Farahnaz Khosrawi, Sweden]	Editorial. See response to comment 9-1813.
9-1816	9	65	31	65	49	AustralAsia is a CORDEX domain, so it would be appropriate to cite a study carried out by an Australian group. An example is: Evans et al. (2012) Evaluating the performance of a WRF physics ensemble over South-East Australia, <i>Clim. Dyn.</i> , 39, 1241–1258, doi: 10.1007/s00382-011-1244-5. [Government of Australia]	Taken into account. Indeed, it would have been very appropriate to also cite CORDEX results for AustralAsia, but results have so far not been published. However, the Evans et al. (2012) that concerns a single model with which various parameterisation options have been explored is cited in 9.6.5.
9-1817	9	65	33	65	33	Add reference to "Nguyen et al., 2011" [Government of Australia]	Rejected. The reference is to a study with an AGCM, not a regional climate model. (Cf. comment 9-1818.)
9-1818	9	65	37			Include mention of the Pacific with reference to: Nguyen et al., 2011 [Government of Australia]	Rejected. The reference is to a study with an AGCM, not a regional climate model.
9-1819	9	65	38			Please add the Atlantic basin TCs to this list (Knutson et al. 2007; Bender et al. 2010). References: Knutson, T R., J J Sirutis, S T Garner, I M Held, and R E Tuleya, 2007: Simulation of the Recent Multidecadal Increase of Atlantic Hurricane Activity Using an 18-km-Grid Regional Model. <i>Bulletin of the American Meteorological Society</i> , 88(10), doi:10.1175/BAMS-88-10-1549. and Bender, M A., T R Knutson, R E Tuleya, J J Sirutis, G A Vecchi, S T Garner, and I M Held, 2010: Modeled impact of anthropogenic warming on the frequency of intense Atlantic hurricanes. <i>Science</i> , 327(5964), doi:10.1126/science.1180568. [Government of United States of America]	Taken into account. Bender et al. was cited in conjunction with value added section.
9-1820	9	65	44			Add reference to Déqué et al (2012), see chapter 14, p 106. It is about uncertainties, the Christensen et al paper is about weighting [Erik Kjellström, Sweden]	Taken into account. Déqué et al. (2012) is now cited in Section 9.6.6, given its focus on reasons for spread (of results) among RCMs.
9-1821	9	65	47			Please explain the ERA-Interim statement (boundary condition, ref. data, performance measured against a third data set?) [Peter Braesicke, United Kingdom]	Taken into account. The authors used ERA-Interim as boundary conditions, as stated in the section. The simulation results were evaluated against several observational datasets. The sentence has been clarified.
9-1822	9	65	53	66	28	This subsection focuses almost exclusively on dynamical downscaling methods. Studies of the skill of statistical downscaling methods also exist for regions already undergoing change (see, e.g., Charles et al. (2007). Effect of GCM bias on downscaled precipitation and runoff projections for the Serpentine catchment, Western Australia. <i>Int. J. Climatol.</i> , 27, 1673-1690, 10.1002/joc.1508). [Government of Australia]	Noted. The Chapter is about climate models, but information on statistical downscaling is covered as it is a parallel method to obtain regional-scale information. Chapter 9 concerns model evaluation, projections being the remit of other Chapters, as well as WGII.
9-1823	9	66	10	66	11	Is it possible to improve the skill of the downscaling method via bias correction? Or is bias correction used to	Noted. Discussion on bias correction is removed from

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						fit the resulting values to a given (may be observed) value? Bias correction is the last. Please rewrite. [Frank Kreienkamp, Germany]	Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of model outputs.
9-1824	9	66	10	66	18	A couple of lines stating the problems of bias correction is missing. Many articles are much more critical about the 'improvements' due to bias correction. (Ehret? et al 2012 Hydrol. Earth Syst. Sci. Discuss) If there is a discussion about the problems some where please add a link [Frank Kreienkamp, Germany]	Noted. Discussion on bias correction is removed from Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of model outputs.
9-1825	9	66	10	66	18	I find this paragraph a bit too unprecise. I would distinguish between different error types which may or may not be correctable by bias correction. E.g., it is very unlikely that bias correction can correct for GCM circulation errors, but it is very likely that it can correct for local scale orographic and parameterisation effects. Unfortunately the literature on this topic is still very sparse. One good example is the paper by Eden et al. (2012): Skill, correction and downscaling of GCM-simulated precipitation, J Climate, published online, DOI:10.1175/JCLIM-D-11-00254.1. They argue explicitly that bias correction can be used for correction of precipitation parameterisations and orographic effects (see conclusions). [Douglas Maraun, Germany]	Noted. Discussion on bias correction is removed from Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of model outputs.
9-1826	9	66	10		18	In this section it is talked about "bias correction". This terminology implies that the resulting data from the models are "correct" in some sense. As observations do contain errors this is a misleading term and depending on the degree of realism and representativity in the observations the "corrected" model output may be more or less erroneous. There are plenty of examples in the literature about deficiencies in often used data sets like the E-OBS data in Europe, in particular in relation to the quality of extreme precipitation. One vry good example is Lenderink et al (2010) showing that model performance in terms of precipitation extremes is better in areas where the density of the observational network is higher. (Lenderink, G. 2010. Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations. Clim. Res. 44: 151-166) [Erik Kjellström, Sweden]	Noted. Discussion on bias correction is removed from Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of model outputs.
9-1827	9	66	10			"Correcting known biases": in the RCM and/or GCM? Please clarify. [Government of United States of America]	Noted. Discussion on bias correction is removed from Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of model outputs.
9-1828	9	66	12	66	13	change "(e.g., de Elia and Cote, 2010) and (Boberg and Christensen, 2012; Christensen and Boberg, 2012)" to (e.g., de Elia and Cote, 2010; Boberg and Christensen, 2012 and Christensen and Boberg, 2012)" [SAMIR POKHREL, INDIA]	Editorial.
9-1829	9	66	13	66	13	add "Maraun, 2012" to the two listed references (already in reference list) Citation: Maraun, D. (2012): Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. GEOPHYSICAL RESEARCH LETTERS 39: L06706, DOI: 10.1029/2012GL051210 [Douglas Maraun, Germany]	Noted. Discussion on bias correction is removed from Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of model outputs.
9-1830	9	66	16	66	16	add a sentence along the lines "Räisänen and Räty (2012) cross validate different temperature bias correction methods in a pseudo reality and conclude that no best performing methods exist, differences between methods arise in particular for extreme temperatures, and uncertainties due to the chosen bias correction method should be accounted for." Citation: Räisänen, J., and Räty, O. (2012): Projections of daily mean temperature variability in the future: cross-validation tests with ENSEMBLES regional climate simulations. Climate Dynamics, published online, DOI 10.1007/s00382-012-1515-9. [Douglas Maraun, Germany]	Noted. Discussion on bias correction is removed from Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of model outputs.
9-1831	9	66	18			Add: The benefit of correcting of the sea surface temperature biases in the AR4 GCMs before using in downscaled simulations was highlighted over the Pacific in Nguyen et al., 2011. [Government of Australia]	Noted. Discussion on bias correction is removed from Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							model outputs.
9-1832	9	66	20	66	20	Performance-based ranking/weighting is just another form of stationarity assumption, but calibrated on the large scales and therefore a less robust assumption. See Reifen and Toumi, 2009, GRL 36 doi:10.1029/2009GL038082 for a similar discussion. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted. Performance-based weighting of regional climate models is explorative. The citation mentioned is not added as it concerns global models.
9-1833	9	66	25	66	27	"Coppola et al. (2010), Kjellström et al. (2010) and Sobolowski and Pavelesky (2012) demonstrated, for two different regions, that weighted sets of RCMs outperformed sets without weighting for both temperature and precipitation." This statement is not clear. How were sets of RCMs "weighted"? The following sentence talks about "skill-based weighting" which could be interpreted to mean that individual models were calibrated against local observed data where these were available. Is that what is meant here? [Government of Canada]	Taken into account. The weights used by Coppola et al and Kjellström et al were some of those reported by Christensen et al. cited earlier in the paragraph. The text has been clarified, and is now contained in Section 9.6.3, following some restructuring of the overall Section 9.6.
9-1834	9	66	27	66	27	Pls change "(Wehner, 2012)" to "Wehner, 2012" [HASIBUR RAHAMAN, India]	Editorial
9-1835	9	66	30	68	43	Overstructured and repetitive; skill, value added, uncertainties, performance are not independent enough. Please amalgamate and improve the relevance to the report, move towards a more integrated view! [Peter Braesicke, United Kingdom]	Taken into account. There has been general restructuring of the sections 9.6.2-9.6.6 in the Final Draft, towards a more consistent flow of presentation.
9-1836	9	66	30			please rename the subsection title to: 'Value added through RCM' or move this subsection into the subsection 9.6.2.2 If you want to keep this subsection as a general subsection add something about Sds. [Frank Kreienkamp, Germany]	Taken into account.
9-1837	9	66	30			Running regional models at higher resolution for all the major land regions (and a bit of the surrounding seas) of the Earth is probably less efficient in itself than running a single global model at higher resolution. Added value comes in part from being able to run for one or two regions at very much higher resolution than is possible globally. But perhaps the major added value is that regional climate modelling opens up the field to a wider set of modellers, who may bring specific experience of regional phenomena to the processes of model development and evaluation. This perhaps could be brought out in this section. [Adrian Simmons, United Kingdom]	Noted. The computational efficiency issue is noted, but it is not relevant in this context, as is also suggested in the comment. Global climate models are seldom run (due to computational constraints) at the resolution of RCMs. RCMs are also often used in region-specific research and, as suggested in the comment, by groups without the capacity to run global models.
9-1838	9	66	32	67	3	The section remains not completely solid yet. I would particularly suggest the statement at lines 32-33 be followed by added value definition given at lines 46-47. This will support a logical flow. [IGOR SHKOLNIK, RUSSIAN FEDERATION]	Accepted. Text modified.
9-1839	9	66	35	66	38	Consider taking the above into account in the 1st paragraph of the section, which I find too negative regarding the possibility of improvements in large scales. I believe that none of the tests reporting on degradations in large scales were done driving RCMs with climate forecasts and verifying against analyses or reanalyses. [Fedor Mesinger, United States, and Serbia]	Taken into account. The text was amended, including additional citations of recent work.
9-1840	9	66	35	66	38	If upscale interactions are important, is it reasonable to run RCMs without this coupling? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted. It is not evident whether RCMs should be allowed to modify the large-scales within their domain, if such scales are already representatively simulated by the global model (cf. Section 9.6.1.1). When upscale interaction affects the driving GCM (in case of two-scale coupling set-up), it is outside the regional domain.
9-1841	9	66	47	67	3	These skill metrics (I assume) are all for climatologies, not predictions of change. Given the use to which RCMs are being put it is essential that predicted changes in RCMs be more skillful than the changes predicted by the driving GCM. This needs to be explicitly discussed. [Gavin Schmidt, United States of America]	Noted. The assessment of climate models' skill, for both global and regional models, when it comes to climate projections remains elusive per definition. The role of RCMs is not least to address aspects that are not at the "scale" remit of global models.
9-1842	9	66	47			Another supporting reference is Otte et al., J. Climate, 2012, who showed improvements in mesoscale spatial variability. [Government of United States of America]	Taken into account. Cited now in the section 9.6.5 in conjunction with the discussion of nudging.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1843	9	66	53	66	53	The term "cutoff low" is not clear. Consider explaining or using different terminology. [Government of Canada]	Rejected. "Cutoff low" is established meteorological terminology and refers to a specific phenomenon: "A closed upper-level low which has become completely displaced (cut off) from basic westerly current, and moves independently of that current".
9-1844	9	66	53	66	54	Consider inserting two or three sentences something like: "In most cases researchers have by experiment design looked at value added at smaller scales, this being the objective of RCMs. But experiments reported in Veljovic et al. (2010) and Mesinger et al. (2012) demonstrate that it is possible to achieve value added also at the largest scales, upper tropospheric jet stream winds. This ought to lead to greater improvements in small scales than if jet steam winds were to stay the same as in the driver data." [Fedor Mesinger, United States, and Serbia]	Taken into account. We noted that this aspect has not been explored as much as value added on small-scale phenomena. Added a citation to Diaconescu and Laprise who build on the findings of Veljovic et al. and Mesinger et al., in a study with an RCM.
9-1845	9	66	53			Please add the Atlantic basin TCs to this list (Knutson et al. 2007; Bender et al. 2010). References: Knutson, T R., J J Sirutis, S T Garner, I M Held, and R E Tuleya, 2007: Simulation of the Recent Multidecadal Increase of Atlantic Hurricane Activity Using an 18-km-Grid Regional Model. Bulletin of the American Meteorological Society, 88(10), doi:10.1175/BAMS-88-10-1549. [Government of United States of America]	Indeed, there are yet additional papers and examples that could be cited. The Bender et al. was cited in te Final Draft.
9-1846	9	66				Reference not in the SOD: [Fedor Mesinger, United States, and Serbia]	Noted. Understood to belong together with the comment 9-1844 and 9-1847 (see the responses to these).
9-1847	9	66				Mesinger, F., K. Veljovic, M. J. Fennessy, and E. L. Altshuler, 2012: Value added in regional climate modeling: Should one aim to improve on the large scales as well? In "Climate Change: Inferences from Paleoclimate and Regional Aspects", Ed. by A. Berger, F. Mesinger, and Dj. Sijacki, Springer, 201-214 (doi:10.1007/978-3-7091-0973-1_15). [Fedor Mesinger, United States, and Serbia]	Noted. Belongs together with the comment 9-1844. See the relevant response.
9-1848	9	67	7	67	8	Structural uncertainty also arises from the selection of SD method(s) as well as the predictors. [Government of Australia]	Taken into account. The structural uncertainty is inherent in climate modelling and SD.
9-1849	9	67	7	67	42	It may be appropriate for the authors to discuss the issue that statistical predictors may not capture the key relationship for climate change even though they capture it for interannual variability. For an example, see Vecchi et al. 2008. Reference: Vecchi, Gabriel A., K L Swanson, and Brian J Soden, November 2008: Whither hurricane activity. Science, 322(5902), DOI:10.1126/science.1164396. [Government of United States of America]	Noted. The reference has been cited in a subsection 9.6.2.1 before as it is relevant to evaluation of SD through physical process. See also Comment 9-1860.
9-1850	9	67	12	67	25	The discussion of the use of nudging for regional climate modeling is very negative in this section. While it is still not universally accepted, several published papers support its use for regional climate modeling, especially when consistency with the driving fields is desired, as in environmental impact assessments. A counterpoint to L18-20 is that Otte et al. (J. Climate, 2012) nudging improved the representation of the means of 2-m temperature and precipitation; variability of temperature, water vapor mixing ratio, and winds aloft; and extremes of 2-m temperature and precipitation using a 20-year reanalysis-driven study. When nudging was not used, inaccurate extremes were produced by the RCM in that study. In addition, Bowden et al. (Clim. Dyn., 2012) used the same period as Otte et al., and they showed that nudging improved the representation of the large-scale circulation--specifically the Bermuda high, which strongly influences summertime weather and air quality in the eastern United States. We urge the authors to include published evidence of the benefits of using nudging for regional climate modeling, as the weight of evidence has increased substantially since AR4. [Government of United States of America]	Accepted. The key assumption behind the use of spectral nudging is that more is expected to be gained than to be lost compared to RCMs with unconstrained large scales, a view not shared by all, which is contained in the literature and the references in this section. We have balanced the text further, including citing Otte et al. as an example. .
9-1851	9	67	13	67	15	I strongly disapprove of the concluding part of the sentence of lines 13-15. Why should RCMs be blamed of being more prone to the growth of systematic errors than the driver data? This makes no sense. Presumably we are talking about climate change projections, aren't we? Errors grow in both climate change projection runs and in the nested RCMs. But various forcings presumably make this error grows constrained so as to have the result still contain the message. [Fedor Mesinger, United States, and Serbia]	Accepted. "Systematic errors" is replaced by "inconsistencies" which refers to differences between the data that underlies the boundary conditions and the RCM result.
9-1852	9	67	15	67	20	In Section 9.6.5, 2nd paragraph, there is a very brief discussion of the use of spectral nudging. The text cites a few studies that have shown detrimental effects of using nudging, but there have been numerous other	Accepted. The text has been amended.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						studies showing more positive results. For example, spectral nudging improves representation of climate downscaled from global reanalyses, including temperature and precipitation extremes (Otte et al., J Clim 2012), large-scale circulation patterns (Bowden et al., Clim Dyn. 2012), and tropical cyclones (Feser and Barcikowska, Env Res Lett 2012). One criticism of nudging, as cited in the text (Veljovic, 2010), is that a regional climate model can in some cases improve on the large scales in the driving fields. Veljovic et al. (2010) used for its lateral boundary conditions a free-running AGCM. Unlike a global reanalysis with assimilated observational data, an AGCM is subject to significant drift over the course of a simulation. The conclusion that the RCM was able to improve on the large-scale as represented by the AGCM may not hold if the RCM had been driven by more accurate coarse-scale fields. Perhaps it would be sufficient to add a caveat that the desirability of using spectral nudging may depend on the particular application (e.g., Bowden et al., Clim. Dyn. 2012; Rummukainen, 2010; Giorgi, J. Phys. IV, 2006). [Government of United States of America]	
9-1853	9	67	19	67	20	"For example, it may lead to deterioration of features such as precipitation extremes (Alexandru et al., 2009; Kawazoe and Gutowski, 2012). Also, while the smaller scale features of precipitation improved in larger domains, the time correlation of precipitation decreased." Assuming the second sentence is related to the first one, suggest rewording: "For example, it may lead to deterioration of features such as precipitation extremes (Alexandru et al., 2009; Kawazoe and Gutowski, 2012). These authors [if only one of the two studies, name them!] also found that smaller scale features of precipitation would improve in larger domains, but the time correlation of precipitation was degraded." [Government of Canada]	Taken into account. The text has been revised, but also shortened in the continued writing process.
9-1854	9	67	20			"time correlation of precipitation": correlation of precipitation with what? [Government of United States of America]	Noted. The statement does not appear in the revised (shortened) text.
9-1855	9	67	22	67	25	Should we mention here the effects that are found negligible at climatic scales? The sources of RCM errors and uncertainties are clearly outlined above (same page, see lines 8-10). [IGOR SHKOLNIK, RUSSIAN FEDERATION]	Accepted.
9-1856	9	67	29	67	29	add after "elevation": "and Maraun et al. (2012) evaluated regional climate models according to their ability to correctly represent the relationship between the large scale atmospheric circulation and local precipitation extremes". Citation: Maraun, D.; Osborn, T. J.; Rust, H. W. (2012): The influence of synoptic airflow on UK daily precipitation extremes. Part II: regional climate model and E-OBS data validation. CLIMATE DYNAMICS 39(1-2): 287-301, DOI: 10.1007/s00382-011-1176-0 [Douglas Maraun, Germany]	Noted. This section has been shortened and focused on processes that couple different climate variables/process-descriptions, rather than different spatial scales etc.
9-1857	9	67	44	68	43	This may not always be true over some specific regions, e.g. East Asia dominated by monsoon climate. As reported by Gao et al. (2008, 2012), the high resolution RCM not only introduces a fine scale topographically-induced structure in the climate change signal, but it also simulates some significantly different change patterns compared to the driving GCM, mostly due to the stronger RCM topographic forcing and the resulting circulation changes. [Xuejie Gao, China]	Taken into account. Gao (2012) was cited in 9.6.4 in conjunction with the discussion on added value.
9-1858	9	67	46			Is '... past performance..' a good wording? Is a '... reproduction of observed conditions ...' better? [Frank Kreienkamp, Germany]	Taken into account. Reworded.
9-1859	9	67	49	67	53	Are the stationarity hypothesis for SDs and the stationarity of biases coupled? I think not. Please separate those statements more clearly. [Frank Kreienkamp, Germany]	Noted. Stationarity of biases is no longer in this part (the bias correction it no more discussed in this part)
9-1860	9	67	51			Please add Vecchi et al. 2008 cite here. Reference: Vecchi, Gabriel A., K L Swanson, and Brian J Soden, November 2008: Whither hurricane activity. Science, 322(5902), DOI:10.1126/science.1164396. [Government of United States of America]	Accepted.
9-1861	9	67	55	68	11	The discussion should include a reference to Køltzow et al, (2011), which explored the relative importance of the quality of the oceanic surface forcing (SST and sea-ice) and the lateral boundary condition data as a function of the size of the integration domain. This was done by switching ocean-surface data and large-scale atmospheric data from different models when provided to the RCM. With Norway as the scope for the results, changes in the lateral boundary data and the size of the integration domain turned out to be more important than the quality of the ocean surface data. Køltzow, M.A.Ø., T. Iversen, J. E. Haugen (2011) The Importance of Lateral Boundaries, Surface Forcing and Choice of Domain Size for Dynamical Downscaling of Global Climate Simulations. Atmosphere 2011, 2, 67-95; doi:10.3390/atmos2020067. [Trond	Accepted. Cited in 9.6.5.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Iversen, United Kingdom of Great Britain & Northern Ireland]	
9-1862	9	67	55	68	43	Discussion of biases in RCMs and their causes. Needs emphasising as is important for decision makers. [European Union]	Noted.
9-1863	9	67	55			What about the stationarity hypothesis for all parametrisations from RCMs? Have I missed something ? If not please add. [Frank Kreienkamp, Germany]	Noted. Climate model parameterisations (both for global and regional climate models) are based on physical understanding of the underlying processes and phenomena, rather than statistical fits between different sets of data.
9-1864	9	68	5			Reference is made to RCMs tending to dry out the soil too effectively at high temperatures. Is this behaviour specific to RCMs, or does it apply to climate models in general? If it is behaviour specific to RCMs, some discussion of the possible reason should be given. [Adrian Simmons, United Kingdom]	Noted. The results showed (cf. cited literature, Figure) that the tendency to increased warm bias with higher temperature is common to both regional and global models.
9-1865	9	68	8			"ENSEMBLES": Is there a need to use upper case here. If this is referring to a particular project, please add a reference. [Government of United States of America]	Taken into account. Indeed, ENSEMBLES is a specific project, but does not necessarily need to be mentioned in the text here. "ENSEMBLES" was removed.
9-1866	9	68	20	68	20	Pls give gap after km [HASIBUR RAHAMAN, India]	Editorial
9-1867	9	68	25			Similarly, Nguyen et al. (2011) showed the climate change signal from six downscaled GCMs using one RCM produced broadly similar climate change signals, implying the dominant climate change signal (in the Pacific) was primarily coming through changes in ocean temperatures. [Government of Australia]	Noted. The assessment nature of the AR5 precludes from citing all papers.
9-1868	9	68	31	68	31	Pls give gap after °C [HASIBUR RAHAMAN, India]	Editorial
9-1869	9	68	33			In addition, Bennet et al. 2012 showed the benefit of bias correcting the outputs of down-scaled data for hydrological modelling. Reference: Bennett, J. C., Ling, F. L. N., Post, D. A., Grose, M. R., Corney, S. P., Graham, B., Holz, G. K., Katzfey, J. J., and Bindoff, N. L., 2012: High-resolution projections of surface water availability for Tasmania, Australia, Hydrol. Earth Syst. Sci., 16, 1287-1303, doi:10.5194/hess-16-1287-2012. [Government of Australia]	Noted. Discussion on bias correction is removed from Chapter 9, which is in line with its focus on evaluation of downscaling methods, whereas bias correction is a step done after in order to enhance the quality of model outputs.
9-1870	9	68	34	68	34	Chapter 9: Section 9.6 (Downscaling and Simulation of Regional Scale Climate). It is recommended that the last paragraph of this Section should become a new subsection 9.6.7 titled Summary. [Government of Spain]	Rejected. We appreciate the suggestion, but it would introduce inconsistency across the Chapter 9 subsections. Summary statements are provided in the Executive Summary.
9-1871	9	68	38	68	38	This is not justified. There may be high evidence for added value in producing climatologies, but there is no evidence described that demonstrates added value for climate change experiments (or even for ENSO fingerprints or the response to Pinatubo). Surely this literature must exist? [Gavin Schmidt, United States of America]	Taken into account. These statements wrap up earlier sections of 9.6 and were in the Final Draft moved to 9.6.4. Otherwise, Chapter 9 is about evaluation of climate models, and not about climate change projections.
9-1872	9	68	41			High agreement and medium evidence should be quantified. Please associate a confidence statement with this statement (see figure 1.12). [Government of United States of America]	Taken into account. The assessment gives "high confidence". (The value-added assessment summary was moved to section 9.6.4.)
9-1873	9	68	42			Limited evidence? How about the Emanuel et al 2008 and Knutson et al. 2007 references mentioned earlier? These seem to show added value of downscaling extremes, at least for the case of tropical cyclones. There is also some new work on statistical models for tornado activity from environmental parameters which could also be mentioned in the chapter. Reference: Tippett, M. K., A. H. Sobel, and S. J. Camargo (2012), Association of U.S. tornado occurrence with monthly environmental parameters, Geophys. Res. Lett., 39, L02801, doi:10.1029/2011GL050368. [Government of United States of America]	Noted. "Limited evidence" implied (cf. the AR5 confidence language) that the number of studies is limited (but that there are some), which is the result of the assessment. However, we agree on the number of works to be relatively extensive (more are now cited in 9.6), which can support "medium evidence".
9-1874	9	68	45	69	4	Section 9.7 : I find the association of the two topics (« Understanding model performance » and « climate sensitivity ») very confusing (and the lines 50-52 are not helpful with that regard, and the reading of the whole	Taken into account - the first topic, which was actually a synthesis of earlier sections, is shortened and put

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						section neither). I would suggest to include the information about the first topic within the previous relevant sections (with some references to support the text), and to make the part on « Climate sensitivity » separate or to redistribute it across several other chapters (chapters 7 and 12 in particular). [Sandrine BONY, France]	into a box. The section is devoted to the climate sensitivity issue.
9-1875	9	68	45	70	52	9.7 When relating model performance to climate sensitivity, we recommend that the authors consider a discussion of how cloud and water vapor errors in simulating the current climate transforms into the climate projection uncertainty. In particular, recent studies (e.g. Jiang et al. 2012) have shown that the largest model spread and model error in simulating cloud and water vapor are in the upper troposphere. Study by Su et al. (2008) show that the upper troposphere cloud contributes about 3W/m2 in net cloud radiative forcing. Comparing to ~-20W/m2 in total net cloud radiative forcing, the upper troposphere cloud contribute about 15% total cloud climate sensitivity. For water vapor, the upper troposphere water vapor exerts the strongest greenhouse gas effect (Soden et al., 2005; Su et al. 2006a). Thus the largest CMIP5 model spread in simulating upper troposphere cloud and water vapor is a concern that we recommend is addressed here. Reference: Soden, B. J., D. L. Jackson, V. Ramaswamy, M. D. Schwarzkopf, and X. Huang, The radiative signature of upper tropospheric moistening, Science, 310(5749), 841844, doi:10.1126/science.1115602, 2005. Su, H., J.H. Jiang, Y. Gu, J.D. Neelin, B.H. Kahn, D. Feldman, Y.L. Yung, J.W. Waters, N.J. Livesey, M.L. Santee, and W.G. Read, "Variations of tropical upper tropospheric clouds with sea surface temperature and implications for radiative effects," J. Geophys. Res. 113, D10211, doi:10.1029/2007JD009624, 2008. Su, H., W.G. Read, J.H. Jiang, J.W. Waters, D.L. Wu, and E.J. Fetzer, "Enhanced positive water vapor feedback associated with tropical deep convection: New evidence from Aura MLS," Geophys. Res. Lett. 33, L05709, doi:10.1029/2005GL025505, 2006a. [Government of United States of America]	Noted - the cloud issue is covered in Chapter 7.
9-1876	9	68	45	77	43	Chapter 9: Section 9.7 (Understanding Model Performance and Climate Sensitivity). It is recommended to include a Summary subsection at the end of this Section for a better understanding of its contents. [Government of Spain]	Taken into account - the first topic, which was actually a synthesis of earlier sections, is shortened and put into a box - thus easier to read.
9-1877	9	68	49	68	50	I do not think that there is consensus that inter-model spread represents a minimum of model uncertainty. This discussion is partially tackled in chapter 11page 26, line 4-16 and treted with more detail in Chapter 12 12.2.2. [Ramon de Elia, Canada]	Noted - that part of the text is deleted due to reorganization.
9-1878	9	68	54	69	4	I could not understand the classification of model results error into the two kinds talked about here, basically for being puzzled as to what is meant by "true model solution". I looked therefore into the two papers cited. The first one refers to "modeling error for certain classes of problems in linear and nonlinear mechanics". Essential components of climate models are not problems in mechanics. With all due respect, I wonder if the two papers cited are helpful for understanding of climate model errors; it seems to me that they are not. I could think of other classification of climate model errors which would be more helpful. My suggestion is to delete this paragraph. [Fedor Mesinger, United States, and Serbia]	Noted - that part of the text is deleted due to reorganization.
9-1879	9	68		78		As a policy maker I would like to see a summary of this section which is of very technical nature: What are the main remaining gaps/deficiencies of existing state-of-the-art climate models and how likely are those to lead to a substantial over- or underestimation of key indicators of climate change. Also I would want to know which indicators of climate change are systematically inaccessible to current model approaches? [Jochen Harnisch, Germany]	Taken into account - the first topic, which was actually a synthesis of earlier sections, is shortened and put into a box - thus easier to read.
9-1880	9	69	2	69	2	"When we assess the causes of errors in current climate models..." Consider clarifying whether this mean climate models that are "current" as in "state of art", or climate models that are being used to simulate "current climate" as in "contemporary climate" rather than future climate? Consider clarifying. [Government of Canada]	Noted - that part of the text is deleted due to reorganization.
9-1881	9	69	4	70	52	This entire sections cites no references at all and reads almost like a personal review on the subject. There are no explicit links to other sections of the chapter. It needs to include references throughout to support statements and would benefit from relating to the range of sensitivities in the following section. [Government of Australia]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1882	9	69	6	69	52	Section 9.7.1, Understanding Model Performance: this section currently is not very well linked to the previous sections, and we therefore suggest to add specific references to the relevant subsections to better connect cross sections. In addition, we note that there is some repetition with information that has come already in earlier sections, and this should be avoided so that the section is condensed as much as possible. It might	Accepted - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						actually be worth considering gathering this material within a Box, given that this section does not contain anything in the way of assessed findings. The lack of cited literature in this section is also very obvious. [Thomas Stocker/ WGI TSU, Switzerland]	
9-1883	9	69	6	69	57	9.7.1 This section gives an excellent accounting of the sources of uncertainty in various components. However, not a single reference is cited. Much of the attributions/claims can also be found in previous sections, but not all of them. This section is not intended to summarize the related uncertainty attributions loosely described in previous sections either. It is supposed to be an important standalone section to provide insights for interested readers. Therefore, it will be very useful to cite some references for claims that are based on other publications and are not obvious in the context, to enlighten the readers. [Government of United States of America]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1884	9	69	6	70	52	This section lacks references. [Thierry Fichefet, Belgium]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1885	9	69	6	70	52	9.7.1 Compared to the other sections of Chapter 9, section 9.7.1 stands out as being totally devoid of references. Consider adding references to that section to parallel other sections. [Government of United States of America]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1886	9	69	6			This has already happened for RCMs, maybe merge [Peter Braesicke, United Kingdom]	Rejected - we decided to put all the issues particular to downscaling in Section 9.6.
9-1887	9	69	8	69	57	The whole section is reading like a text-book. It does not contain any reference. There are many statements which must be linked to actual literature! [Martin Dameris, Germany]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1888	9	69	8	69	57	Section 9.7.1.1 Please add that ENSO prediction is severely limited because despite the characteristics being recognised, the driver of ENSO is poorly understood. [John McLean, Australia]	Rejected - the section is now shortened and put into a box. As it should be very concise, there is no space for additional discussion.
9-1889	9	69	8			Strange there are no references here in the section. I think some things could (should) be referenced. [Andrew Gettelman, United States of America]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1890	9	69	10	69	57	There is quite a lot of repetition of points made earlier in this subsection 9.7.1.1, as issues in parameterizing processes is discussed earlier. The other subsections of 9.7.1 are quite short, so maybe 9.7.1.1 could be shortened. [Adrian Simmons, United Kingdom]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1891	9	69	10			the uncertainties are causing deviations; systematic errors and random errors (see earlier comments) - the scene setting for error should be done early on [Peter Braesicke, United Kingdom]	Rejected - we think the current structure is fine.
9-1892	9	69	15	69	27	A regime-dependent error diagnostic framework is put forward by Su et al., (2012), which decomposes model simulated cloud errors into the large-scale error, cloud parametrization error and co-variation. Using multiple large-scale regime parameters, the dominant source of error is identified (the cloud parametrization error) and different dependency of cloud simulations on large-scale parameters is shown. [Government of United States of America]	Noted - the cloud issue is covered in Chapter 7.
9-1893	9	69	15	69	44	We recommend adding references to these two paragraphs. [Government of United States of America]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1894	9	69	25	69	26	As in my comment (18), these two processes are not necessarily the same and should not be lumped together in this way. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Noted - that part of the text is deleted due to reorganization.
9-1895	9	69	26			Example of a reference: for 'drizzle bias': Stephens et al 2010 (Ref in ch 7) [Andrew Gettelman, United States of America]	Noted - that part of the text is deleted due to reorganization.
9-1896	9	69	27			add (self) reference [Peter Braesicke, United Kingdom]	Taken into account - the section was actually a

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1897	9	69	27			I might be helpful to give a definition of shallow and deep convection. Overall this section is extremely technical. Who is the reader you are focusing on. [Frank Kreienkamp, Germany]	Noted - that part of the text is deleted due to reorganization.
9-1898	9	69	31			I suggest adding the lack of proper overflows as a source of bias for intermediate and deep layer in the Labrador Sea [RYM MSADEK, United States of America]	Rejected - the section is now shortened and put into a box. As it should be very concise, there is no space for additional discussion.
9-1899	9	69	36	69	37	What is the basis for this statement? Even atmosphere-only models, driven by prescribed SST, produce insufficient marine stratocumulus (e.g., Gibson JK, Källberg P, Uppala S, Hernandez A, Nomura A, Serrano E (1997) ERA description. ECMWF reanalysis project report 1, ECMWF, Reading, UK, 72 pp). So the warm SST bias in coupled models is not caused by weak upwelling, but by to high solar radiation due to to few clouds. [Andreas Sterl, Netherlands]	Noted - that part of the text is deleted due to reorganization.
9-1900	9	69	36			does South America well up? Sounds strange to me [Frank Kreienkamp, Germany]	Noted - that part of the text is deleted due to reorganization.
9-1901	9	69	49			Please provide a reference for the improvement in sea-ice decline simulation due to improved melt pond representation. [Government of United States of America]	Noted - that part of the text is deleted due to reorganization.
9-1902	9	69	51		59	One weakness concerns the biogeochemical component for which different structures (or lackoff) have been included and might explain the spread between model simulations. This should be part of the "uncertainty" discussion. Moreover information given in Table 9.1 (p138) on ocean biogeochemical models is insufficient to estimate how comparable these model components are. [European Union]	Noted - Table 9.1 is already huge and we think it is difficult to include further details. Description papers of each models should be referred to if one really want to compare model components.
9-1903	9	69	51			Reference needed for this sentence. Particularly what is the reference for the importances of forest fires on long (multidecadal and longer) time scales? [Ronald Stouffer, United States of America]	Taken into account - an internal citation of Section 9.4.5.1 is added. Original sources can be found there.
9-1904	9	69	55			"...Atlantic Multidecadal Oscillation, although the extent of influence of anthropogenic aerosols vs natural variability (including natural variations is aerosols and dust) remains uncertain." [Government of United States of America]	Noted - that part of the text is deleted due to reorganization.
9-1905	9	69				section 9.7.1: give references [Barbara Früh, Germany]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1906	9	69				This section is interesting but the absence of any citations to support the statements and generalisations is a problem. [Steven Sherwood, Australia]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1907	9	70	2	70	52	Good material, most of this should come earlier to set the scene; 9.7.1.3: how are the phenomena linked? [Peter Braesicke, United Kingdom]	Rejected - we think the current structure is fine.
9-1908	9	70	2			Section 9.7.1.2: This should include mention of feedbacks between tropical SST biases and rainfall biases (in both directions, e.g. excessive rainfall causing cooler SSTs which then modulate excessive rainfall; SST biases affecting temperature gradients which then affect rainfall distributions, monsoon circulations, etc); refs including Bollasina and Nigam (2009; Clim. Dyn. 33, 1017-1032); Levine et al. (2012; Clim. Dyn. revised)), and also between climate mean state and vegetation distribution when DGVMs are included (e.g. Martin and Levine, 2012 (Earth Sys. Dyn. 3, 245-261)). [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Rejected - the section is now shortened and put into a box. As it should be very concise, there is no space for additional discussion.
9-1909	9	70	4			"Cause[s] of one model bias" [Andrew Wittenberg, United States of America]	Editorial.
9-1910	9	70	13	70	14	"A decreasing trend in September Arctic ice extent tends to be underestimated when sea ice thickness is overestimated". Can you please include a reference? Thanks. [François Massonnet, Belgium]	Taken into account - the section was actually a synthesis of earlier sections. It is now shortened and put into a box with relevant internal citations.
9-1911	9	70	17	70	28	Only improvements introduced by increasing resolution are mentioned. In fact increasing resolution can also	Noted - that is true. But, there is no space for

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						lead to a deterioration of some aspects of the simulation, see for example Delworth, Thomas L., Anthony Rosati, Whit G Anderson, Alistair Adcroft, Ventakramani Balaji, Rusty Benson, Keith W Dixon, Stephen M Griffies, H C Lee, Ronald C Pacanowski, Gabriel A Vecchi, Andrew T Wittenberg, Fanrong Zeng, and Rong Zhang, April 2012: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. Journal of Climate, 25(8), doi:10.1175/JCLI-D-11-00316.1. [Sonya Legg, United States of America]	additional discussion. The section is now shortened and put into a box. It should be very concise.
9-1912	9	70	17			Section 9.7.1.3: Relative words like "higher" and "improved" are not helpful in discussions of resolution as the impacts very often depend on what the resolution is before and after the increase, how this relates to the scale of the phenomena in question, and, for vertical resolution, at what height/depth in the atmosphere/ocean the change is made. It is essential that discussions of the impacts of resolution changes are specific about the actual changes being discussed. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Noted - we agree. But, here the space permits only short and general discussion. The section is now shortened and put into a box. It should be very concise.
9-1913	9	70	21			storm->storm [Reto Knutti, Switzerland]	Editorial.
9-1914	9	70	22			"tropical cyclone intensity and structure." [Government of United States of America]	Accepted.
9-1915	9	70	30			This section about uncertainty in observations should include a statement about the need for high-resolution high-quality data needed for evaluation of high-resolution regional (and eventually global) models. There are examples in the literature where observational data even in well-sampled regions like Europe shows potential errors both not just for extremes but also for long-term seasonal averages (see for instance p189-190 in Christensen et al. 2010 for examples). [Erik Kjellström, Sweden]	Rejected - the section is now shortened and put into a box. As it should be very concise, there is no space for additional discussion.
9-1916	9	70	37			"...extreme precipitation and tropical cyclone activity trends." [Government of United States of America]	Rejected - the tropical cyclone activity trends are not included in the subjects of model evaluation in this Chapter.
9-1917	9	70	42	70	43	Newly available vertical resolved global cloud and water vapor information, such as those from the NASA A-Train satellites datasets (L'Ecuyer and Jiang, 2012), has highlighted difficulties in simulating clouds and water vapor fields in models, especially in the upper troposphere (Jiang et al. 2010;2012). References: L'Ecuyer, T.S., and J.H. Jiang, "Touring the atmosphere aboard the A-Train," Physics Today 63, 7, 36-41, 2010. Jiang, J.H., et al., "Five-year (2004-2009) Observations of Upper Tropospheric Water Vapor and Cloud Ice from MLS and Comparisons with GEOS-5 analyses," J. Geophys. Res. 115, D15103, doi:10.1029/2009JD013256, 2010. Jiang, J.H., et al., "Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations," J. Geophys. Res. 117, D1410, 24 PP, 10.1029/2011JD017237, July 2012. [Government of United States of America]	Noted - the cloud issue is covered in Chapter 7.
9-1918	9	70	43			We suggest appending this sentence in the following way: "However, new or improved observing capabilities can also lead to homogeneity problems when assessing trends or comparing modeled and observed trends." [Government of United States of America]	Rejected - the section is now shortened and put into a box. As it should be very concise, there is no space for additional discussion.
9-1919	9	70	45	70	52	It must be stated that "boundary conditions" (... how the models are forced) also include "emissions". In particular in models containing chemistry (feedbacks) or at least effects of volcanic eruptions (aerosols) it is another source of uncertainty. [Martin Dameris, Germany]	Taken into account - though it doesn't have to be a complete list, 'aerosols' is changed to 'aerosols emissions', considering that many AOGCMs specify emissions for aerosols (and their precursors) instead of aerosols concentrations.
9-1920	9	70	45			Here again the inherent limitation of comparing short records is not discussed. Even for a perfect model the agreement would not be perfect in many variables because internal variability is random. For some fields this contribution is large. [Reto Knutti, Switzerland]	Noted - the intention is that it is implied by 'insufficient length'. As the section is now put into a Box and should be very brief, we give up elaborating it further here.
9-1921	9	70	47	70	48	"Models with prescribed or simulated stratospheric ozone perform better than fixed ozone". It is confusing to use the words "perform better" since the whole chapter stresses that different metrics for different processes can be set up. I would be very surprised of such a strong conclusion (in which case I would like the authors to include a reference). I may agree that some climate features are better simulated with prescribed/simulated ozone, but the authors should temper this statement. For Antarctic sea ice and CMIP5 models, the trends in winter and summer sea ice extent over the past decades are undistinguishable between the models with and without interactive chemistry (Zunz et al., 2012). [François Massonnet, Belgium]	Noted - that part of the text is deleted due to reorganization.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1922	9	70	51	70	51	Again, mention here that guidelines for correct modelling are being developed (see my previous comment for references) [Juan Blanco, Spain]	Rejected - the section is now shortened and put into a box. As it should be very concise, there is no space for additional discussion.
9-1923	9	70	55	77	43	Section 9.7 : Parts of this section significantly overlap -not completely- with those of Chapter 7 (e.g. 9.7.2.2.1, 9.7.2.2.3 on water vapor, lapse-rate and cloud feedbacks, 9.7.2.1 on methodological aspects of the diagnostic of forcings, adjustments and climate sensitivity, 9.7.2.2.3 on the diagnostic of cloud feedbacks and observational constraints). If they moved to chapter 7, they would nicely complete the assessment of chapter 7 (provided that chapter 7 can get longer!). On the other hand, the physical interpretation of the multi-model spread (in MME, not PPE which is better discussed in chapter 9) of climate sensitivity and cloud feedbacks is as little developed in chapter 9 than it is in chapter 7 (SW cloud feedbacks are mentioned as the primary source of spread, but there is no assessment of the origin of this spread in terms of regions, cloud types or regimes, local or remote source, etc). For instance (among others), the study by Vial et al. (Clim. Dyn, 2012) could help ; Brient and Bony (GRL, 2012) propose a physical mechanism through which biases in the simulation of low-cloud radiative effects in the current climate might affect climate change low-cloud feedback, and suggest that the quasi-systematic 'too few, too bright low-cloud problem' of CMIP models (e.g. Nam et al, GRL, 2012) might amplify the spread of cloud feedbacks in climate change. [Sandrine BONY, France]	Accepted -- overlap with chapter 7 has been minimized by eliminating duplicative material.
9-1924	9	70	55			Suggest to refer to Chapter 12, Box 12.2; ensure consistency in assessment with Ch12 [Thomas Stocker/ WGI TSU, Switzerland]	Agreed - reference to chapter 12 box 12.2 added, consistency checked against Ch 12.
9-1925	9	70	56			Suggest to move the statement from page 77, lines 40-43 ("An overall assessment of climate sensitivity and transient response is given in....) to here. We think it's better to mention upfront that Chapter 12, Box 12.2 is where all the information from various chapters will be compiled. [Thomas Stocker/ WGI TSU, Switzerland]	Agreed -- statement on page 77 moved to page 70 (SOD page #ing)
9-1926	9	71	4	71	5	This definition is incomplete since it does not account for the time scale on which equilibrium is found nor on the components included. Models with dynamics vegetation, or land ice components for instance will have a very different ECS than identical models that have fixed veg and LI. If you mean the Charney sensitivity only including fast atmospheric processes (but no ozone or aerosols, or ocean dynamics), say so. Otherwise this definition is meaningless in terms of comparing different models. The Paleosens paper (Nature, 2012) has some interesting things to say about this issue. [Gavin Schmidt, United States of America]	Noted -- the definition of ECS given here is consistent with its (abbreviated) definitions in the literature
9-1927	9	71	5	71	5	"(ECS) represents the single most important measure of climate response" This is a debatable, non-consensus statement. Many argue TCR is more important because the ECS will not be approached in the coming century. And TCR does not necessarily scale with ECS. Either way, determining the "most important measure of climate response" is not solely in the realm of the physical basis for climate change (WG1). [Timothy Merlis, United States of America]	Accepted -- the characterization that "most important measure" has been deleted.
9-1928	9	71	8	71	8	Pls give gap after °C [HASIBUR RAHAMAN, India]	Editorial
9-1929	9	71	18			The disagreement can apparently be much larger than this. For example, the table 9.4 lists the ECS of GFDL AM2.1 (via. ESM2G or ESM2M) as 2.4C, while the published value based on a slab model is 3.4C (Stouffer et al J Clim, 2006). This is a 30% difference! [Government of United States of America]	Taken into account -- combined with comment 9-2305.
9-1930	9	71	18			"Yu, 2003; Danabasoglu and Gent, 2009; Li et al., 2012a)", though individual feedbacks (water vapor and lapse rate) are larger in AOGCMs than SOM-AGCMs (Shell, 2013). [Shell, Karen M., 2012: Consistent Differences in Climate Feedbacks Between Atmosphere-Ocean GCMs and Atmospheric GCMs with Slab-Ocean Models, accepted pending minor revisions J. Clim., http://people.oregonstate.edu/~shellk/AOGCM_AGCM_jclim.pdf ; To be re-submitted early Dec.] [Karen Shell, United States of America]	Noted
9-1931	9	71	19	71	19	Keith Williams et al (J Climate, 2008) also made such a comparison, finding that slab and AOGCM ECS could differ by 0.5 K. [Jonathan Gregory, United Kingdom]	Accepted -- The reference to Williams et al has been added.
9-1932	9	71	22	71	22	Pls change "Gregory et al (2004)" to "Gregory et al., (2004)" [HASIBUR RAHAMAN, India]	Noted.--It's automatically generated by the IPCC endnote style.
9-1933	9	71	26	71	32	This is difficult to visualize. Consider including a diagram which shows how F and ECS can be obtained from a	Noted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						graph of the GCM output of radiative imbalance against ΔT . [Government of Canada]	
9-1934	9	71	26			Often recent studies use ' ΔR ' instead of ' N ' and ' λ ' instead of ' α ' in their notation. [Thorsten Mauritsen, Germany]	Noted - maerial deleted
9-1935	9	71	28	71	29	Remove 'radiative' and just use 'adjusted forcing'. It would be useful to refer to Chapter 8 in addition to 7 for AF. [Gunnar Myhre, Norway]	Noted - maerial deleted
9-1936	9	71	32	71	32	The forcing in the extrapolated value of N , when $T=0$. This is important, as there can be significant uncertainties because of this extrapolation. [Government of Australia]	Noted - maerial deleted
9-1937	9	71	32	71	32	It might be helpful to note that F is the adjusted forcing (see ch8) when diagnosed by this method. [Jonathan Gregory, United Kingdom]	Noted - maerial deleted
9-1938	9	71	35			The logarithmic dependency of direct radiative from CO_2 dates back to Arrhenius (1896). Pierrehumbert (2011) went at great lengths to explain it too. The references could be added, and the level of confidence raised. However, point out that direct radiative forcing is different from adjusted forcing, cf. next comment. [Thorsten Mauritsen, Germany]	Noted - maerial deleted
9-1939	9	71	35			Maybe the reference to Manabe and Bryan(1985) could be omitted at the end of this line, as the paper is discussed in a qualified way in the following paragraph. [Adrian Simmons, United Kingdom]	Noted - maerial deleted
9-1940	9	71	36			At this point it would be worthwhile outlining (since this is dealt with in other chapters) how we interpret forcing and feedback, and how that relates to AR4. AR4 assumed that CO_2 forcing consisted of direct forcing and stratospheric adjustment, and that all change due to e.g. clouds is part of the respective feedback. The text on page 75 lines 9-19 does some of this, and could effectively be moved up, after some modification. [Thorsten Mauritsen, Germany]	Noted - maerial deleted
9-1941	9	71	37	71	37	No it doesn't. The Andrews et al definition is just another definition, and there are clear differences in both the Andrews inferred forcing and ECS than the actual adjusted forcing and the eventual temperature. [Gavin Schmidt, United States of America]	Noted - maerial deleted
9-1942	9	71	39	71	48	Non-linear behavior in the abrupt4x CO_2 simulations is more the rule than an exception among CMIP5 models, hence the use of 'not always valid' is a too weak statement, in my opinion. Some studies argue that the non-linearity is due to decadal timescale adjustments in the forcing due to deep ocean heat uptake modulating the distribution of SST warming (Senior and Mitchell 2000, Williams 2008, Winton et al. 2010). In MPI-ESM-LR, however, the behavior is better explained by a systematic state-dependent positive shift in the feedback factor at warming temperatures, because stand-alone atmospheric simulations with uniformly increased SST's and coupled to a mixed-layer ocean exhibits the same type of non-linearity (Block and Mauritsen 2012). The argumentation with Li et al. (2012a) is a little imprecise because focus is on temperature change, not adjusted forcing, and so one cannot rule out state-dependent feedback. Further it is not clear what aspect is accurate to within 10 percent. Block and Mauritsen (2012) do find a slight increase in adjusted forcing in 4x CO_2 relative to 2x CO_2 (determined using fixed SST runs), which is consistent with a rise in 'traditional' CO_2 forcing (direct + stratospheric adjustment) found Colman and McAvaney (2009) in another model. [Thorsten Mauritsen, Germany]	Noted - maerial deleted
9-1943	9	71	39			Senior and Mitchell 2000 GRL and Gregory GRL 2004 were the first to show that α is not constant in some models. [Reto Knutti, Switzerland]	Noted - maerial deleted
9-1944	9	71	43	71	46	A strictly logarithmic response is certainly not expected. This is because climate feedbacks change with temperature. So even if forcing were to scale exactly (which it doesn't) the climate response certainly does not. This is discussed at length in Colman, R.A., and B.J. McAvaney, 2009: Climate feedbacks under a very broad range of forcing, Geophysical Research Letters, 36, L01702, doi: 10.1029/2008GL036268. [Government of Australia]	Noted - maerial deleted
9-1945	9	71	46	71	48	The Li et al (2012) study used a regression over a 1000 year period to achieve a 10% accurate ECR estimate. It should be noted that this is much longer than the 150 years used in the Andrews et al (2012) values that are cited in the table. [Government of United States of America]	Noted - maerial deleted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1946	9	71	46			Jonko et al. (2012) also found that climate sensitivity increases with larger forcings for CCSM. Climate sensitivity increases by 23% between the first and third CO2 doublings. [Jonko, Alexandra, Karen M. Shell, Benjamin M. Sanderson, and Gokhan Danabasoglu, 2012: Climate feedbacks in CCSM3 under changing CO2 forcing. Part II: Variation of climate feedbacks and sensitivity with forcing, J. Clim, in press.], though this result seems to be model-dependent. [Karen Shell, United States of America]	Noted - maerial deleted
9-1947	9	71	50	71	50	Pls change "Hansen et al (2005)" to "Hansen et al., (2005)" [HASIBUR RAHAMAN, India]	Noted - we can not change it because it's automatically generated.
9-1948	9	71	50			Here it might be useful to mention if the Hansen et al. (2005) based forcing estimates include correction for the small surface temperature change that does occur in these kinds of simulations. [Thorsten Mauritsen, Germany]	Noted - maerial deleted
9-1949	9	71	54	71	54	Andrews et al (2012, GRL) compare the results of the two methods for all CMIP5 models. [Jonathan Gregory, United Kingdom]	Noted - maerial deleted
9-1950	9	71	54	71	54	I believe the AR5 terminology is just "adjusted forcing" (AF), without "radiative". [Jonathan Gregory, United Kingdom]	Noted - maerial deleted
9-1951	9	71				Section 9.7.2.1: To repeat many of my comments above: the large range of climate sensitivity and the lack of improvement since CMIP3 is a very major problem and which needs more emphasis. It gets lost in a carefully worded paragraph in sub-sub-section 9.7.2.1 ... even though it is in the first paragraph of the sub-sub-section. Section 9.7.2 actually states many (all?) of the problems with the climate models. At this stage of development of climate change science it should have a much higher impact than it does here, (lost at the back of the filing cabinet). I really liked figure 9.44 and some aspects of the associated discussion. I think that the approach should help focus ideas on where to concentrate the research effort in future. It would be nice if this figure or something similar could be given a place in the technical summary. [David Webb, United Kingdom]	Noted - this section has been shortened to increase the clarity of its statements relative to CMIP5.
9-1952	9	72	2	72	4	TCR definition is very unclear. [Stephen Gaalema, United States of America]	Accepted -- definition of TCR has been clarified and reconciled with its definition in the glossary.
9-1953	9	72	4	72	8	Can you reconcile the statement here that ECS spread is the dominant source of TCR spread with the finding by Winton et al (2010, not cited) that ocean heat uptake efficacy has an equally large influence on TCR variation? Might the short regressions used to estimate the ECS's here give a spuriously strong correlation between TCR and ECS estimates since Williams et al (2008) and Winton et al (2010) find that many models exhibit increasing sensitivity over time? Reference: Winton, Michael, Ken Takahashi, Isaac M. Held, 2010: Importance of Ocean Heat Uptake Efficacy to Transient Climate Change. J. Climate, 23, 2333-2344. [Government of United States of America]	Accepted -- omission of ocean heat uptake efficiency as an explanation for the spread has been corrected.
9-1954	9	72	6			Fig 9.43a is not explained. There appears to be a small but interesting negative correlation between GMST and ECS. [Stephen Gaalema, United States of America]	Accepted -- Figure 9.43a is now called out and explained separately from Figure 9.43b.
9-1955	9	72	6			They may appear to be linearly correlated but we know the relationship is nonlinear, see e.g. chapter 10 in AR4 or Knutti et al. GRL 2005. The dependence of TCR on ECS weakens for high ECS. [Reto Knutti, Switzerland]	Accepted -- the departures from linearity are now noted in the text at this point.
9-1956	9	72	10			How do the findings in this paragraph relate to model performance? [Peter Braesicke, United Kingdom]	Accepted -- goal is to explain how well an emergent property, the climate sensitivity, can be explained in terms of feedbacks from individual processes that have been assessed in this chapter. This connection is now made explicit.
9-1957	9	72	20	72	21	Pls give gap after °C [HASIBUR RAHAMAN, India]	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1958	9	72	24	72	25	Why would ESMs be expected to change the range?. Carbon cycle (the main advance in ESMs) does not enter into the picture, as CO2 is doubled, then held fixed in these idealised experiments. Section needs clarity on this. [Government of Australia]	Accepted -- the fact that the c-cycle feedbacks would not affect the ECS in these idealized experiments is clarified.
9-1959	9	72	24	72	27	What does the "overall spread" reflect? If the ECS has a large spread even with simple models, this suggests that a very fundamental component/characteristic of the system is responsible for the wide variability (in which case additional resolution and process representation is unlikely to further constrain the ECS). [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted
9-1960	9	72	31	72	31	Not just conceivable, but demonstrated within a single model. See Colman, R.A., and B.J. McAvaney, 2009: Climate feedbacks under a very broad range of forcing, Geophysical Research Letters, 36, L01702, doi: 10.1029/2008GL036268 [Government of Australia]	Accepted -- reference to Colman and McAvaney added.
9-1961	9	72	31	72	34	You may not want to overgeneralize the conclusion about ECS as a function on global mean temperature based on Fig 9.43a, as the range of global temperature explored in the figure is considerably less than the projected changes over the next few centuries in high emission scenarios. [Government of United States of America]	Taken into account -- combined with comment 9-1964.
9-1962	9	72	31	72	34	Twisting this observation to be a way of enhancing confidence in models seems inappropriate. [Thorsten Mauritsen, Germany]	Taken into account -- combined with comment 9-1964.
9-1963	9	72	31	72	34	This is a strange sentence - our confidence is enhanced by a result which is against physical intuition? Perhaps instead of "enhancing confidence" this observation "suggests that ECS is less sensitive to errors in the time-mean than to other sources of uncertainty" [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Accepted -- wording altered as suggested.
9-1964	9	72	31			This appears to be the case in the narrow range of global temperature that CMIP5 is covering, but it is important to note that based on processes we would not expect this to hold for very strong warming (see the work by Boer and Yu). It also is not true for feedbacks at LGM (Hargreaves 2007 Climate of the Past). The current wording is a bit misleading and suggests that fundamentally the control bias doesn't matter, which is not the case (see the discussion above of the Boberg/Christensen 2012 paper). [Reto Knutti, Switzerland]	Accepted - caveats introduced to as suggested to minimize risk of misinterpretation of the figure.
9-1965	9	72	31			Lapse rate and cloud feedbacks could depend on temp as well [Karen Shell, United States of America]	Agreed -- possibility of temperature dependence of lapse rate and cloud feedbacks noted here.
9-1966	9	72	34	72	34	This seems the first mention (in the sections) of 'errors in the time-mean', of global mean temperature. These are also evident from Fig. 9.8, but apparently not referenced. Earlier discussion of these errors in section 9.4.1 is needed. Should be 'current climate' rather than 'time mean' [Government of Australia]	Accepted -- wording changed from "time mean" to "current climate".
9-1967	9	72	37	72	43	"Adjusted forcing", not "radiative forcing". [Jonathan Gregory, United Kingdom]	Accepted -- terminology changed from RF to AF.
9-1968	9	72	45	72	48	Please use more CMIP5 model results in Figure 9.43, including those from FGOALS-g2, FGOALS-s2 and so on. [Bin Wang, China]	Taken into account -- we have used all models analyzed in the Andrews and Forster papers.
9-1969	9	72	48	72	48	Please consider to add "(cf. Table 9.4)". [Government of Germany]	Editorial
9-1970	9	72	48	72	48	Andrews et al. don't give TCR; I think Tim Andrews provided you with them directly. They are also in Kuhlbrodt and Gregory (2012). [Jonathan Gregory, United Kingdom]	Accepted -- sourcing of TCR corrected and Kuhlbrodt and Gregory cited.
9-1971	9	72	50	73	42	I miss to read about the role of the biosphere on the different feedbacks (e.g. albedo, water vapour, etc.) or an attempt to include biosphere feedbacks as a lumped estimated in Table 9.4 [European Union]	Taken into account -- combined with comment 9-38.
9-1972	9	72	52	72	57	Here it might be useful to remind the reader that the method of evaluating feedbacks differ considerably between AR4 and AR5. [Thorsten Mauritsen, Germany]	Accepted -- Reader is directed to Chapter 8 for updated definitions.
9-1973	9	72	54	72	55	Clarify that you are discussing ensemble mean feedback values here. [Government of Australia]	Accepted -- clarification that these are ensemble mean values has been introduced here.
9-1974	9	72	54	72	56	Another example of compensating errors where a headline figure is used for calibration. Does this mean that previous estimates of the sum of the feedbacks were right-for-the-wrong-reason? [Erica Thompson, United	Noted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Kingdom of Great Britain & Northern Ireland]	
9-1975	9	72	56	72	57	'Advances in determining' sounds like advances in methods of evaluating feedbacks. If this means advances in understanding and in our estimates of the feedback strength, it needs to be clear. [Government of Australia]	Accepted -- language clarified to indicate advances in understanding.
9-1976	9	72				Fig. 9.43: What is the meaning of the open symbols in panel A? [William Merryfield, Canada]	Accepted -- open symbols now explained in the figure legend.
9-1977	9	73	1	73	8	Fig 9.44 needs considerable tidying up: (1) Clarify meaning of all symbols, (2) fix axes ,including x axis title, y-axis number of sig figs (3) change 'squares' to points so they don't overlap into a single 'block' (4) make x axis 'heavy' to highlight (5) make units on y axis K not C. [Government of Australia]	Accepted -- figure has been completely revised.
9-1978	9	73	1	73	8	The labels in this figure are much too small! [Andreas Sterl, Netherlands]	Accepted -- font size for labels has been increased.
9-1979	9	73	2	73	8	This figure is very hard to read. The letters would need to be larger, consider to reorder the legend to increase clarity. [Government of Germany]	Accepted -- figure formatting has been harmonized and improved.
9-1980	9	73	10	73	22	This discussion duplicates a more detailed discussion in Chapter 7 Section 2. Things currently missing from Chapter 7 which should be retained here are the comparisons between CMIP3 and CMIP5, the context of these feedbacks in terms of overall climate sensitivity, and the bearing of paleoclimate which we hardly discuss. [Steven Sherwood, Australia]	Accepted -- duplicative material with chapter 7 has been deleted.
9-1981	9	73	10			Such relationships have been used (sometimes wrongly) to estimate ECS. These are discussed in chapter 10 and this section should at least refer to it. [Reto Knutti, Switzerland]	Noted -- this material has been deleted.
9-1982	9	73	10			strengthening of the links to the assessment provided in earlier sections is needed here. Section 9.7 (and 9.8, in fact) seem rather unconnected to all the earlier pieces of model evaluation. [Thomas Stocker/ WGI TSU, Switzerland]	Noted -- this material has been deleted.
9-1983	9	73	20	73	22	"AOGCMs can reproduce the positive rate of increase in specific humidity with increased SST...". Consider simplifying to: "AOGCMs can reproduce the increase in specific humidity with increasing SST..." Possibly the point is that the rate of increase itself is increasing? In this case suggest: "AOGCMs can reproduce the accelerating increase in specific humidity with increasing SST..." [Government of Canada]	Noted -- this material has been deleted.
9-1984	9	73	25		26	This reference would be better replaced by Ingram, W., 2012. Some implications of a new approach to the water vapour feedback. Climate Dynamics, doi 10.1007/s00382-012-1456-3 (out online 4 months ago, not yet on paper). [William Ingram, United Kingdom]	Accepted -- reference updated.
9-1985	9	73	34			I would say 'are partly due to'. I am pretty much convinced that optical properties of snow is not the main issue, rather the simulated extent. Something which is also confirmed a little later in the text. [Thorsten Mauritsen, Germany]	Accepted -- wording changed.
9-1986	9	73	44	73	49	A comparison between observed NH surface albedo feedback and modeled global dito is not reasonable. In fact there are good reasons to believe that the former should be larger than the latter. [Thorsten Mauritsen, Germany]	Rejected -- both feedbacks are true albedo feedbacks due to changes in surface properties and therefore the comparison is valid.
9-1987	9	73	47	73	49	This seems important but doesn't get much attention. What are the implications of this? [Government of United States of America]	Noted -- this sentence is speculative and has been altered to indicate this is one of several possibilities.
9-1988	9	73	51			Would it not be appropriate to say something about cloud adjustments here? At the very least, it would be useful to mention them, if only to refer to the more detailed discussion in Chapter 7. [Mark Webb, United Kingdom of Great Britain & Northern Ireland]	Noted -- in the context of this section, in which cloud adjustments have been implicitly removed by adjusted forcing, it would not.
9-1989	9	73	56	74	2	Role of diurnal modelling in the spread of SW low-cloud feedback ? [Robert Kandel, France]	Taken into account -- combined with comment 9-880.
9-1990	9	73	56			(Volodin, 2008c) missing in references [Barbara Früh, Germany]	Accepted -- references corrected / added.
9-1991	9	73	56			This needs rewording – the difference between ECS and TCR is mainly due to ocean heat uptake. What you probably meant to say is the relative difference or model ranking or something like that. [Steven Sherwood, Australia]	Accepted -- language has been changed to relative differences.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-1992	9	74	2			It deserves mention that often PPE studies perturb mainly parameters concerning cloud processes, so care should be taken when concluding as is done here. [Thorsten Mauritsen, Germany]	Agreed -- caveat that these results are contingent on the parameter space perturbed has been introduced here.
9-1993	9	74	4	74	17	This discussion should be reconciled with a parallel discussion in Chapter 7 section 2. The following two paragraphs also duplicate discussions there. [Steven Sherwood, Australia]	Accepted -- paragraph now shortened to just focus on sign of feedbacks. Subsequent paragraph has been eliminated (9-2002).
9-1994	9	74	4			I would claim that the radiative kernel method is 'practical'. By saying it is 'well suited' I get the impression that it is an accurate method. However, for instance Jonko et al. (2012) and Block and Mauritsen (2012) found considerable state-dependency of the calculated kernel, which means that beyond a doubling of CO2 errors are sizeable. Further, applying a single kernel to an MME can give misleading results concerning the inter-model spread of cloud feedback (Crook et al. 2011). [Thorsten Mauritsen, Germany]	Accepted -- description of method is now "practical" rather than "well suited", and caveats from these references have been noted.
9-1995	9	74	7			See earlier comment re: surface air temp. [Karen Shell, United States of America]	Noted
9-1996	9	74	8	74	8	The first term is not derived from basic physical principles, but rather derived from running the radiation code in a single GCM, applying it to idealised perturbations of properties (such as water vapour amount) in the current climate. [Government of Australia]	Agreed -- reference to basic physical principles modified to reflect actual procedure for derivation.
9-1997	9	74	8			It is 'assumed to be derivable' [Thorsten Mauritsen, Germany]	Editorial
9-1998	9	74	15	74	15	Critical relationships are not with SSTs (e.g. a regional variation) but with global temperatures. Relationships between SSTs and clouds cannot be extrapolated to global changes, as circulation patterns obviously have a major impact on the former, not just the temperatures. [Government of Australia]	Taken into account -- combined with comment 9-2001.
9-1999	9	74	17			An observational data analysis of upper tropospheric cloud variations with SST also suggested a positive upper tropospheric cloud feedback in the current climate (Su et al., 2008). Su, H., J. H. Jiang, Y. Gu, J. D. Neelin, B. H. Kahn, D. Feldman, Y. L. Yung, J. W. Waters, N. J. Livesey, M. L. Santee, and William G. Read, 2008: Variations of tropical upper tropospheric clouds with sea surface temperature and implications for radiative effects, J. Geophys. Res., 113, D10211, doi:10.1029/2007JD009624. [Government of United States of America]	Taken into account -- combined with comment 9-2001.
9-2000	9	74	17			The citation to Zelinka et al. (2012) is incorrect, as they do not assess interannual cloud feedbacks. The correct citation is: Zelinka, M.D. and D.L. Hartmann, 2011: The Observed Sensitivity of High Clouds to Mean Surface Temperature Anomalies in the Tropics. J. Geophys. Res., 116, D23103, doi:10.1029/2011JD016459. [Government of United States of America]	Taken into account -- combined with comment 9-2001.
9-2001	9	74	17			However, Dessler (2010) found no correlation between model cloud feedback derived from his 'observational method' and the models actual climate change cloud feedback. The same goes for Lindzen and Choi (2011). Thus, one should be very careful in putting too much weight on observational evidence for cloud feedback. There was an extensive discussion on this in Chapter 7.2.4.3.7. [Thorsten Mauritsen, Germany]	Accepted -- The sentence beginning "This finding is consistent" at line 14, p. 74 has been replaced with a sentence referring the interested reader to Chapter 7.2.4.3.7.
9-2002	9	74	22	74	24	In the Clement et al. study there were actually only 2 models that reproduce the observed relationships. One of them was dismissed by Clement et al. for being weird in other ways, leaving one model that they concluded got it right for a good reason. Furthermore, Broccoli and Klein (2010, Science) concluded that the analysis of the GFDL model by Clement et al. was flawed because the GFDL cloud signal came from high clouds rather than the low clouds that Clement et al. assumed were dominating all the models. [Anthony Del Genio, United States of America]	Accepted -- because of the problematic nature of the model analysis in Clement et al (2009), this paragraph has been dropped.
9-2003	9	74	27	74	32	I think it would be useful to note that the studies by Ogura et al refer mainly to mixed phase cloud feedbacks in the mid-latitudes. [Mark Webb, United Kingdom of Great Britain & Northern Ireland]	Noted --- this material has been deleted
9-2004	9	74	27	74	32	The discussion of high cloud feedbacks here is brief and mainly focused on microphysical feedbacks ; other feedback mechanisms (e.g. the FAT hypotheses are discussed in Chapter 7.) It would I think be useful to note this here. [Mark Webb, United Kingdom of Great Britain & Northern Ireland]	Accepted -- other possible feedback mechanisms including the FAT hypothesis are now noted.
9-2005	9	74	34	74	44	What does this mean? Do we know the forcing? How large are the error bars? [Peter Braesicke, United Kingdom]	Rejected. The sources of the forcing are explained in the figure caption and the errors in this estimation

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							process are discussed at the end of the paragraph in question.
9-2006	9	74	39	74	42	Pls change "Andrews et al (2012a)" to "Andrews et al., (2012a)" [HASIBUR RAHAMAN, India]	Editorial
9-2007	9	74	44			The largest error is about 50 percent, which deserves mention. [Thorsten Mauritsen, Germany]	Agreed -- upper bound on errors in estimation of ECS from forcing/feedback ratios is now noted.
9-2008	9	74	46	74	46	9.7.2.2.4 Relationship of feedbacks to uncertainty in modelled climate sensitivity should be: 9.7.2.2.5 Relationship of feedbacks to uncertainty in modelled climate sensitivity [Koichi Sakaguchi, the United States]	Editorial
9-2009	9	74	49	74	51	Unclear what this sentence means. [Government of Australia]	Agreed -- this sentence has been merged with the subsequent (primary) sentence regarding shifts to lower sensitivity.
9-2010	9	74	51	74	51	"If structural biases are shared" should read "If shared structural biases are identified"? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Editorial
9-2011	9	74	52	74	52	The abbreviation MME has not been introduced yet. [Farahnaz Khosrawi, Sweden]	Accepted -- acronym spelled out.
9-2012	9	75	4	75	5	This last sentence seems to not add much to the argument. [Thorsten Mauritsen, Germany]	Agreed -- this sentence has been dropped and the Sexton paper is now addressed in greater detail in 9.7.3.3.
9-2013	9	75	5	75	5	The abbreviation PPE has not been introduced yet. [Farahnaz Khosrawi, Sweden]	Accepted -- acronym spelled out.
9-2014	9	75	7			If they used a SOM-GCM, making a statement about AOGCMs requires extra steps. Maybe just say "GCM climate sensitivity" [Karen Shell, United States of America]	Noted -- this material has been deleted
9-2015	9	75	8	75	8	delete "cloud" from cloud entrainment [David Sexton, United Kingdom]	Noted -- this material has been deleted
9-2016	9	75	8	75	9	Pls put "., " after "et al" [HASIBUR RAHAMAN, India]	Noted -- this material has been deleted
9-2017	9	75	9	75	19	It is worth double checking with mark.webb@metoffice.gov.uk or Jonathan Gregory, but I don't think the issue of fast cloud adjustment in a slab model affects the equilibrium sensitivity (ECS). If you use the Gregory and Forster method, sure there is an instantaneous adjustment due to fast processes and this changes the intercept. But the Gregory and Forster line goes through (ECS, 0) and a change in intercept does not change that. Instead the slope changes i.e. the climate feedback parameter. So I think it is correct to say the cloud adjustment is an additional source of uncertainty in equilibrium climate feedback in a slab experiment, it does not affect equilibrium climate sensitivity. [David Sexton, United Kingdom]	Noted -- this material has been deleted
9-2018	9	75	9			This should say 'an AOGCM': the statement does not apply generally. [Andrew Gettelman, United States of America]	Accepted
9-2019	9	75	9			From 'An additional...' and onwards is a different topic from the first part of the paragraph. This information should also occur much earlier as suggested in my comment 13. [Thorsten Mauritsen, Germany]	Taken into account -- combined with statement 9-1940.
9-2020	9	75	23	75	24	Pls give gap after °C and % [HASIBUR RAHAMAN, India]	Editorial
9-2021	9	75	23	75	38	Am I over-reacting by being quite concerned at the fact that models apparently reproduce the warming from 1850-2005 to a better agreement than equilibrium climate sensitivity would suggest? Can the authors reassure the reader that the tuning of climate models referred to earlier in the chapter does not extend to tuning the models to get about the right degree of warming from 1850-2005? I hope so, as statements elsewhere imply we have a degree of confidence in climate models because of their ability to reproduce past change. Please also note later comment 301, concerning the discussion of ECS in both chapters 9 and 10. [Adrian Simmons, United Kingdom]	Noted -- however, the text already makes it clear implicitly that in CMIP5 there is no indication of tuning by aerosol forcing. And as stated in Box 9.1, the details of how modelling centres tune their models are often not known. Therefore, the current text continues to rely on the published literature. Introducing language referring back to Box 9.1 would be more speculative and impede clarity.
9-2022	9	75	24	75	27	Please add to this sentence "or that because the AOGM's are deficient in other areas, the influence of CO2 has been over-emphasised." This statement is consistent with the mismatch between model-predicted and	Rejected -- This subject is covered in detail in the new "hiatus" box 9.2 added to the chapter.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						observed temperatures over the last 16 years. [John McLean, Australia]	
9-2023	9	75	33	75	38	This is a strong statement. If correct, is there an alternate hypothesis to explain the large spread in ECS and narrow spread in 20th century change? [Jean-Christophe Golaz, United States of America]	Taken into account -- the strong concluding statement has been modified to be fully backed by the published literature. Paragraph has also been re-ordered.
9-2024	9	75	33			"comparable range of sensitivities" may be unclear to the reader. We feel it may be more appropriate to say "wide range of sensitivities". [Government of United States of America]	Editorial
9-2025	9	75	40			The EMIC discussion here is extremely short and not informative. The conclusion drawn about the EMIC, even though I agree with it, is not demonstrated or discussed here. [Reto Knutti, Switzerland]	Taken into account -- combined with comment 9-44.
9-2026	9	75	41	75	41	update or skip Zickfeld et al., submitted [European Union]	Accepted
9-2027	9	75	43	75	43	Pls change "(Eby et al., 2012)" to "Eby et al., (2012)" [HASIBUR RAHAMAN, India]	Editorial
9-2028	9	75	44	75	44	The lower panel of Figure 9.8 shows results from EMICs until 2005 only. Thus figure 9.8 does not present the "relevant results" as claimed on line 44. [European Union]	Accepted -- sentence is qualified with an end date of 2005 for the EMIC results shown in Figure 9.8.
9-2029	9	75	45	75	46	I found little supporting evidence for making such a strong statement in that particular figure. [Thorsten Mauritsen, Germany]	Accepted -- this statement has been removed.
9-2030	9	75	48			I believe the authors mean Table 9.5. [Thorsten Mauritsen, Germany]	Accepted -- reference corrected to Table 9.5
9-2031	9	75	53	75	53	update or skip Zickfeld et al., submitted [European Union]	Taken into account -- combined with comment 9-2026
9-2032	9	75	56	76	4	Should be Table 9.6. This table is missing in the table section starting on p. 138. [European Union]	Accepted -- table number corrected to 9.6, and table moved to table section at the end of the chapter.
9-2033	9	76	10			Hargreaves et al (GRL 2012 in press) presents an analysis based on the PMIP2 models and the most recent proxy data, and tentatively provides an estimate of the climate sensitivity. [James Annan, Japan]	Taken into account -- combined with comment 9-2035.
9-2034	9	76	10			Section 9.7.2.3.2: State that climate sensitivity is state dependent: non linear nature of water vapor and wv feedbacks is one cause. Snow and ice albedo feedbacks are another relevant for LGM. Thus a glacial world with more ice and less water vapor should have a different sensitivity (probably lower) [Andrew Gettelman, United States of America]	Accepted -- the dependence of climate sensitivity on the initial climate state subject to perturbation is now explicitly noted.
9-2035	9	76	14	76	14	Hargreaves et al. (2012) "Can the last glacial maximum constrain climate sensitivity?" Geophysical Research Letters, doi:10.1029/2012GL053872 confirms a good correlation of climate sensitivity with tropical LGM temperature change but not with Antarctic LGM temperature change for PMIP2 multi-models. [Masakazu Yoshimori, Japan]	Accepted -- this qualification regarding the poorer correlation for Antarctica from more recent model intercomparisons is now cited here.
9-2036	9	76	26	76	26	... is the dominant... [European Union]	Editorial
9-2037	9	76	27	76	27	"for" in "Climate Sensitivity for the Last Glacial Maximum" sounds a bit strange. [Masakazu Yoshimori, Japan]	Editorial
9-2038	9	76	27	76	27	The reference to Sect. 5.3.3 is needed here. [Masakazu Yoshimori, Japan]	Accepted -- reference to section 5.3.3 on "Glacial Climate Sensitivity" has been added.
9-2039	9	76	29			This section does not seem to be as up to date as section 9.2.2.7 missing some important work in this area by Sanderson (2012) and Sexton et al (2012), as cited in section 9.2.2.7. In particular, these papers have recognised that a single PPE is not sufficient as stated on p.77 line 13, and so have used both PPE and MME data. Also both papers provide uncertainty ranges on climate sensitivity accounting for both PPE and MME, and come up with relatively tight ranges because of the work done by the observational constraints. These two results are at odds with lines 12-13 and lines 19-21. [David Sexton, United Kingdom]	Accepted -- recent Sanderson and Sexton work concerning the tighter ranges of climate sensitivity from recent analyses of PPE and MME.
9-2040	9	77	2			"The perturbation of atmospheric and surface albedo feedbacks". Perturbation of feedbacks does not make sense. We recommend that the authors use "parameters" instead of "feedbacks" to make the phrase clearer. [Government of United States of America]	Noted -- this material has been deleted
9-2041	9	77	5	77	39	Pls give gap after °C [HASIBUR RAHAMAN, India]	Noted -- this material has been deleted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2042	9	77	7			For the other studies the endpoints of the interval is given rather than just the range. [Thorsten Mauritsen, Germany]	Noted -- this maertial has been deleted
9-2043	9	77	10	77	11	Sanderson (2012) shows that by removing members from climateprediction.net PPE which are poor (woefully out of balance at TOA) that MME is reasonably consistent with emergent relationship from the subset of relatively good ensemble members. It would be good to balance the negative sentiment of line 10 with this result. [David Sexton, United Kingdom]	Noted -- this maertial has been deleted
9-2044	9	77	11	77	11	Rougier et al (2009b) should be deleted as it did not consider CMIP3 data. Also Rougier et al (2009b) is the same as Rougier et al (2009a) so needs to be removed on page 122. [David Sexton, United Kingdom]	Noted -- this maertial has been deleted
9-2045	9	77	12			The text suggests that because PPE relationships between climatological quantities and climate sensitivity are not robust when extended to PPEs from other climate models, that this implies that the underlying model structure is important. While that is one possible conclusion: it is actually not the most likely. A more likely conclusion is that relationships between climatological mean quantities (i.e. base climate state) are just very poorly related to climate sensitivity as a physical principle. There is no physical equation or principle that says they should be closely related. Instead, this just supports the idea that how well a climate model matches observed mean climatologies is a very weak indication of how well it matches climate sensitivity of the Earth. See Klocke, Pincus, and Quaas, J. Climate, 2011 for examples. [Government of United States of America]	Noted -- this maertial has been deleted
9-2046	9	77	15			See earlier discussion of natural fluctuations as analogues. [Karen Shell, United States of America]	Noted -- this maertial has been deleted
9-2047	9	77	16	77	16	Wu and North (2003) do not use CMIP3 simulations. (It was published too early for that too have been possible) [Timothy Merlis, United States of America]	Noted -- this maertial has been deleted
9-2048	9	77	19	77	21	Another negative sentence focussing on "most studies" which are not cited and therefore we cannot judge the assumptions or methods used. It would be worth citing some papers here where that is the case. However, there are those papers that have (probably) put more effort into trying to do a more comprehensive job of combining structural and parametric uncertainty plus including a more sophisticated methods on observational constraints. I would say they have narrowed the range somewhat. Sanderson (2012) and Sexton and Murphy (2012) are both about 4.3K for 95%-ile. OK, it may not be "significant" but it is nevertheless counter the negative sentiment of this sentence based on unknown references. [David Sexton, United Kingdom]	Noted -- this maertial has been deleted
9-2049	9	77	22	77	24	Huber et al (2011) do not account for predictor screening. For example, one expects to achieve "5% significance" (the filtering criteria used in the paper) by chance 85 times when 1700 predictors are tested (this is the number of predictors examined in the paper). One would have to account for predictor screening to produce climate sensitivity estimates and confidence intervals that are comparable to the range of CMIP models. [Timothy Merlis, United States of America]	Noted -- this maertial has been deleted
9-2050	9	77	23	77	23	"likely or very likely range of about 2 - 5 degrees...." -> So is this range a likely range or very likely range? Please specify. This current wording gives the wrong impression that the chapter can not agree on the likelihood level that should be used. [Thomas Stocker/ WGI TSU, Switzerland]	Noted -- this maertial has been deleted
9-2051	9	77	39			I would be careful with coining this 'the consensus range'. [Thorsten Mauritsen, Germany]	Taken into account -- combined with comment 9-2052.
9-2052	9	77	39			"consensus range" -- we suggest to specify that this range represents the (IPCC) consensus on what is "likely" the range of climate sensitivity. As such, the consensus also is that lower/higher values can not be excluded. This should be clarified. [Thomas Stocker/ WGI TSU, Switzerland]	Accepted -- the term "consensus range" has been clarified.
9-2053	9	77	43			This is now Section 10.8.2. It is distracting to have the ECS discussion so split up among different chapters. [Government of United States of America]	Accepted -- reference to section in Chapter 10 corrected.
9-2054	9	77	47			Section 9.8.1 presents a summary or synthesis of the assessment presented in Ch9. It would thus make sense to us to revise the title to reflect this synthesis-status (Figure 9.45 is also introduced as "Summary of the findings in Chapter 9...": suggest to change to "Summary/Synthesis Assessment of Model Performance" [Thomas Stocker/ WGI TSU, Switzerland]	Accepted - title of the section changed to 'Synthesis Assessment of Model Performance' as suggested.
9-2055	9	77	49	77	54	Not everything was strictly 'quantitative', I would drop the word; ECV concept should be introduced early and	Accepted - 'quantitative' deleted; 'ECVs' also deleted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						used consistently. [Peter Braesicke, United Kingdom]	since the term is not introduced in Chapter 2 and correspondingly not in Chapter 9.
9-2056	9	77	51			"ability to simulate" here, and throughout the chapter (but in this section on credibility in particular) there is a need to distinguish clearly what is "included" or "simulated" from what is simulated realistically. A clear and open discussion on potential applications where current models are judged NOT credible would be a great aid toward understanding the credibility of CMIP5 model output. [Leonard Smith, United Kingdom]	Noted - The summary figure that is presented in this section shows what this comment asks for, namely it summarizes what the models do well and where they still have problems.
9-2057	9	77	52	77	52	Range in model quality or range in performance metric quality (or both)? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted - but sentence has been deleted.
9-2058	9	77				Section 9.8.1: This is not a good assessment of model performance as the table does not allow for the fields used to tune each model or the number of parameters available for tuning. [David Webb, United Kingdom]	Noted - As highlighted in the Box 9.1, modelling centres do not rigorously document their tuning procedure. It is therefore not possible to indicate in the table what was tuned. The text throughout the chapter has been sharpened to indicate what are weak and strong tests (now discussed in 9.2.3), which should help put the results in this table into perspective.
9-2059	9	78	5	78	5	... The x-axis... [European Union]	Accepted -changed as suggested.
9-2060	9	78	5	78	5	"y-axis" - should be x-axis? [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Accepted -changed as suggested.
9-2061	9	78	5	78	5	Pls change "Mastrandrea et al. 2011)" to "Mastrandrea et al., (2011)" [HASIBUR RAHAMAN, India]	Editorial
9-2062	9	78	5	78	5	"y-axis" should read "x-axis" [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Accepted -changed as suggested.
9-2063	9	78	5			"The y-axis refers to ..." should be "The x-axis refers to ...". [Weiping Li, China]	Accepted -changed as suggested.
9-2064	9	78	5			"y-axis" should be replaced by "x-axis" [RYM MSADEK, United States of America]	Accepted -changed as suggested.
9-2065	9	78	6	78	7	In Mastrandrea 2011 the indices of interest are "agreement" and "amount of evidence". In figure 9.45 are "model quality" and "confidence", which I do not think work as well. For example, the lower right region is a bit paradoxical: very high confidence achieved with low model quality. In addition, the shading just repeats information present in x-axis. These kind of graphs are very challenging to do and may leave a lasting impression on readers; so extra care should be taken to be sure that they really work. [Ramon de Elia, Canada]	Taken into account - We have improved the description and slightly changed the style of the summary figure for clarity.
9-2066	9	78	7			Please change "level of evidence and the level of agreement" to the terms introduced in the Uncertainty Guidance Note: "amount of evidence and the degree of agreement". [Thomas Stocker/ WGI TSU, Switzerland]	Taken into account - we now use language consistent with the AR5 guidance notes. Note, however, that Co-chair here suggests AR4 language. In AR5 guidance notes, it is "level of evidence and degree of agreement".
9-2067	9	78	13	78	13	Surface air temperature is referred to as "SAT" in text while figure 9.45 and caption call it "TAS" [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Accepted - inconsistency removed.
9-2068	9	78	13	78	14	The notations need to be consistent between the text and the figure. The mean surface air temperature is referred to as "SAT" in the text and "TAS" in the figure and caption [RYM MSADEK, United States of America]	Accepted - inconsistency removed.
9-2069	9	78	21	79	28	The subsection numbers are sometime not specific enough (e.g., Large scale precipitation (Section 9.4.1) should be Large scale precipitation (Section 9.4.1.1)). I believe this will and needs to be taken care of at the professional editing as mentioned in the invitation letter. [Koichi Sakaguchi, the United States]	Accepted - Section numbering now more specific where appropriate.
9-2070	9	78	21			Figure 9.45: need to check carefully the entries in this figure. For panel 'C' the entries are partly inconsistent with the information provided in Table 9.3 (e.g., SAM, PDO, AMOC are all given as low confidence in the table, but are in the medium confidence sector here). [Thomas Stocker/ WGI TSU, Switzerland]	Accepted - inconsistency removed.
9-2071	9	78	28	79	28	Don't put this list in the figure caption. An extra table would be more adequate. [Farahnaz Khosrawi, Sweden]	Rejected - the author team believes the figure is clearer if the list is in the caption.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2072	9	78	60			Are you sure this is total ozone, it looks strictly like stratospheric by it name: StratO3 [Michael Prather, United States of America]	Accepted - abbreviation changed to TotalO3 to make this clear.
9-2073	9	79	30	79	30	Section 9.8.2: Please refer to chapter 10 to make the link to the D&A in AR5. [Thomas Stocker/ WGI TSU, Switzerland]	Accepted
9-2074	9	79	30	80	2	This section leaves me feeling cold as it is direct over-reach into the direct charge of the following Chapter. This is just an invitation to critics to find differences and rubbish the whole report. I would strongly suggest this be removed or shipped to Chapter 10 to reduce such propensity for cross-talk and unfair criticism of either / both chapters and the report as a whole. [Peter Thorne, United States of America]	rejected. The section is co-authored with chapter 10.
9-2075	9	79	33			"therefore" is inappropriate. In fact the preceding clause makes it clear why their use in attribution and detection is nontrivial. [Leonard Smith, United Kingdom]	Taken into account.the sentence has been modified
9-2076	9	79	34	79	34	"key" should read "required basis" [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account.sentence deleted
9-2077	9	79	37	79	38	Example is a bit disingenuous; the quantity is insensitive to model uncertainty, when the model uncertainty is small. What about things (storms, sea ice, extremes, regional circulation) less directly related to single simple physical processes? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. the section has been slightly reorganised. However given space constrains it is not possible to provide a statement for all variables and only a few representative examples are provided
9-2078	9	79	40	79	40	The uncertainties in stratospheric ozone forcings are discussed in section 9.4.1.3.5, thus in this paragraph a reference to this section is missing. [Farahnaz Khosravi, Sweden]	Accepted. reference added
9-2079	9	79	47	79	49	"the better representation of the intraseasonal to decadal variability in terms of pattern and magnitude, suggest with medium confidence that model better reproduce internal variability": It would be helpful to explain which variability this refers to. [Government of Canada]	Accepted. A reference to section 9.5.2 and 9.5.3 has been added
9-2080	9	79	49			"that model[s] better reproduce" [Andrew Wittenberg, United States of America]	Accepted.
9-2081	9	79	54	79	57	So are these "improved detections and attribution results" improved to the point of being useful or do the "major biases" still preclude meaningful D+A at regional scales? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. It is difficult however to make a short statement valid for all variables
9-2082	9	79	56	79	59	It should be mentioned here that the D&A approach is conditional upon using statistical frameworks that make reliable assumptions about climate model biases (e.g. that output can be linearly scaled to match observations). Hence, model evaluation is a prerequisite for designing well-specified statistical frameworks (see 9.8.3) that can then be used in D&A. [David Stephenson, United Kingdom of Great Britain & Northern Ireland]	Rejected. This is not the place to explain the D&A methodology. Reference is made to chapter 10
9-2083	9	79	57	80	2	Please add references for "Recent studies..." [Government of United States of America]	Taken into account.reference to section 9.5.4 where extremes are discussed has been added
9-2084	9	79	57			Again, Hansen et al. (2012, PNAS) should be mentioned here (recent studies of climate extremes ...). [Martin Dameris, Germany]	Taken into account.reference to section 9.5.4 where extremes are discussed has been added
9-2085	9	79				Section 9.8.2: I thought that this was a weak and verbose section. For example I cannot see how the verbose second sentence is a consequence of the first. In the next sentence the argument that variability gives a better indication of detection and attribution than the mean state appears to be a hope not a fact. [David Webb, United Kingdom]	Taken into account. The first paragraph has been modified, the whole section reduced and slightly reorganised to provide a better statement.
9-2086	9	80	4	80	17	Section 9.8.3 : Our confidence in model projections does not only rely on this. As discussed in a WCRP position paper (Bony et al., Springer Monograph, in press and available on the website of the WCRP Open Science Conference), it also (primarily?) relies on our physical understanding of how the climate system works, which depends on a large spectrum of models, theories and observations, and of our assessment of the strengths and limitations of climate modelling. [Sandrine BONY, France]	Taken into account - the discussion on confidence in climate model projections has been extended along these lines.
9-2087	9	80	6	80	6	"... cannot be directly evaluated without waiting multiple decades and hoping that the scenarios were	Taken into account – in response to other comments,

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						reasonably close.” (Not quite the same thing). [Gavin Schmidt, United States of America]	the text has been revised to clarify that our confidence in model projections does not only rely on the direct evaluation of climate models with observations, but is drawn on multiple lines of evidence (see reviewer comment 9-2086). We also now refer to Figure 1.4 that directly evaluates projections from previous assessment reports.
9-2088	9	80	6	80	14	Therefore, "the" uncertainty in future climate projections is not an objectively knowable quantity. Subjective assessments of model quality and therefore projection uncertainty can be based on, but not equal to, top-down evaluations of historical simulations and bottom-up evaluations of physical process representation. Subjective assessments of projection uncertainty can also incorporate the possibility of rapid or unexpected change, although meaningful quantification of probabilities is challenging. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Noted - text has not been changed since this is discussed already in the SOD (lines 12-14)
9-2089	9	80	6	80	17	The references mentioned here do not fit: Eyring et al. is related to CCM studies (which are in this case not climate models) and Knutti et al. is not a peer-reviewed paper. [Martin Dameris, Germany]	Rejected - IPCC reports may use literature that is not peer-reviewed (Appendix A to the Principles Governing IPCC Work). The IPCC guidance papers (i.e. the Knutti et al. paper cited here) in fact are reviewed by targeted reviewers, so they are peer-reviewed in the same way as an IPCC chapter is. They are just not published in a journal. We have removed the first reference although in the context used here we do not agree that this is not an appropriate reference.
9-2090	9	80	7	80	8	This type of broad-brush statement about climate models is highly inadvisable. Certain aspects of climate model projections are in fact based on clear physical principles that give us fairly high confidence that the specific predictions are correct at least in a qualitative sense. I would put the GHG radiative forcing, the Planck response, and to at least some extent the water vapor feedback in that category. Other aspects of climate change are based on parameterization assumptions that are highly uncertain and in some cases may not be physically realistic, e.g., the processes that control low cloud feedback, or else they may be qualitatively realistic but quantitatively in error such that they interact with the general circulation in incorrect ways and produce incorrect or at least uncertain projections of regional climate change (drought/flood patterns, hurricane intensity changes, etc.). Some things are correct about climate models, some things incorrect. We are best served by clarifying which things we are certain of and which we are uncertain about and not making grand statements about the models overall that will be (rightly) criticized later. [Anthony Del Genio, United States of America]	Taken into account - the discussion on confidence in climate model projections has been extended.
9-2091	9	80	25	80	28	It would be useful to cite here this recent article, which critically reviews the implicit assumptions involved in the one vote per model averaging approach: Sansom, P. G., D. B. Stephenson, C.A.T. Ferro, G. Zappa, and L.C. Shaffrey, 2012: A simple framework for weighting climate change projections in multi-model ensembles. Journal of Climate. Submitted. [David Stephenson, United Kingdom of Great Britain & Northern Ireland]	Accepted - reference added. Please note that the discussion on independence of the models has been moved to Section 9.2.2 to address another reviewer comment.
9-2092	9	80	32	80	33	several word repetitions in this sentence. [Farahnaz Khosrawi, Sweden]	Editorial. Please note that the discussion on independence of the models has been moved to Section 9.2 to address another reviewer comment.
9-2093	9	80	35			The “Despite this lack of independence” clause doesn’t seem appropriate here and would be better omitted. The rank histogram approach can easily be applied to a weighted ensemble (eg for the purpose of downweighting related models, as in Hargreaves et al, Clim. Past Discuss. 8, 3481-3511 2012 currently in review). The debate over independence doesn't really have anything to do with rank histograms per se. [James Annan, Japan]	Accepted. The text was revised to remove the mismatch between discussing model independence and rank histogram methods. Please note that the discussion on independence of the models has been moved to Section 9.2.2 to address another reviewer comment.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2094	9	80	37	80	40	<p>Sakaguchi et al. (2012) recently used this approach (rank histogram) for surface air temperature 'trend' instead of climatological fields as done in the cited references. I'd like to suggest to add one sentence mentioning this study, following L37-40 as:</p> <p>"..., but existing single-model-based ensembles tend to be too narrow (Yokohata et al.,2012a; Yokohata et al., 2012b). When this method is applied to surface air temperature trend, for example on global average, CMIP3/5 ensembles are reliable for some temporal scales (e.g., 10-year trend) but not necessarily so for other scales (e.g., 50-year trend) due to scale-dependent mean bias and ensemble dispersion (Sakaguchi et al., 2012). This "reliability test" highlights..."</p> <p>Sakaguchi, K., X. Zeng, and M. A. Brunke, 2012: The hindcast skill of the CMIP ensembles for the surface air temperature trend. J. Geophys. Res., 117, D16113, doi:10.1029/2012JD017765. [Koichi Sakaguchi, the United States]</p>	Accepted. The reference was added. Please note that the discussion on independence of the models has been moved to Section 9.2.2 to address another reviewer comment.
9-2095	9	80	42		43	<p>Mentioning the parallel work on the LGM (ie testing a climate change not just a steady climate) should be helpful here. The results for the LGM are consistent with the findings of Yokohata et al. Although the analysis of a climate change strengthens the result, it is not as powerful as one might hope because of the relatively large uncertainty in the paleo observations compared to historical observations. BTW: even the multi-model ensemble fails for 6ka. Hargreaves et al 2011 CP, Hargreaves et al 2012 CPD. [Julia Hargreaves, Japan]</p>	Rejected. Length constraints do not permit adding another aspect, especially since it does not strengthen the assessment from contemporary climate.
9-2096	9	80	46	80	46	<p>Should cite Reifen and Toumi here (doi:10.1029/2009GL038082) [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]</p>	Accepted - reference added. Please note that the discussion on independence of the models has been moved to Section 9.2 to address another reviewer comment.
9-2097	9	80	46	80	57	<p>Besides model ensemble weighting, the authors should also mention that another frequently used approach to improve the projections: re-calibration, that is, scaling the model outputs to a given observed value. For example, Wang and Overland (2009,2012) shift the CMIP3 and CMIP5 summer sea ice extents as to match the present value, and then infer the possible dates of ice-free conditions; Mahlstein and Knutti (2012) adjust the CMIP3 simulated September Arctic sea ice extent to the observed value, and then only explore its sensitivity to global temperature to derive the critical temperature threshold after which ice-free conditions are likely to occur. Boé et al. (2009) re-calibrate the CMIP3 simulated trend in September sea ice extent on the observed value to obtain the percentage of remaining ice in the future. While the present remark also applies to Chapter 12, I think at least mentioning the existence of such methods would be appreciable. [François Massonnet, Belgium]</p>	Accepted - studies on re-calibration mentioned.
9-2098	9	80	46	80	57	<p>Please add a sentence to say that climate models are explicitly tuned to match historical observations as closely as possible. This tuning is inevitable if models are to be regarded as producing "ball park" results and worthy of improvement. [John McLean, Australia]</p>	Noted - As highlighted in Box 9.1 models are tuned to match a small number of global constraints. They are not tuned to match all historical observations. Given the complex design of the models, this would not be possible. The section on ensembles has been moved into the new Section 9.2.2.
9-2099	9	80	50			<p>Knutti et al. (2010b) is not a peer-reviewed paper. [Martin Dameris, Germany]</p>	Rejected. IPCC reports may use literature that is not peer-reviewed (Appendix A to the Principles Governing IPCC Work). The IPCC guidance papers in fact are reviewed by targeted reviewers, so they are peer-reviewed in the same way as an IPCC chapter is. They are just not published in a journal.
9-2100	9	80	51			<p>Waugh and Eyring (2008) belongs to AGCMs with chemistry, not climate models. [Martin Dameris, Germany]</p>	Rejected. The Waugh and Eyring (2008) paper is among the pioneering work on performance metrics, so should be cited here. It is also one of the first that applied performance metrics in a process-oriented model evaluation approach. The method is the same, whether or not applied to chemistry-climate or climate models.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2101	9	80	52	80	54	It might be worth mentioning here that although the mean responses are found to be similar the precision of these estimates can be more sensitive to the choice of weighting (e.g. Bracegirdle and Stephenson, 2012a). [David Stephenson, United Kingdom of Great Britain & Northern Ireland]	Taken into account - text expanded to comment on precision.
9-2102	9	80	54	80	57	But we still don't know if it is a better forecast! Describe the role of outliers (and what would we do if one model loses sea-ice too quickly). [Peter Braesicke, United Kingdom]	Noted - this is further discussed in Chapter 12 which the text refers to.
9-2103	9	80				<p>Section 9.8.3: As far as I can tell this is the first place in the whole document that it states that confidence in the models "cannot be directly evaluated and so our confidence ... must be built upon the knowledge that these models are based on physical principals and a demonstration of how well models represent ... especially those related to important feedbacks". It also emphasises how much climate change projections "push the models into unknown territory".</p> <p>This is all good stuff (despite being a bit long-winded) the main bit left out being the problem of not using the data that was used to tune each model.</p> <p>It is wrong that this section is lost back here after pages of verbose description. It needs to be a much more visible part of both this chapter and the Technical Summary.</p> <p>However I am then disappointed that instead of some hard results there is then a discussion of Multi-Model Ensembles - a statistical approach which may be a good first stage in judging models but, as here, has a habit of not looking much further into the physics.</p> <p>There is then a rather rambling discussion of "emergent constraints" (page 81, line 4) and the rather shattering statement that "many studies have failed to find strong relationships between observables and projections". If the latter is true all of the concerns with "how well models represent ... processes" (page 80, line 8 and in many other sections throughout the whole document), is a waste.</p> <p>[David Webb, United Kingdom]</p>	Taken into account - We agree that the text on independence of multi-model ensembles is not correctly placed here and have moved it to Section 9.2. The text on emergent constraints has been revised and provides some more concrete examples, that are for example used in Chapter 12 to explore the value of weighting for one particular application. With regard to model tuning, we have sharpened the discussion on weak and strong tests given the existence of tuning in Section 9.2.3. Most of the model results assessed in this chapter are not directly tuned. As pointed out, only a small number of global constraints are subject to tuning, although the lack of documentation makes it difficult to assess what was tuned in individual models.
9-2104	9	81	2	81	35	Important discussion of snow-albedo errors in models which could be fixed, reducing spread in results between models. Needs emphasising as an area where research could be targeted to improve model simulations at high northern latitudes. [European Union]	Rejected - this is a climate assessment not a document that emphasises future research needs.
9-2105	9	81	4	81	4	It is shown in Tsushima et al.(accepted) that models, which have better agreement with observations in their representation of the seasonal variation of the cloud regimes, tend to have higher climate sensitivity. Although the result is based on five CMIP5 models, this study can be sited in this sentence. [Yoko Tsushima, United Kingdom of Great Britain & Northern Ireland]	Accepted - reference added.
9-2106	9	81	4	81	6	The sentence is difficult to read as it is now. I would suggest to rephrase it. One could write "e.g." instead of "where". However, then one would have a repetition with "For example" in the sentence starting at line 6. [Farahnaz Khosrawi, Sweden]	Accepted - sentence reworded.
9-2107	9	81	4	81	7	Consider adding the paper of Massonnet et al. (2012; cited in Chapter 9) in the references. [Thierry Fichefet, Belgium]	Noted - the paper by Massonnet et al. (2012) is already cited in the previous paragraph on sea-ice.
9-2108	9	81	4	81	23	Analogous relationships however to those of Hall and Qu for snow albedo, have not been found with other feedbacks as shown in: Colman, R.A. and L.I. Hanson, 2012: On atmospheric radiative feedbacks associated with climate variability and change. Climate Dynamics, doi 10.1007/s00382-012-1391-3, and Dessler, A.E. and S. Wong (2009): Estimates of the water vapour climate feedback during El Niño-Southern oscillation. J. Climate, 22, doi: 10.1175/2009JCLI3052.1. [Government of Australia]	Taken into account - discussion extended.
9-2109	9	81	4	81	23	This paragraph needs to include a statement about ENSO modelling needing to improve. McLean et al (2009) showed that the Southern Oscillation Index provides a good indicator of global average lower tropospheric temperatures about seven months later. (The paper was criticised but the criticism didn't focus on the Discussion and Conclusions, and it contained several blatantly false claims about what the paper said. The journal refused to extend the basic and common courtesy of allowing the paper's authors to respond, and	Rejected. The existence of this emergent constraint on tropical carbon loss evidently does not depend on the quality of ENSO simulation, as models clearly exist on a common line in Figure 9.46b regardless of differences in their simulation of ENSO.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						surely you don't condone that refusal?). The fact that SOI changes predate LTT changes and the relationship is consistent for most of the last 60 years (except for short periods following tropical volcanic eruptions) shows that ENSO is very likely the primary driver of temperature. It therefore follows that ENSO prediction needs to improve before temperature predictions can improve. [John McLean, Australia]	
9-2110	9	81	6			Eyring et al. (2007) does neither belong to the topic of mean climate nor to (past) climate trends and variability. It deals with future projections of stratospheric ozone. [Martin Dameris, Germany]	Rejected - it does belong to the topic of emergent constraints, where variations in model projections are found to be related to inter-model variations in the past. Projections of stratospheric ozone and associated impacts on surface climate are discussed in subsequent chapters and stratospheric ozone is evaluated in Chapter 9.
9-2111	9	81	8	81	8	Pls change "Hall and Qu 2006" to "Hall and Qu (2006)" [HASIBUR RAHAMAN, India]	Editorial
9-2112	9	81	25	81	35	Delete this paragraph. No credible evidence has been shown to support the contention that CO2 has a significant influence on temperature. The only so-called evidence has been the output of climate models that IPCC 4AR showed to be flawed and incomplete. The primary force on temperature may well be something that is poorly modelled. In this paragraph you are implying factual evidence that does not exist. [John McLean, Australia]	Rejected. There are robust reasons for why CO2 affects temperature, and ample empirical evidence to back this up, as summarised throughout this report.
9-2113	9	81	27	81	28	Pls give gap after 30 [HASIBUR RAHAMAN, India]	Editorial. Rejected 30°N is quite commonly written without a gap.
9-2114	9	81	44	81	45	Should not one of this be the x-axis? [Government of Germany]	Accepted - this has been corrected.
9-2115	9	81	44	81	45	Not the x- and y axes are calculated, but the values on those axes. [Andreas Sterl, Netherlands]	Accepted - this has been corrected.
9-2116	9	81	51	81	53	"Scherrer et al. (2011) find no robust relationship between the ability of the CMIP3 models to represent interannual variability and their large scale transient projections." This seems like an important statement but it is not elaborated here. What is it in the transient projections that should be correlated with a realistic representation of (observed) variability? And why should the lack of such a relationship be significant? Consider explaining further. [Government of Canada]	Noted – further details are described in the paper.
9-2117	9	81	53	81	53	Pls change " Räisänen et al. (2010)" to "Räisänen et al., (2010)" and also give gap after % [HASIBUR RAHAMAN, India]	Editorial. Accepted - this has been corrected.
9-2118	9	81	56	81	57	"difficulty...is" should read "difficulties...are" [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Editorial. Accepted - this has been corrected.
9-2119	9	81	56			"main difficulty" surely large systematic errors ("biases") deserve mention within this list, especially (but not only) when they impact the mechanisms underpinning the target climatological data. [Leonard Smith, United Kingdom]	Accepted – biases mentioned in the paragraph on emergent constraints.
9-2120	9	82	12	82	14	Questions that should be asked and clarified at the end: Do correlations stay the same under climate change? Should they? Do we understand the underlying physics? [Peter Braesicke, United Kingdom]	Noted - these questions are addressed in the chapter but the chapter does not necessarily need to end with this.
9-2121	9	82	18	82	22	FAQ 9.1 : Climate models are sophisticated pieces of software...but they are primarily constructions that encapsulate our understanding of the physical (and chemical or biochemical) processes and drivers of the climate system. [Sandrine BONY, France]	taken into account. Text revised.
9-2122	9	82	18	83	24	Discussion of improvements in climate models. Emphasise these points as they are important for decision makers. [European Union]	noted.
9-2123	9	82	18	83	24	This section provides quite some text without much content and its structure could be improved. Please provide concrete information about the improvements of model (more processes, higher resolution, improved numerical methods (?), more and better obs to evaluate models. Please be aware, that the FAQs are written for non-experts and try to take their perspective when writing the text (introduction, key messages, provide	taken into account. Text revised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						information in a more logical structure). [Government of Germany]	
9-2124	9	82	18	83	24	This was by far the nicest FAQ of the set I read. It was clear, concise and clearly targeted. Changes should be limited to wordsmithing. [Peter Thorne, United States of America]	noted. Thank you.
9-2125	9	82	18	83	26	FAQ 9.1 (Model improvements): I suggest you consider adding one or more paragraphs outlining some of the major improvements made to models and the dates, which will illustrate to the reader how models certainly have improved over time in terms of moving from equilibrium to to transient models, coupling the atmosphere and the ocean, including more processes, ... [David Wratt, New Zealand]	taken into account. Text revised. However, it must be noted that model improvement is incremental, and proceeds at different rates and by different increments at different institutions, so it is difficult to generalize.
9-2126	9	82	18			FAQ 9.1: Which improvements have been made? The entire FAQ seems to us to be very vague regarding specific improvements of models. We suggest to use further examples in the text to illustrate improvements in representations of climate subsystems, physical processes, e.g., aerosols, sea level (ice sheet dynamics) etc. [Thomas Stocker/ WGI TSU, Switzerland]	taken into account. Text revised. Detail is limited owing to space constraints and the perceived audience (non-technical).
9-2127	9	82	20	82	20	What does "fidelity" mean in the context of a models? Does it not rather qualify the characteristics of human relation ships, or with pets? [Government of Germany]	noted. Definition as generally used in science and technology relates to correspondence with some reference, the accuracy with which something is reproduced (such as 'high-fidelity' sound recordings).
9-2128	9	82	20	82	20	Insert a short information on what climate models are based. E.g. insert in "...pieces of software that simulate...) the following: "... pieces of software that are based on equations which mainly refer to known physical laws or relations derived from observations. They simulate..." [Urs Neu, Switzerland]	taken into account. Text revised.
9-2129	9	82	20	82	20	Could 'extremely' be dropped from this opening sentence? [Thomas Stocker/ WGI TSU, Switzerland]	noted. We felt the adjective was important to convey to non-specialist readers that this 'software, is much more sophisticated than, say, a mobile phone "app". The code for an ESM typically comprises a million or more lines of code and is the end result of hundreds of person-years of dedicated scientific and technical effort.
9-2130	9	82	20			The wording "with as much fidelity as possible" should be changed, or at least qualified to explain what is meant by "possible". We all (I hope) believe that climate models will get better in the future, so they currently do not simulate the atmosphere with as much fidelity as is possible in an absolute sense. One could say that modelling centres develop models that reproduce climate with as much fidelity as can be achieved given limitations in personnel numbers, computing resources, understanding, and time available for model development given the need to produce projections to schedule for IPCC assessments. [Adrian Simmons, United Kingdom]	taken into account.Text revised.
9-2131	9	82	20			"as much fidelity as possible": without some qualification regarding current computational resources etc, this statement is obviously false. [Leonard Smith, United Kingdom]	taken into account.Text revised.
9-2132	9	82	21	82	21	"possible" is inappropriate. If we were already having climate models that perform as well as it is possible, there would of course be no reason to continue improving them. "currently possible", or similar. [Fedor Mesinger, United States, and Serbia]	taken into account.Text revised.
9-2133	9	82	21	82	22	Many (most) climate models do not model the ecosystem and biology, so this seems a strange up front definition [Government of Australia]	rejected. Many of the models whose results are assessed here are 'Earth System Models', as defined near the beginning of the chapter. Such models do indeed represent aspects of the terrestrial and oceanic ecosystem relevant to the carbon cycle.
9-2134	9	82	24	82	24	We suggest to not only refer to FAR (improvements using this baseline seem obvious), but also broaden this statement to include improvements made throughout all the IPCC assessment. A stronger chapeau would provide a statement referring to improvements that have been made since the AR4. [Thomas Stocker/ WGI TSU, Switzerland]	taken into account. text revised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2135	9	82	24	82	25	The argument that complexity is in any way better is silly and should not be in a report like this one. [David Webb, United Kingdom]	rejected. The ability to represent, in a more comprehensive way, the many complex interactions between components of the Earth System represents vital and important progress. This allows potentially important feedbacks to operate and correspondingly reduces the possibility that important processes or feedbacks have been ignored.
9-2136	9	82	24	82	28	I started marking up this paragraph, and found lots of things that might make it speak to lay readers a bit more smoothly, so here is a suggestion for how to redraft. "The complexity of climate models, and their completeness, has increased substantially since the first IPCC assessment in 1990, so in that sense, current Earth system models are vastly 'better' than the models of that era. Other factors have also improved models. More powerful computers allow current models to resolve much finer spatial details. Today's models also reflect the improved understanding of how climate processes work that has been obtained from the past two decades of research and analysis of more comprehensive observations," [Francis Zwiers, Canada]	taken into account. /Text revised. Thank you for the concrete suggestions.
9-2137	9	82	24	82	33	FAQ 9.1 : I would suggest to put more emphasis (by mentioning it earlier) upon the benefit of « research into various processes, etc » rather than on the increased complexity (which is not an end in itself). Moreover, the comparison of CMIP3 and CMIP5 models suggests that adding complexity (ESMs) has not changed much the performance of models (with some exceptions of course for particular phenomena), and that it has not increased considerably the spread of projections neither. [Sandrine BONY, France]	taken into account. Text revised. Complexity is not portrayed as an end unto itself, but rather as a description of the increasingly comprehensive manner in which the Earth System, and its various complex interactions and feedbacks, are treated in the latest generation of models.
9-2138	9	82	24			What is the role of numerics? Could we create methods that converge? [Peter Braesicke, United Kingdom]	noted. There is not compelling evidence that numerics are a crucial source of uncertainty or error in climate models. Not clear if comment refers to numerical convergence or convergence of model results. If the latter, parameterization of physical processes is undoubtedly more important than numerical discretization/methods used to solve equations of motion. In any case, in the interests of accessibility of the FAQ to a non-specialist audience, discussion of numerics is avoided.
9-2139	9	82	30	82	30	"Climate models of today are, in principal, better than their predecessors". What principal can possibly support this statement given the continuing large spread in climate sensitivities, the use of a large and increasing number of model parameters, and the continuing poor representation of clouds and rainfall? [David Webb, United Kingdom]	noted. The principle is that contemporary models are more comprehensive in their representation of the complex Earth-System interactions that ultimately shape the climate and its response to external forcing. The point is that this improvement "in principle" has not always translated directly into improvement in representing many important climate processes/quantities as the reviewer notes.
9-2140	9	82	30	82	33	This section is very unspecified. "Climate models of today are ... better than their predecessors". What does this mean? How do you "measure" this? What are the major improvements and what are the reasons for the improvements? [Martin Dameris, Germany]	taken into account. Text revised to provide more concrete examples.
9-2141	9	82	30	82	33	In a physical sense, surely these new interactions improve the simulation of the other components, even if in a RMS error sense it gets worse. So if the uncertainty does increase, perhaps that is a reflection of previous overconfidence not a "worse" outcome. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	noted.
9-2142	9	82	31	82	31	I would not use error here. To the lay audience error = wrong and you do not want to imply that here. Uncertainty? Complexity? [Peter Thorne, United States of America]	taken into account. Changed to 'possible error'.
9-2143	9	82	31	82	31	Insert "possible" before "error" (one does not purposefully introduce error). [Francis Zwiers, Canada]	accepted. Thank you.
9-2144	9	82	31	82	33	The reader might wonder why these potential errors are introduced. Perhaps it should be stated that new	taken into account. Text revised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						variables and interactions can potentially give a more realistic response under a changed climate. For the present climate they may replace a constraint that is tuned, so it may appear there is a reduction in performance for the present climate. [Government of Australia]	
9-2145	9	82	31			errors: possible errors and certainly uncertainties [Peter Braesicke, United Kingdom]	taken into account. Text revised.
9-2146	9	82	35	82	40	FAQ 9.1 : The fact that models are now evaluated on a much wider range of time and space scales (processes, weather, modes of variability, volcanic eruptions, long-term trends, paleo changes, simulation of the climate of other planets, etc) is key when one wants to assess the predictive capabilities of models. [Sandrine BONY, France]	taken into account. Text revised.
9-2147	9	82	35	82	40	We recommend mention of the importance of sustained satellite missions to enable accurate evaluation of climate variability. Improvement of model performance is guided by process-oriented observational data analysis. Availability of long-term observational record is a necessary prerequisite for future model improvements. [Government of United States of America]	taken into account. Text revised to note importance of observations in evaluating and guiding model improvement, but specific mention of satellite missions did not seem to fit in this FAQ.
9-2148	9	82	35	82	45	Insert between these two paragraphs - "The most important factor in the evaluation of climate models is whether they accurately include all climate factors. While the literature indicates an improvement over the last 20 years the level of scientific understanding remains poor in several areas (IPCC 4AR)." [John McLean, Australia]	taken into account. Text revised to mention limitations of scientific understanding in some areas.
9-2149	9	82	37	82	37	add „be“ between „not“ and „sufficient“ [Douglas Maraun, Germany]	editorial
9-2150	9	82	37	82	40	"Whereas...": sentence is difficult to understand [Barbara Früh, Germany]	taken into account. Text revised.
9-2151	9	82	40	82	40	Add to paragraph something like: "In addition, the climate system has - du to its partially 'chaotic' nature - an accidental component and can reach different states in similar conditions. While this can be simulated by several runs of the same model - which give different outcomes and thus different 'possible' states of the climate system - the Earth's climate only delivers the values of one possible 'real' state. Since we don't know if this 'realization' of current climate lies on the lower or the upper side of the possible range (e.g. for temperature), certain disparities of model results to observations do not necessarily mean that the model is 'wrong'. Only if the observations are out of the range of a considerable number of model simulations, the model is probably insufficient. On the other hand, good agreement of model simulations with observations is not a concluding proof of its validity, since models can be 'right', but for the wrong reason." [Urs Neu, Switzerland]	taken into account. Text revised to mention internal variability.
9-2152	9	82	44	82	44	I suggest this line be reworded as "Past reports have typically provided a rather broad survey of model performance, ASSESSING differences between model-calculated ..." (ie change "showing" to "assessing"). [David Wratt, New Zealand]	rejected -- past reports have indeed 'shown' such differences (by way of figures and accompanying text)
9-2153	9	82	47	82	52	Not only does no model emerge as "the best" overall, but "the best" as defined by these methods of "synthesizing" observations to boil them down to a "simple numerical score" is meaningless when it comes to differentiating divergent model predictions of climate change. The answer to FAQ 9.1, if "better" means more reliable predictors of climate change, is for the most part, "We don't know because we don't have any way to tell what makes for an accurate prediction of the future beyond just the broad conclusion that the climate is going to continue to get warmer." [Anthony Del Genio, United States of America]	noted. As indicated in responses to other comments by this reviewer, the remit of this chapter (and hence this FAQ) is to assess model performance in terms of their ability to reproduce *observed* aspects of the climate system. This is a necessary test of model quality. It is important to know whether models are able to simulate the climate system as we observe it, and whether that capability is improving or not. A distinct question is whether some metric of model quality, as judged against historical performance, is a reliable indicator of future performance. There is little in the way of published literature on this latter topic (as the reviewer has noted repeatedly). But it remains important to document, at least, the extent to which models can or cannot reproduce what has been observed.
9-2154	9	82	51	82	51	Suggest to replace the text in parenthesis with "(which have their own uncertainty ranges)" [Thomas	taken into account. Text revised.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Stocker/ WGI TSU, Switzerland]	
9-2155	9	82	54	82	54	I wonder if there is any value in attempting to "connect the dots" between successive generations of climate models from the same centre. This might show a contrast between established groups and those with less experience. [Francis Zwiers, Canada]	taken into account. We have discussed this at length in constructing the figure, and for many reasons this is fraught with difficulty (e.g. multiple model versions from the same institution).
9-2156	9	82	54	83	6	FAQ 9.1, Fig. 1 is too difficult. Please consider chapter 1, Fig 1.14 for an easier representation of model improvement, or part of Fig 9.45 for an easier representation of model performance. [Government of Germany]	taken into account. A new version of the figure has been constructed.
9-2157	9	82	54			FAQ 9.1 Figure 1: We appreciate it is noted that this figure is only a preliminary draft, however, the current caption does not seem well connected to the figure. There is no mention of the red/green precip/temperature components of the figure. What are the sketches that are referred to in the lower panel of the figure? Both the figure and caption will need considerable attention to be suitable for the FAQ audience. [Thomas Stocker/ WGI TSU, Switzerland]	taken into account. Unfortunately the wrong figure version was attached to the SOD (an error that was not discovered until much later). The correct version did indeed include sketches to illustrate model complexity. A new version of the figure has been constructed for the final draft.
9-2158	9	83	3	83	4	This part of the figure is probably missing from the current draft (?). Note that the model development is shown in Figure 1.14. You might want to refer to this figure as well. [Government of Germany]	taken into account. Unfortunately the wrong figure version was attached to the SOD (an error that was not discovered until much later). The correct version did indeed include sketches to illustrate model complexity. A new version of the figure has been constructed for the final draft.
9-2159	9	83	4	83	4	don' t understand "Across the bottom of the figure, the sketches indicate..."? [Barbara Früh, Germany]	taken into account. Unfortunately the wrong figure version was attached to the SOD (an error that was not discovered until much later). The correct version did indeed include sketches to illustrate model complexity. A new version of the figure has been constructed for the final draft.
9-2160	9	83	8	83	9	A more critical conclusion is required! An improvement of the model performance (grade) does not necessarily mean that you now have a better model, and also the other way around would be a wrong conclusion. A model result can be in good agreement with observation, but for the wrong reasons or because two or more errors are eliminating themselves. The cancelation of a specific model error can therefore lead to less agreement with observations, although the model system has been improved. I would like to see a brief discussion about such points. [Martin Dameris, Germany]	noted. While it is true that improved performance may result from cancellation of errors in a particular model, this is unlikely in the multi-model ensemble as a whole. The issue of model improvement leading to temporary reduction in model performance is discussed (text revised) earlier in the FAQ.
9-2161	9	83	8	83	11	This paragraph covers only model assessments up to 2005 (ie CMIP1, CMIP2 and CMIP3). Hopefully in line with draft Figure FAQ 9.1 Fig 1 you will be able to bring this up to date by also covering CMIP5 ? [David Wratt, New Zealand]	taken into account. New figure has been constructed and text updated to reflect this.
9-2162	9	83	8			Does this mean a convergence of models with observations, does it happen for the right reason? [Peter Braesicke, United Kingdom]	noted, although comment seems to refer to an incorrect line number. Text related to figure has been revised.
9-2163	9	83	9	83	9	Steady improvement in simulating the current climate, but little further insight into the accuracy of predictions of climate change. [Anthony Del Genio, United States of America]	noted.
9-2164	9	83	9	83	9	This is a bit awkward. Suggest replacing "Evaluation against a 'performance index' has shown a" with "Comparisons using such a performance metric show that there has been" [Francis Zwiers, Canada]	taken into account. Thank you for the concrete suggestion.
9-2165	9	83	10	83	11	Should CMIP5 be mentioned? Would there be any opportunity to include CMIP1 in the figure? [Francis Zwiers, Canada]	taken into account. New figure has been constructed to include CMIP5 and text updated to reflect this. For technical reasons related to availability of consistent model output, CMIP1 is not included.
9-2166	9	83	13	83	13	This is also a bit awkward - suggest replacing this sentence with "While each generation exhibits a range of	taken into account. Thank you for the concrete

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						model performance, the average model performance has improved with each generation." [Francis Zwiers, Canada]	suggestion.
9-2167	9	83	13	83	18	IPCC TAR and 4AR each showed a collection of climate forces and the levels of scientific understanding (TAR as a graphic, 4AR as a table). You have a similar Figure in chapter 8 so cross-reference it at this point. [John McLean, Australia]	taken into account, however, the performance measure shown in the figure is not sensitive to climate forcing because it is a comparison of 'climatology' rather than change.
9-2168	9	83	14	83	15	Amend "changes model performance" by "changes in model performance" [Urs Neu, Switzerland]	editorial
9-2169	9	83	18	83	18	Improvements is misspelled. [Francis Zwiers, Canada]	editorial
9-2170	9	83	18	83	18	It seems to me that there are two other points that could be made. Achieving the same spread in model performance as with previous generations, but using with more complete, and therefore less strongly constrained, models that represent process and parameter uncertainty more completely seems to be to be a big success. Very early climate models were constrained by observed SSTs and prescribed clouds, subsequently by specified deep ocean properties, subsequently by flux adjustments, then by fully flux adjusted carbon cycles (i.e., GHG concentrations were specified), etc - and yet, despite relaxing all of these constraints, model performance has slowly improved. I think that is an enormous success story. A second point is that climate modelling, over its history, has continually attracted new entrants, and it takes time for these new participants to gain experience and develop state-of-the-science models. Surely this phenomenon maintains model spread as well. [Francis Zwiers, Canada]	taken into account. Text revised to try to convey these points.
9-2171	9	83	18			... improvements (typo) [Government of France]	editorial
9-2172	9	83	20	83	21	So climate models are getting better at simulating the past, but we want to know how good they will be at simulating the future, and these are not the same thing. It would be helpful perhaps to give an idea of the space and time scales for which we expect the output of climate models to be informative about future climate. For regional precip there would be a much smaller range than for global mean temperature anomaly, for instance. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	noted. This issue is discussed to the extent possible (given existing literature on the subject) in the body of the chapter. The scope of the FAQ is more limited.
9-2173	9	83	20			delete colloquial 'so, yes', mention conceptual understanding and/of underlying physics [Peter Braesicke, United Kingdom]	noted, but direction was to write FAQs in a manner accessible to a general audience.
9-2174	9	83	21	83	23	If climate models were based on physically verifiable principles they would agree to a much greater extent than they actually do. But they are not completely based on physically verifiable principles. The fundamental result that the climate is going to warm through the 21st Century is based mostly on physically verifiable principles. So is the prediction of stratospheric cooling. So is the prediction of the component of sea level rise due to thermal expansion of ocean water. So is the qualitative idea that sea ice will decrease. To a decent extent so is the sign and general magnitude of water vapor feedback (i.e., strong as opposed to weak). Much of the rest is based on parameterization assumptions that may or may not be proven to be correct and must be regarded circumspectly at the present time. [Anthony Del Genio, United States of America]	taken into account. Text modified.
9-2175	9	83	21			please consider commenting on the improvement of the observational data-base and its effect on the improved model performance. [Thomas Stocker/ WGI TSU, Switzerland]	taken into account. This is noted explicitly in the preceding paragraph.
9-2176	9	83	23	83	23	"scientifically sound preview" sounds much too confident. I would prefer a phrase like "set of possible evolutions of the future climate, each largely consistent with our current physical understanding but limited by known model inadequacies" [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	taken into account. Text revised.
9-2177	9	83	23	83	24	"Preview of the climate to come" implies projections WILL occur. Instead, should say something like "information about possible future climates under a range of emission scenarios" to emphasise the possibility of avoiding these climates. [Josephine Brown, Australia]	taken into account. Text revised.
9-2178	9	84	1			References: Some of them appear twice. [Farahnaz Khosrawi, Sweden]	Editorial
9-2179	9	84		137		References for Table 9.1 Dunne, J. P., et al., subm.: GFDL's ESM2 global coupled climate-carbon Earth System Models. Part II: Carbon system formation and baseline simulation characteristics. Journal of Climate.	Taken into account. We have added the references to Table 9.1.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						<p>Dunne, J. P., et al., 2012: GFDL's ESM2 global coupled climate-carbon Earth System Models Part I: Physical formulation and baseline simulation characteristics. <i>Journal of Climate</i>, doi:10.1175/JCLI-D-11-00560.1.</p> <p>Dunne, J. P., A. Gnanadesikan, J. L. Sarmiento and R. D. Slater, 2010 Technical description of the prototype version (v0) of Tracers of Phytoplankton with Allometric Zooplankton (TOPAZ) ocean biogeochemical model as used in the Princeton IFMIP model, <i>Biogeosciences Supplement</i>, 7, 3593.</p> <p>Henson, S. A., J. P. Dunne, and J. L. Sarmiento, 2009: Decadal variability in North Atlantic phytoplankton blooms. <i>J. Geophys. Res.</i>, 114, C04013, doi:10.1029/2008JC005139.</p> <p>Hallberg, R., 1995: Some aspects of the circulation in ocean basins with isopycnals intersecting the sloping boundaries. Ph.D. thesis, University of Washington, 244 pp. [Available from University Microfilms, 1490 Eisenhower Place, P.O. Box 975, Ann Arbor, MI 48106.]</p> <p>Hallberg, R. W., 1997: Stable split time stepping schemes for large-scale ocean modeling. <i>J. Comp. Phys.</i>, 135, 54-65.</p> <p>Hallberg, R., and A. Adcroft, 2009: Reconciling estimates of the free surface height in Lagrangian vertical coordinate ocean models with mod-split time stepping. <i>Ocean Modelling</i>, 29, 15-26. [Jasmin John, United States of America]</p>	
9-2180	9	85	13	85	16	Duplication of Andrews et al. 2012 [SAMIR POKHREL, INDIA]	Editorial
9-2181	9	85	38	85	42	Duplication of Arora et al. 2011 [SAMIR POKHREL, INDIA]	Editorial
9-2182	9	86	36	86	39	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2183	9	87	21	87	22	The present correct reference is: Bentsen, M., I. Bethke, J. B. Debernard, T. Iversen, A. Kirkevåg, Ø. Seland, H. Drange, C. Roelandt, I. A. Seierstad, C. Hoose, and J. E. Kristjánsson: The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation. <i>Geosci. Model Dev. Discuss.</i> , 5, 2843-2931, 2012. doi:10.5194/gmdd-5-2843-2012 [Trond Iversen, United Kingdom of Great Britain & Northern Ireland]	Editorial
9-2184	9	87	21			Suggest referring to: Bennett, J. C., Ling, F. L. N., Post, D. A., Grose, M. R., Corney, S. P., Graham, B., Holz, G. K., Katzfey, J. J., and Bindoff, N. L., 2012: High-resolution projections of surface water availability for Tasmania, Australia, <i>Hydrol. Earth Syst. Sci.</i> , 16, 1287-1303, doi:10.5194/hess-16-1287-2012. [Government of Australia]	Rejected. To include another reference requires a scientific justification for changing the assessment text, not just a suggestion for extending the bibliography.
9-2185	9	87	38			Suggest referring to: Nguyen, K. C., Katzfey, J. J., & McGregor, J. L., 2012: Global 60 km simulations with CCAM: evaluation over the tropics. <i>Climate Dynamics</i> , 39(3), 637-654. [Government of Australia]	Rejected. To include another reference requires a scientific justification for changing the assessment text, not just a suggestion for extending the bibliography.
9-2186	9	88	3	88	10	Duplication of reference Boe et al 2009a and 2009c [SAMIR POKHREL, INDIA]	Editorial
9-2187	9	89	30			Please change to JN Brown instead of J Brown, to distinguish from other J Brown papers. [Jaclyn Brown, Australia]	Editorial
9-2188	9	90	28	90	31	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2189	9	92	16	92	19	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2190	9	92	16	92	25	(Collins et al 2012a) is identical to (Collins et al 2012b) change references [Barbara Früh, Germany]	Editorial
9-2191	9	92	20	92	25	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2192	9	92	57	92	57	Incomplete reference [SAMIR POKHREL, INDIA]	Editorial - Duplication of reference, deleted

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2193	9	93	3	93	7	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2194	9	94	22			We believe the reference should be 2012 not 2011 (Delworth et al.). [Government of United States of America]	Editorial
9-2195	9	95	42	95	45	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2196	9	96	33	96	36	- (Eyring et al 2012a) and (Eyring et al 2012b) are identical - change reference on p.29 I.23, I.36, I.40, I.48; p.30 I.9 [Barbara Früh, Germany]	Editorial
9-2197	9	96	33	96	36	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2198	9	96	43	96	43	Please change “Fangli, Q.” into “Qiao, F.” [Bin Wang, China]	Editorial
9-2199	9	97	25	97	28	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2200	9	97	39	97	40	- (Fischer et al. 2007a) is identical to (Fischer et al. 2007b); skip lines p.97 I. 39-40 since in line 41-42 the entry is more complete; - change reference on p.11 I.6 and p.41 I.41 to (Fischer et al. 2007) [Barbara Früh, Germany]	Editorial
9-2201	9	97	39	97	42	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2202	9	98	27	98	27	Add references: Frederiksen, C.S., J.S. Frederiksen, J.M. Sisson and S.L. Osbrough, 2011a: Australian winter circulation and rainfall changes and projections. <i>Int. J. Clim. Change Strat. Mang.</i> , 3, Issue 2, 170-188. Frederiksen, C.S., J.S. Frederiksen, J.M. Sisson and S.L. Osbrough, 2011b: Changes and Projections in the Annual Cycle of the Southern Hemisphere Circulation, Storm Tracks and Australian Rainfall. <i>Int. J. Clim. Change Impacts Responses</i> , 2, 143-162. Frederiksen, J.S., C.S. Frederiksen, S.L. Osbrough and J.M. Sisson, 2010: Causes of changing Southern Hemispheric weather systems. Chapter 8, <i>Managing Climate Change</i> , Eds. I. Jupp, P. Holper and W. Cai, CSIRO publishing, pp85-98. [Jorgen Frederiksen, Australia]	Taken into account - The papers have been assessed and their findings included as appropriate.
9-2203	9	99	48	99	50	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2204	9	101	46			correct reference: <i>Clim. Dyn.</i> , 39, 2611-2629, doi: 10.1007/s00382-011-1228-5 [Andreas Sterl, Netherlands]	Editorial
9-2205	9	102	26	102	29	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2206	9	102	39	102	47	Duplication of reference Holland et al 2012a and 2012b [SAMIR POKHREL, INDIA]	Editorial
9-2207	9	103	1	103	9	Among those 4 reference to Hourdin et al... : 1) the second one must be removed (From LMDZ5A to LMDZ5B...). It correspond to a first draft of the first reference (LMDZ5B...). 2) The DOI of the last one (line 9) is wrong. It should be : 10.1007/s00382-012-1411-3 [Frédéric HOURDIN, France]	Editorial
9-2208	9	103	18	103	18	In the Hunke reference, "Retorspective" --> "Retrospective". Note that the reference below is the same, with the same typo. [François Massonnet, Belgium]	Editorial
9-2209	9	103	18	103	21	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2210	9	104	11	104	12	The present correct reference is: Iversen, T., M. Bentsen, I. Bethke, J. B. Debernard, A. Kirkevåg, Ø. Seland, H. Drange, J. E. Kristjánsson, I. Medhaug, M. Sand, and I. A. Seierstad: The Norwegian Earth System Model, NorESM1-M – Part 2: Climate response and scenario projections. <i>Geosci. Model Dev. Discuss.</i> , 5, 2933-2998, 2012. doi:10.5194/gmdd-5-2933-2012 [Trond Iversen, United Kingdom of Great Britain & Northern Ireland]	Editorial
9-2211	9	104	31	104	35	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2212	9	105	1	105	2	Incomplete reference [SAMIR POKHREL, INDIA]	Editorial
9-2213	9	105	9	105	14	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2214	9	106	46	106	47	The present correct reference is: Kirkevåg, A., T. Iversen, Ø. Seland, C. Hoose, J. E. Kristjánsson, H. Struthers, A. M. L. Ekman, S. Ghan, J. Griesfeller, E. D. Nilsson, and M. Schulz: Aerosol-climate interactions in the Norwegian Earth System Model – NorESM. Geosci. Model Dev. Discuss., 5, 2599-2685, 2012.doi:10.5194/gmdd-5-2599-2012 [Trond Iversen, United Kingdom of Great Britain & Northern Ireland]	Editorial
9-2215	9	107	11	107	14	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2216	9	107	11	107	15	The three Klocke references may all refer to the one paper: Klocke, Daniel, Robert Pincus, Johannes Quaas, 2011: On Constraining Estimates of Climate Sensitivity with Present-Day Observations through Model Weighting. J. Climate, 24, 6092–6099. doi: http://dx.doi.org/10.1175/2011JCLI4193.1 [Anthony Hirst, Australia]	Editorial - duplication of reference
9-2217	9	109	53	110	9	There is some confusion and errors in the representation of the Li et al. references, of which there are 4 that might be relevant. NOTE all were submitted or in press by the deadline. I) CMIP5 & CMIP3 RADIATION EVALUATION (and secondarily its connection to model representation of clouds and precipitation)----- Li, J.-L. F., D. E. Waliser, G. Stephens, S. Lee, T. S. L'Ecuyer, S. Kato, and N. Loeb (2012), Characterizing and Understanding Radiation Budget Biases in CMIP3/CMIP5 GCMs, Contemporary GCMs and Reanalyses, Journal of Geophysical Research, Submitted with revisions. II) CMIP5 & CMIP3 CLOUD ICE EVALUATION ----- Li, J.-L., D. E. Waliser, W. T. Chen, B. Guan, T. Kubar, G. Stephens, H-Y Ma, D. Ming, L. Donner, C. Seman, and L. Horowitz (2012), An Observation-Based Evaluation of Cloud Ice Water in CMIP3 and CMIP5 GCMs and Contemporary Analyses, Journal of Geophysical Research, In Press. III) CMIP5 & CMIP3 CLOUD LIQUID EVALUATION ----- Li, J.-L., D. E. Waliser, S. Lee, B. Guan, G. Stephens, M. Christensen, J. Teixeira, and H.-Y. Ma (2012), Assessment and Comparison of Cloud Liquid Water in CMIP3, CMIP5 GCMs and Contemporary GCMs and Analyses as well as Observations Journal of Geophysical Research, Submitted with Revisions. IV) CMIP3 CLOUD LIQUID EVALUATION ----- Li, J.-L. F., D. E. Waliser, and J. H. Jiang (2011), Correction to "Comparisons of satellites liquid water estimates to ECMWF and GMAO analyses, 20th century IPCC AR4 climate simulations, and GCM simulations", Geophysical Research Letters, 38(24), doi:10.1029/2011GL049956. The first author's contact info is: "Jui-Lin (Frank) Li" <Juilin.F.Li@jpl.nasa.gov> [duane waliser, United States of America]	Editorial.
9-2218	9	110	4	110	9	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2219	9	111	56	111	57	The Manzini et al is correct for the text at page 14. But the Manzini et al reference for CMCC model reference (Tab 9.1) is an additional one: Manzini, E., C. Cagnazzo, P. G. Fogli, A. Bellucci, and W. A. Müller (2012), Stratosphere-troposphere coupling at inter-decadal time scales: Implications for the North Atlantic Ocean, Geophys. Res. Lett., 39, L05801, doi:10.1029/2011GL050771. [Elisa Manzini, Germany]	Editorial
9-2220	9	113	1	113	6	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2221	9	113	1	113	8	It is very hard to tell whether or to what extent the models and observations overlap in these figures because the model results appear to be overwriting some of the observations. [Melissa Free, United States of America]	Taken into account. Figures have been modified.
9-2222	9	113	7	113	12	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2223	9	114	35	145	35	Wrong reference (wrong name and missing page number). The right reference is: Salas-Melia, D., 2002: A global coupled sea ice-ocean model. Ocean Modelling 4, 137-172. [DAVID SALAS Y MELIA, France]	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2224	9	117	22	117	22	correct reference [Barbara Früh, Germany]	Taken into account - The correct reference should be: Oleson, K.W., Dai, Y., Bonan, G., Bosilovich, M., Dickinson, R., Dirmeyer, P., Hoffman, F., Houser, P., Levis, S., Niu, G.-Y., Thornton, P., Vertenstein, M., Yang, Z.-L., and Zeng, X. 2004. Technical description of the Community Land Model (CLM). NCAR Technical Note NCAR/TN-461+STR. National Center for Atmospheric Research, Boulder, CO, 173 pp.
9-2225	9	117	22	117	24	Incomplete reference [SAMIR POKHREL, INDIA]	Editorial, combine with 9-2224
9-2226	9	117	30	117	33	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2227	9	117	41	117	44	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2228	9	119	3	119	8	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2229	9	119	15	119	18	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2230	9	120	39	120	43	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2231	9	120	49	120	52	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2232	9	121	18	121	21	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2233	9	121	48	121	50	Correct reference is: Rotstayn, L. D., Jeffrey, S.J., Collier, M.A., Dravitzki, S.M., Hirst, A.C., Syktus, J.I., Wong, K.K., 2012: Aerosol- and 49 greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: a 50 study using single-forcing climate simulations. Atmos. Chem. Phys., 12, 6377-6404. [Mark Collier, Australia]	Editorial
9-2234	9	122	47	122	50	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2235	9	122	49			Duplicate reference [Karen Shell, United States of America]	Editorial
9-2236	9	123	6	123	20	Duplication of reference Santer et al 2012 (line 6) and Santer at al 2012 (line 19) [SAMIR POKHREL, INDIA]	Editorial
9-2237	9	123	8	123	9	- (Santer et al 2009a) is identical to (Santer et al 2009b) on p.123 l.16-18; skip one of them and check if it is really necessary to cite a non-reviewed conference abstract in this context; see also (Santer et al 2012) - change reference on p.16 l.1, p.18 l.41 and p.19 l.15 [Barbara Früh, Germany]	Editorial
9-2238	9	127	1	127	5	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2239	9	127	9	128	9	Paper is now published. Full reference: Stephenson, D.B., Collins, M., Rougier, J. C. and Chandler, R.E. (2012), Statistical problems in the probabilistic prediction of climate change. Environmetrics, 23: 364–372. doi: 10.1002/env.2153 [David Stephenson, United Kingdom of Great Britain & Northern Ireland]	Editorial
9-2240	9	127	35	127	38	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2241	9	128	16			Please also the reference on Page 9-128, Line 16 to the following: Su, H., W.G. Read, J.H. Jiang, J.W. Waters, D.L. Wu, and E.J. Fetzer, "Enhanced positive water vapor feedback associated with tropical deep convection: New evidence from Aura MLS," Geophys. Res. Lett. 33, L05709, doi:10.1029/2005GL025505, 2006a. Su, H., D.E. Waliser, J.H. Jiang, J-L. Li, W.G. Read, J.W. Waters, and A.M. Tompkins, "Relationships of upper tropospheric water vapor, clouds and SST: MLS observations, ECMWF analyses and GCM simulations," Geophys. Res. Lett. 33, L22802, doi:10.1029/2006GL027582, 2006b [Government of United States of America]	Editorial
9-2242	9	129	14	129	19	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2243	9	130	11	130	12	The present correct reference is: Tjiputra, J. F., C. Roelandt, M. Bentsen, D. M. Lawrence, T. Lorentzen, J.	Editorial

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Schwinger, Ø. Seland, and C. Heinze: Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM). Geosci. Model Dev. Discuss., 5, 3035-3087, 2012. doi:10.5194/gmdd-5-3035-2012. [Trond Iversen, United Kingdom of Great Britain & Northern Ireland]	
9-2244	9	131	14	131	18	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2245	9	133	20	133	21	The reference is incomplete. I. G. Watterson and P. H. Whetton (2011) Distributions of decadal means of temperature and precipitation change under global warming. JGR, 116, D07101, doi:10.1029/2010JD014502 [Ian Watterson, Australia]	Editorial
9-2246	9	133	47	133	52	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2247	9	134	52	134	55	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2248	9	135	3	135	6	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2249	9	136	4	136	9	Duplication of reference [SAMIR POKHREL, INDIA]	Editorial
9-2250	9	138	0			Table 9.1: This table would be more use if it could include additional separation of components and perhaps be rearranged. For example, there is no column for land cover/vegetation so that we could see whether or not a DGVM is included, and there is no column for model top/stratosphere so that high top/low top models can be differentiated. It would help if the entries could be arranged vertically in section according to complexity, e.g. atmosphere-only, AOGCM low top, AOGCM high top, ESM low top, ESM high top, or something similar. This would give the reader a much better idea of what the spread in complexity is within the CMIP5 ensemble, which will inform their interpretation of the results. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Noted. We understand the several comments made here. We have to balance the amount of information we put in the table and the readability. The information on land cover vegetation should be found from the references provided for the Land Surface. Grid top is included in the atmosphere section. The fact that there is no clear cut between AOGCM and ESM makes the separation in the table difficult. We added a simplified table to ease readability.
9-2251	9	138	0			Table 9.1: the Land Ice column is misleading - it does not appear to differentiate models which include ice sheet modelling from those that don't, and some of the ones that say "(1) Land ice" don't seem to have ice sheet models (e.g. HadGEM2-ES has an adjustment freshwater flux applied to take account of the fact that it does not have an ice sheet model). [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Taken into account - table clarified
9-2252	9	138		138		Table9.1: About aerosols, I would suggest listing aerosol species considered, since different aerosols have different climatic effects. Also need to distinguish aerosol direct or indirect effect. Or it can mention if there are other places in this report have such information. [Bin Wang, China]	Rejected - We have to balance the amount of information we put in the table and the readability. This information can be found on the online documentation associated with the ESG data distribution system (metadata viewer).
9-2253	9	138		156		In Table 9.1, the distinction "low or high top" is mentioned but not defined. A reference to Wilcox et al. (2012) would provide a precise definition that's consistent with qualitative discussion in the text (Section 9.1.3.2.7). Using this definition, color or typeface could highlight the "high top" models. Similar techniques could mark other important distinctions between models, e.g. full Earth system models including biogeochemistry vs. physical climate system models. [Curt Covey, United States of America]	Taken into account - table clarified. We added a simplified table to ease readability.
9-2254	9	138		156		Suggest to evaluate the independence of models in Tab.9.1. There are 56 models in CMIP5, but only 23 models are independent. So, the reliability assessment of detection, attribution, prediction and projection of the model results will not be robust or have higher reliability than the AR4. At the same time, it is recommended that the description of model details in Table 9.1 should be specification, such as the content involving aerosol should give the direct and indirect climate effects. [Lei Huang, China]	Rejected - even though we agree independence is a real issue, there is no agreed protocol to establish the degree if independence of models. This is still an active area of research
9-2255	9	138		156		Table 9.1 More accurate and more abundant information about each model should be added to this table so that CMIP5 data users are well informed concerning the possible deficiencies of a specific model. [Weiping Li, China]	Rejected - We have to balance the amount of information we put in the table and the readability. More information can be found on the online documentation associated with the ESG data distribution system (metadata viewer). We added a simplified table to ease readability.
9-2256	9	138		156		There is a lot of really important information in here but this table quite frankly would scare the proverbial out	Taken into account. We added a simplified table to

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						of me as a lay reader. I would place this table in an appendix and produce a greatly simplified version of this which presents the key laymans information content only in the main text and then refer the reader to the appendix for both the detailed information and the references. I think with a little thought you could make a simplified table for the main text that communicated the truly essential component information without at the same time scaring the pants off the readers straight up. [Peter Thorne, United States of America]	ease readability.
9-2257	9	138				<ul style="list-style-type: none"> - Table 9-1 header, column 3: "aerosols" is wrong, it is either "aerosol" or "aerosol particles" - header column 8: to print "ocean" in bold font is misleading since it is identical to column 9 header; suggestion: print "marine biogeochemistry" in bold letters - "flux corction implementation" must be desribed - modelID BNU-ESM: the entry for the main references is missing for all components - modelID EC-Earth: the entry for the main references is missing for several components; "Europe" is no institution but a continent - there are many more models with missing references, please, insert them;if there are no then rethink the participation of the models since it is no good scientific practise to publish results from unpublished models [Barbara Früh, Germany] 	Taken into account - table clarified following comments. More up to date information can be found on the online documentation associated with the ESG data distribution system (ES-DOC metadata viewer).
9-2258	9	138				Table 9.1 Caption of Table 9.1, "flux correction implementation (not yet decribed)" needs to be updated. If no model uses flux correction, no need to include that item. On the use of "grid top". It is more commonly referred to as "model top". Any reason having to use "grid top" instead? [Government of United States of America]	Taken into account - table clarified following comments and item removed
9-2259	9	138				Table 9.1. Although it provides a comprehensive set of references to each model, this Table doesn't allow one to pick out key differences between the models' features at a glance, apart from grid choices, resolution and some other structural things. The fact that it is so huge is not helpful in this respect. I'm hoping that a more user-friendly synthesis/presentation of the information will be possible for the Final Draft. [Tim Johns, United Kingdom of Great Britain & Northern Ireland]	Taken into account. We added a simplified table to ease readability.
9-2260	9	138				Table 9.1. The column order (alphabetical) doesn't seem the most natural choice. A better order might be Atmosphere, Aerosols, Atmos Chemistry, Land Surface, Land Ice, Ocean, Ocean Biogeochemistry, Sea Ice. [Tim Johns, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Table modified
9-2261	9	138				Table 9.1: Column 2 and 3 for ACCESS1.3 the reference is not complete [HASIBUR RAHAMAN, India]	Editorial. Corrected
9-2262	9	139				Table 9.1: Third column, second and third row: text misplaced. [Farahnaz Khosrawi, Sweden]	Editorial
9-2263	9	139				Table 9.1: If there are no reference than just write "reference missing" [Farahnaz Khosrawi, Sweden]	Noted. Text modified
9-2264	9	139				Table 9.1 Three references about AVIM, the basis of BCC_AVIM1.0 are recommended to be enlisted in the reference. (1) Ji, J., 1995: A climate-vegetation interaction model: Simulating physical and biological processes at the surface. J Biogeogr, 22: 2063-2069 (2) Lu Jianhua, Jinjun Ji, 2006: A simulation and mechanism analysis of long-term variations at land surface over arid/semi-arid area in north China, J. Geophys. Res., 111, D09306, doi:10.1029/2005JD006252. (3) Ji, J., M. Huang, and K. Li, 2008: Prediction of carbon exchange between China terrestrial ecosystem and atmosphere in 21st century. SCIENCE CHINA Earth Sciences, 51(6): 885-898. [Weiping Li, China]	Taken into account. References added.
9-2265	9	140		140		<p>Please correct the information of BNU-ESM:</p> <ul style="list-style-type: none"> 1) Column 7: "Dai and Trenberth, 2004" is an incorrect reference paper. The references here should be "Dai et al., 2004; Dai et al., 2003." 2) Line 44 on Page 9-93: Please change "2004b" into "2004" [Bin Wang, China] 	Editorial
9-2266	9	140				Table 9.1: CanESM2 and CanCM4 the Ocean Componenet name is given same as column title, Please give the actual Ocean componenet name [HASIBUR RAHAMAN, India]	Editorial
9-2267	9	144		144		Table 9.1, correct references for CMCC-CESM: Fogli et al., 2009; Vichi et al., 2011; Cagnazzo et al.,submitted (see Chapter 12) [Chiara Cagnazzo, Italy]	Taken into account. References added.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2268	9	145	1	145	1	EC-Earth entry: Main reference (column 2) is Hazeleger et al (2011). Reference to HTESEL (column 7) is: Balsamo, G., P. Viterbo, A. Beljaars, B.J.J.M. van den Hurk, A. Betts en K. Scipal (2009), A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System ; J. Hydrometeor., 2009, 10, 623-643, doi:10.1175/2008JHM1068.1. [Bart Van den Hurk, Netherlands]	Editorial
9-2269	9	145		145		References are missing for the ocean and sea ice components of EC-Earth. [Thierry Fichefet, Belgium]	Taken into account. References added.
9-2270	9	145		145		Wrong reference in the sea ice column for CNRM-CM5 model. Please replace Melia (2002) with Salas-Melia (2012) [DAVID SALAS Y MELIA, France]	Taken into account. Reference corrected.
9-2271	9	145				Table 9.1: Reference for CSIRO-Mk3-6-0 should be Rotstayn et al. (2012), not Rotstayn (2012). It is correct in the reference list. [Leon Rotstayn, Australia]	Taken into account. Reference corrected.
9-2272	9	146	1	146	1	The main reference about FGOALS-g2 should be: Li et al., 2012: The Flexible Global Ocean-Atmosphere-Land System Model: Grid-point Version 2: FGOALS-g2, Adv. Atmos. Sci., Accepted and in press. [Tianjun Zhou, China]	Taken into account. Reference corrected.
9-2273	9	146		146		In Table 9.1, the information of FGOALS-g2 is incomplete. The main references in column 2, 4 and 9 should be Li et al., 2012a, Li et al., 2012b/Wang et al., 2004, and Liu et al., 2012, respectively. The fifth choice in column 9 is η -coord (or eta co-ord). Also, some of these references are not included in the reference list. Please add them to the reference list: Li, L. J., et al, 2012a: The Flexible Global Ocean-Atmosphere-Land System Model version g2. Adv. Atmos. Sci. doi: 10.1007/s00376-012-2140-6. Li L. J., et al, 2012b: Development and Evaluation of Grid-point Atmospheric Model of IAP LASG, Version 2.0 (GAMIL 2.0), Adv. Atmos. Sci. (revised). Liu, H. L., P. F. Lin, Y. Q. Yu, and X. H. Zhang, 2012: The baseline evaluation of LASG/IAP Climate system Ocean Model (LICOM) version 2.0. Acta Meteorologica Sinica,26(3), 318-329. [Bin Wang, China]	Taken into account. References corrected.
9-2274	9	146		146		Please correct the information of FIO-ESM v1.0. 1) Column 7: please add the following paper (Bao et al, 2012). Bao Y., Qiao F. L., Song Z. Y. , 2012: Historical simulation and twenty-first century prediction of oceanic CO2 sink and pH change, Acta Oceanologica sinica, 31(5), 87-97. 2) Column 8: please change “(1) POP2.0” into “(1) Modified POP2.0 through incorporating the non-breaking surface wave-induced mixing”. 3) Column 8: please change “ (2) 320*384” into “(2) 1.125 degree in longitude, 0.27-0.64 degree variable in latitude”. 4) Column 8: please change “(4) 5 m” into “(4) 10 m thick with surface variables at 5 m”. 5) Column 8: please add the following paper (Huang et al, 2012). Huang C. J., Qiao F. L., Qi S. Song Z. Y., 2012: Evaluating austral-summer mixed layer response to surface wave-induced mixing in the Southern Ocean---- summer mixed layer in the Southern Ocean, J. Geophys. Res., 117, C00J18, doi:10.1029/2012JC007892. [Bin Wang, China]	Taken into account. Reference corrected.
9-2275	9	146		146		Please add more information of FGOALS-s2: 1) Column 4: please add the following references (Bao et al., 2010; Liu et al., 2012) Bao, Q., et al., 2010: An Introduction to the Coupled Model FGOALS1.1-s and Its Performance in East Asia,	Taken into account. References corrected.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						<p>Adv. Atm. Sci., 27(5), 1131-1142. Liu, Y. M., et al., 2012: Seasonal evolution of the subtropical anticyclones in a climate system model FGOALS-s2, Adv. Atm. Sci. (Accepted)</p> <p>2) Column 7: change the land surface component name with "CLM3; VEGAS2" and add the references (Zeng et al., 2005; Wang et al., 2012)</p> <p>Zeng, N., et al., 2005: Terrestrial mechanisms of interannual CO2 variability. Global Biogeochemical Cycles, 19, GB1016, doi:10.1029/2004GB002273. Wang, J., et al., 2012: The Earth System Model FGOALS-s2: Coupling a dynamic global vegetation and terrestrial carbon model with the physical climate system model, Adv. Atm. Sci. (Revised) [Bin Wang, China]</p>	
9-2276	9	147	0	147	0	<p>There are multiple errors or omissions in the Table 9.1 Entries for GFDL-ESM2G and GFDL-ESM2M. The vintage year for ESM2G is 2012. The main references for both should be Dunne et al., 2012a, J. Climate, doi:10.1175/JCLI-D-11-00650.1 and Dunne et al., 2012b, J. Climate, doi:10.1175/JCLI-D-12-00159.1. The atmosphere column should be the same as for gfdl-cm2.1. The Ocean Biogeochem. model name is TOPAZ, ref, Dunne et al., 2012b. The land model entry should be copied from GFDL-CM3. [Robert Hallberg, United States of America]</p>	Taken into account. References corrected.
9-2277	9	147	0	147	0	<p>The GFDL-ESM2G entry in Table 9.1 has several errors in its "Ocean" column. (1) The model name is GOLD (5) The vertical coordinate is "Isopycnic", and (7) a primary reference for the code is Hallberg, R., and A. Adcroft, 2009, Ocean Modeling, 29, 15-26, doi:10.1016/j.ocemod.2009.02.008 and for this configuration it is Dunne et al., 2012a, J. Climate, doi:10.1175/JCLI-D-11-00650.1. [Robert Hallberg, United States of America]</p>	Taken into account. References corrected.
9-2278	9	147	0	147	0	<p>The GFDL-ESM2M entry in Table 9.1 has several errors in its "Ocean" column. (1) The model name is MOM4.1 (5) The vertical coordinate is "z*", and (7) a primary reference for the code is Griffies, S. M., 2009: Elements of MOM4p1. GFDL Ocean Group Tech. Rep. No. 6, NOAA/GFDL. and for this configuration it is Dunne et al., 2012a, J. Climate, doi:10.1175/JCLI-D-11-00650.1. [Robert Hallberg, United States of America]</p>	Taken into account. References corrected.
9-2279	9	147	0	147	0	<p>The GFDL-CM3 entry in Table 9.1 has several errors in its "Ocean" column. (1) The model name is MOM4.1 (5) The vertical coordinate is "z*", and (7) the best reference is Griffies, S. M., et al., 2011: The GFDL CM3 Coupled Climate Model: Characteristics of the ocean and sea ice simulations. J. Climate, 24, doi:10.1175/2011JCLI39641. [Robert Hallberg, United States of America]</p>	Taken into account. References corrected.
9-2280	9	147	0	147	0	<p>The vintage year for gfdl-cm2p1 in Table 9.1 should be 2006. [Robert Hallberg, United States of America]</p>	Accepted.
9-2281	9	147		147		<p>Table 9.1 Salient features of the AOGCMs and ESMs participating in CMIP5. Add references for GFDL-ESM2M, GFDL-ESM2G. GFDL-ESM2M and GFDL-ESM2G Column 2 (Main reference): Add Dunne et al., 2012 and Dunne et al., subm. Column 4 (Atmosphere): Horizontal grid: 2.5 degree longitude, 2 degree latitude Vertical grid: 24 Grid Top: midpoint of top box is 3.65 hPa Column 7 (Land Surface): References: Milly and Shmakin, 2002; Shevliakova et al, 2009; Donner et al., 2011 Column 8 (Ocean Biogeochemistry): Component Name: TOPAZ2 References: Dunne et al. subm., Dunne et al, 2010, Henson et al., 2009 Column 10 (Sea Ice): Reference: Winton, 2000.</p> <p>GFDL-ESM2G (as for ESM2M, except for ocean component). Column 2 (Main reference): Add Dunne et al., 2012 and Dunne et al., subm. Column 4 (Atmosphere): Horizontal grid: 2.5 degree longitude, 2 degree latitude Vertical grid: 24 Grid Top: midpoint of top box is 3.65 hPa Column 7 (Land Surface): References: Milly and Shmakin, 2002; Shevliakova et al, 2009; Donner et al., 2011</p>	Taken into account. References added.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						Column 8 (Ocean Biogeochemistry): Component Name: TOPAZ2 References: Dunne et al. subm.; Dunne et al, 2010; Henson et al., 2009 Column 9 (Ocean) References: Hallberg, 1995; Hallberg, 1997; Hallberg and Adcroft, 2009 Column 10 (Sea Ice): Reference: Winton, 2000. [Jasmin John, United States of America]	
9-2282	9	147				Table 9.1: Same here, at several columns/rows [Farahnaz Khosrawi, Sweden]	Editorial
9-2283	9	147				Table 9.1: The reference to Delworth et al. (2006) is appropriate for the GFDL CM2.1 model but a reference to Dunne et al. (2012) should be added for ESM2G/ESM2M models: Dunne, John P., and Coauthors, 2012: GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics. J. Climate, 25, 6646–6665. doi: http://dx.doi.org/10.1175/JCLI-D-11-00560.1 [RYM MSADEK, United States of America]	Taken into account. References added.
9-2284	9	147				Table 9.1: GFDL the Ocean Component name is given same as column title, Please give the actual Ocean component name [HASIBUR RAHAMAN, India]	Editorial-refer to comment 9-2277, 9-2278, 9-2279
9-2285	9	147				ESM2M and ESM2G table entries are incorrect. (1) The main reference should be Dunne et al. 2012 (part 1) and Dunne et al. 2012 (Part 2) for both models. (2) The atmospheric details are the same as those for CM2.1 (lower in the table) The reference for the atmospheric component is Delworth et al. 2006 for 2M, 2G and CM2.1.(3)TOPAZ is the name for the biogeochemical model in both 2M and 2G. The reference is Dunne et al 2012 (Part 2). (4) the sea ice reference should be Delworth et al. 2006 for 2M, 2G and CM2.1. If you have questions, contact me (Stouffer). [Ronald Stouffer, United States of America]	Taken into account. References added.
9-2286	9	148				Table 9.1, Row 6 and 7, column 4 and 5: A number (2) appears without any text following. [Farahnaz Khosrawi, Sweden]	Editorial
9-2287	9	149				Table 9.1: Same at column 7, 8, 10, row 2 and 3. [Farahnaz Khosrawi, Sweden]	Editorial
9-2288	9	150				Table 9.1: Same here, plus text misplaced (row 5, 7, 9). [Farahnaz Khosrawi, Sweden]	Editorial
9-2289	9	151				Table 9.1: same here, column 6, row 3 and column 10, row 2. [Farahnaz Khosrawi, Sweden]	Editorial
9-2290	9	151				Table 9.1: In IPSL-CM5A-LR,Column 9 top model level is shown as 0 (zero) is it correct ? If not please give the actual depth [HASIBUR RAHAMAN, India]	Editorial
9-2291	9	152		153		Table 9.1, row for MPI-ESM-P, column for Ocean, (2/horizontal resolution) should be: average 1.5 deg GR15 [European Union]	Accepted.
9-2292	9	153				Table 9.1: In MRI-ESM1,Column 9 top model level is shown as 0 (zero) is it correct ? If not please give the actual depth [HASIBUR RAHAMAN, India]	Editorial
9-2293	9	154				Table 9.1: sam here, last row. [Farahnaz Khosrawi, Sweden]	Editorial
9-2294	9	154				Table 9.1: In MRI-CGCM3,Column 9 top model level is shown as 0 (zero) is it correct ? If not please give the actual depth [HASIBUR RAHAMAN, India]	Editorial
9-2295	9	155				Table 9.1: same here, row 2 and row 10 [Farahnaz Khosrawi, Sweden]	Editorial
9-2296	9	157		159		This information feels more appropriate for an appendix than inclusion in the main text. It is more about an audit trail than providing headline conclusions [Peter Thorne, United States of America]	Rejected - The overview of the observations used to evaluate the models is important information that should remain in main part of the text.
9-2297	9	157				table 9-2: it would be good if the table would also contain the time period of availability of the data set [Barbara Früh, Germany]	Noted - the time period is either noted in the figures or indicated to be what is available for a particular dataset.
9-2298	9	157				Table 9.2: This table is very good as an overview which observational datasets were used. The different datasets however cover rather different timeperiods. It is not tracable, which version of a dataset was used, neither which time period was analyzed. I therefore highly recommend to add additional columns on a) version	Noted - the time period is either noted in the figures or indicated to be what is available for a particular dataset.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						of dataset, b) time period covered for analysis in this chapter [Alexander Loew, Germany]	
9-2299	9	157				Table 9.2: No land variables are analyzed at all. Why? [Alexander Loew, Germany]	Accepted - land variables added to Section 9.4.4
9-2300	9	158				Table 9.2 It is not clear to me, why CERES and ERBE are used for the radiation fluxes, but ISCCP is not mentioned at all. ISCCP provides by far the longest record. It looks a bit like the data was chosen arbitrarily. Please clarify. [Alexander Loew, Germany]	Noted - ERBE provides a more accurate measurement of TOA fluxes than ISCCP which is more important than the time period of availability.
9-2301	9	160		161		I hope that all the models will be included in this table! [Julia Hargreaves, Japan]	Taken into account -- combined with comment 9-2303.
9-2302	9	160				Skips from Table 9.2 to 9.4. Where is 9.3? [Peter Thorne, United States of America]	Accepted -- Table numbering has been corrected.
9-2303	9	161	1	161	1	Table 9.4: Why can't the other models be included here? Andrews et al was published very early in the CMIP5 process and misses many models. If Andrews have not updated their table, the authors should reach out to the other modelling groups and add their results to this table. [Gavin Schmidt, United States of America]	Accepted -- Estimates from other published literature have been included to represent a larger segment of the CMIP5 models.
9-2304	9	161	1	161	5	Why not attempting to include biosphere feedbacks as a lumped estimate [European Union]	Rejected -- These feedbacks are considered and enumerated in chapter 6.
9-2305	9	161		162		Table 9.4 lists the ECS of GFDL AM2.1 (via. ESM2G or ESM2M) as 2.4C, while the published value based on a slab model is 3.4C (Stouffer et al J Clim, 2006). This is a 30% difference! Is this a typo, a calculation error somewhere, or is the approximation method not so good? This affects Fig. 9.43b also. If 3.4 C were used for the GFDL models ESM2M and ESM2G the correlation in the figure would be somewhat worse. [Government of United States of America]	Noted -- For consistency, equilibrium climate sensitivity (ECS), and transient climate response (TCR) are taken from (Andrews et al. 2012b) and (Forster et al., 2013) for all the models listed.
9-2306	9	161				Table 9.4: use a "-" instad of "n.a". This would make the table more concise. [Farahnaz Khosrawi, Sweden]	Editorial -- notation for missing information used in prior versions of the table has been retained.
9-2307	9	161				Table 9.4 should be 9.3, since Table 9.2 does not exist. [HASIBUR RAHAMAN, India]	Taken into account -- combined with 9-2302
9-2308	9	163		165		Similar to my comment on Table 9.1, this is so information dense I wonder whether it would be better in an appendix with a simplified version more accessible to the non-expert being prepared for the main text here. [Peter Thorne, United States of America]	Rejected -- This table contains the data plotted in 9-44 and hence belongs in the main body of the text.
9-2309	9	167				Figure 9-1: the figure is a too short summary for the description of the CMIP5 runs; either explain it in more detail here or give reference to the section where these single runs are explained and discussed in more detail, perhaps a table. [Barbara Früh, Germany]	Accepted -- the explanation for this figure and the historical experiments has been extended in Section 9.3.2.1 Experiments utilized for model evaluation.
9-2310	9	168	1	171	2	The maps in Figs. 9.2-9.5 use a projection centered on the Pacific in which land areas for example in Europe are most distorted, unlike other world maps in this and those in Chapter 12 that center 0° longitude. Consistency would be nice with these other maps. [Stefan Fronzek, Finland]	Taken into account. Figure has been completely revised. We have used the Robinson Projection centered at 10°E for global maps.
9-2311	9	168	1			The values in the top-left figure are in kelvin, but incorrectly listed as degree C below the map and in the caption. [Stefan Fronzek, Finland]	Taken into account. Figure has been completely revised. The unit has been fixed, and we have used °C unit throughout the chapter.
9-2312	9	168	8	168	9	Fig. 9.2 Bottom right panel, the mean reanalysis inconsistency is shown to indicate it is mostly smaller than the multi model mean absolute error in bottom left panel. However, if purely basing on the color scale, it is not obvious for regions that have large model biases. If same color scale is used for the bottom left and right, it will be clearer. [Government of United States of America]	Taken into account. Figure has been completely revised. Same color scale has been used.
9-2313	9	168		170		I would suggest to reorganize and homogenize a little bit the panels in fig 9.2 to 9.4 if possible. If possible, I would appreciate to see here a comparison with the CMIP3 mean bias. The observations could be put in the upper-right panel for fig 9.3 and 9.4 and the CMIP5 multi model bias systematically on the lower left panel. If the CMIP3 ensemble mean bias could be added on the figure, I think that the panels with absolute errors could be removed : ie : the lower left panel of fig 9.2, the the upper right panel of fig 9.3 and the lower left panel of fig 9.4. [Frédéric HOURDIN, France]	Taken into account. There is little qualitative difference between the CMIP3 and CMIP5 MMM. We now compare results of individual CMIP3 and CMIP5 models in Figure 9.6.
9-2314	9	168				Figure 9.2: The projection does not allow to see Africa and part of the Atlantic; Top left: the bar is in K, but C	Taken into account. Figure has been completely

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						appears; bottom left: red colors are saturated; bottom right: Caption: which three reanalysis? [Massimo Bollasina, Italy]	revised. The units have been fixed to °C, more distinguishable color have been used. The three reanalysis data (ERA-Interim, ERA40, JRA25) have been specified in figure caption.
9-2315	9	168				- Figure 9-2: use ERAinterim data for 1980-2005 to compare the historical run for 1980-2005 (do not compare a 15-year with 25-year average!) - unit in upper left panel must be Kelvin [Barbara Früh, Germany]	Taken into account. Figure has been completely revised. 1980-2005 ERA-Interim data has been used to be consistent with model simulations. The unit has been fixed.
9-2316	9	168				Fig. 9.2: It would help to have a consistent colour scheme between the lower two panels. [Anthony Hirst, Australia]	Taken into account. Same color scales have been used.
9-2317	9	168				Fig. 9.2 and all other maps. Suggest removing grid lines, contour labels in figures, using Arial/Helvetica fonts throughout the figure, and smaller and not bold labels. [Reto Knutti, Switzerland]	Taken into account. Figure has been completely revised. Grid lines have been removed, and we have used sans-serif font throughout all the figures.
9-2318	9	168				Figure 9.2: why is a ERA-interim used as a reference here. ERA interim is known to have a temperature bias. For sake of completeness at least the difference between ERA-interim and CRU or GISSTEMP should be shown. [Alexander Loew, Germany]	Taken into account. Surface temperature data is assimilated into 3 different
9-2319	9	168				Fig 9.2: The colorbar ticks indicate units in °C but the values are not consistent. Are the values shown in K or is there a mistake in the colorbar only? [RYM MSADEK, United States of America]	Taken into account. We have fixed this issue and used °C throughout the chapter.
9-2320	9	168				Figure 9.2: Bottom panels, suggest re-scaling (perhaps 0-10 degrees and a separate colour for 10+) so that these can be easily compared. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Figure has been completely revised. The color scales have been changed be more comparable.
9-2321	9	168				It would be more accessible if the bottom 2 panels used the same color scale and gradations. Visually because the schema are so similar I am drawn into making side-by-side comparisons and it seems like this is the intrention but the use of slightly different schema and different scales is an impediment that should be rectified. Alternatively if you do not wish such comparisons then different schema should be used. Also, the top left hand panel is (I hope!) in K and not C! [Peter Thorne, United States of America]	Taken into account. Figure has been completely revised.
9-2322	9	169				Figure 9.3: How significant are these differences? How do they compare to the uncertainties in the observations used (Table 9.2). [Alexander Loew, Germany]	Noted. These plots are entry point into the chapter and as in previous assessments are meant to be a simple as possible. Significance testing appears later in the chapter.
9-2323	9	169				Figure 9.3: Top left panel, please re-scale so that white is 0. All panels, please clarify whether this is the mean of the seasonality of the models, or the seasonality of the mean model. Bottom panels, please rescale so that the colours between -2 and +1 are distinguishable. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Figure has been completely revised. This is the mean of the seasonality of the models.
9-2324	9	169				Figure 9.3, top left panel: Please center this colorbar on zero. [Andrew Wittenberg, United States of America]	Taken into account. Figure has been completely revised using different color centering on zero.
9-2325	9	170	1	170	4	The observation map should be plotted here alongside the multimodel mean precipitation. It is the only sensible way to make a fair comparison. [FRANCISCO J. TAPIADOR, SPAIN]	Taken into account. Figure has been completely revised.
9-2326	9	170	1	170	4	Regarding representation, I believe that contour plots hide the differences. Raster plots are probably more suitable for a fair comparison. Besides, raster plots provide information on the spatial resolution of the data. Therefore, I suggest using raster plots for the representation of the maps in this figure. [FRANCISCO J. TAPIADOR, SPAIN]	Taken into account. Figure has been completely revised.
9-2327	9	170	1	170	4	The intervals in the color bar of the "Multi Model Mean Relative Error" might be considered as misleading by some as they are too wide and the white color gives the impression of a small error. I suggest either using a 10% interval or using a different color for the 0-20% relative error. [FRANCISCO J. TAPIADOR, SPAIN]	Taken into account. Figure has been completely revised using different color and 15% interval.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2328	9	170	1	170	4	The multi model mean absolute error map is not as useful as an error percentage map is. I suggest plotting the differences between models and observations as relative rather than absolute differences. In any case, a choice should be made between plotting the model vs. observations and the model vs. multimodel mean in the same way, either as absolute or as relative differences. [FRANCISCO J. TAPIADOR, SPAIN]	Noted. Figure has been completely revised.
9-2329	9	170	1	170	4	Plotting alongside the multimodel mean absolute error and the multimodel mean relative error is extremely confusing as the reference precipitation is different in both cases. In the first case, the reference is observed precipitation, whereas in the second map the reference is the multimodel mean precipitation. [FRANCISCO J. TAPIADOR, SPAIN]	Noted. Figure has been completely revised.
9-2330	9	170	1	170	4	I strongly suggest to rework figure 9.4 to show: (1) multimodel mean precipitation; (2) observed mean precipitation; (3) multimodel vs. observed mean precipitation (in %); (4) multimodel mean absolute error (compared with observations); (4) multimodel mean absolute error (compared with multimodel mean). The reason for this is to highlight the differences and to provide a fairer appraisal of the model ability to model precipitation. After doing so, a note explaining that precipitation observations over ocean are not as reliable as over land would be necessary. [FRANCISCO J. TAPIADOR, SPAIN]	Noted. Figure has been completely revised. This is difficult to address as we have received many conflicting suggestions through the first and second order draft.
9-2331	9	170				Figure 9.4: numbers appear on the panels; bottom left: colors are saturated; top right: white color is not convenient for the 1-2 range; caption: specify the observations [Massimo Bollasina, Italy]	Taken into account. Figure has been completely revised and the observation data GPCP has been specified in figure caption.
9-2332	9	170				Fig. 9.4 We recommend that the figure and caption should be changed to reflect the fact that precipitation "rates" are shown. [Government of United States of America]	Noted. Figure has been completely revised.
9-2333	9	170				Fig. 9.4: It would aid display to shift the colour shading scheme such that the zero line was between the light red red shading and the white shading, rather than the medium red shading and the light red shading. [Anthony Hirst, Australia]	Noted. This was considered but not implemented.
9-2334	9	170				Figure 9.4 The upper right color bar for multi model mean bias of precipitation might be modified so that the "brown" color and "blue" color are separated around zero rather than the current around 1. [Weiping Li, China]	Taken into account. Figure has been completely revised with color scale centering on zero.
9-2335	9	170				Figure 9.4: How significant are these differences? How do they compare to the uncertainties in the observations used (Table 9.2). [Alexander Loew, Germany]	Noted. This was considered but not implemented.
9-2336	9	170				Fig 9.4. This is a very important figure. I call attn to the bias in multi-model mean. An error of 1 mm day-1 corresponds to 36 cm yr-1, which might be related to ann avg precip in agricultural regions of say 100 cm yr-1. Ask any farmer whether a change in precip of 36 cm yr would be significant to his crops and he will tell you in no uncertain terms. And this is the multi-model mean. It is certainly worse in individual models. It is essential to show this for each of the models as was done in Suppl to chapter 8 in AR4. My guess is that errors of up to 50 cm yr-1 will not be uncommon. In this case one must ask whether the models are ready to incorporate agriculture and terrestrial vegetation. My view would be an emphatic no, but I think the report must do an assessment of the consequences of such errors on vegetation models and justify its conclusions. Text p 9.21 simply says simulation of precipitation is a tough test for models but does not speak to the implications of the errors. [Stephen E Schwartz, United States of America]	Rejected. In the model evaluation our focus is not on impacts.
9-2337	9	170				Figure 9.4: Top right and bottom right panel, please rescale so that white is 0. Bottom left panel, this scale is too long (difference between 4 and 6 not visible), suggest adding colours. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Figure has been completely revised.
9-2338	9	171		171		Figure 9.5: Units are missing in plots and caption, (W/m2) [European Union]	Taken into account- We have added the units.
9-2339	9	171				Figure 9.5: left panels: unit is missing; right panel: replace the grey lines with the model spread [Massimo Bollasina, Italy]	Taken into account- units has been added.
9-2340	9	171				Figure 9-5: in the caption it is referred to CMIP3 whereas in the text describing the figure (p.9-23 l.11 & 13) it is referred to CMIP5 [Barbara Früh, Germany]	Taken into account-The figure has been updated with CMIP5 data.
9-2341	9	171				Fig. 5: Will this figure be updated with CMIP5 data? If possible, we recommend that this figure be updated with	Taken into account- Combined with comment 9-2340.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						these data. [Government of United States of America]	The figure has been updated with CMIP5 data.
9-2342	9	171				This figure should use the same projection and style as Figure 7.6 (or vice versa) [Steven Sherwood, Australia]	Accepted - We have used the Robinson Project centered at 10°E for global maps.
9-2343	9	171				Figure 9.5: Left panels, which obs dataset is used? Also which models? Text (9-23, lines 11 and 13) says CMIP5, fig caption says CMIP3. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account - Observations data has been specified. The figure has been updated with CMIP5 data.
9-2344	9	172				Figure 9-6: explain abbreviation "20C3M" [Barbara Früh, Germany]	Taken into account. Figure has been completely revised.
9-2345	9	172				Figure 9.6 is not easy to understand. It is difficult to distinguish between CMIP3 and CMIP5 on this figure. Please consider editing the figure for clarity. [Government of United States of America]	Taken into account. Figure has been completely revised.
9-2346	9	173	1	173	1	GISS-E2-R_CC is exactly the same in these metrics than GISS-E2-R (physics-version=1). However, p2 and p3 are distinct. [Gavin Schmidt, United States of America]	Noted. Figure has been completely revised.
9-2347	9	173	1	173	1	Relative error is simply a model beauty context. Absolute error relative to some objective criteria is much more useful at judging whether models are going to be useful or whether there is a practically important range of skill. [Gavin Schmidt, United States of America]	Rejected. Relative error does provide quantitative information about how well models agree with observations, and in particular how this agreement varies across models.
9-2348	9	173				Figure 9-7: which data set in table 9-2 is used as default or alternate data sets? [Barbara Früh, Germany]	Taken into account. Now indicated in Table 9.3
9-2349	9	173				Figure 9.7: Perhaps I have misunderstood this figure, but I cannot see the multi-model mean, as referred to in line 4 of page 25, in this Figure. [Gill Martin, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Now in the figure.
9-2350	9	173				Figure 9.8: Useful to have the absolute values at top. Could the same be included for the EMICs? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account - We have included absolute values for EMICs.
9-2351	9	174	1	174	1	What is the meaning of the magnified time (reference period) and what is x-axis here? Not clear to me [European Union]	Taken into account - The top inset shows 1961-1990 mean temperature, the x-axis represent the temperature value. To make it clearer, we have moved the top inset to the right of the time series panel, and have added y-axis string to indicate it.
9-2352	9	174	1	174	1	What does GISS-E2-# refer to? GISS-E2-H? [Gavin Schmidt, United States of America]	Editorial - it should be GISS-E2-R, fixed.
9-2353	9	174				Figure 9.8: is the drift removed from the time series? Colors are poorly visible [Massimo Bollasina, Italy]	Taken into account - the drift is not removed. We have redrawn the figure and adjusted the colors.
9-2354	9	174				Fig. 9.8: This is a valuable figure. However, as implied in the caption, it is not quite comparing like (models) with like (observations). The model results are for surface air temperature globally, whereas the observations are for merged sea surface temperature over the ocean and surface air temperature over land. This is common practice, but it seems not quite best practice, which would seem to be to do for the models exactly as for the observations (i.e., SST over ocean and surface air temperature over land). [Anthony Hirst, Australia]	Taken into account. We have revised the figure and use ERA-Interim data for surface air temperature which is T(2m) over land and ocean, consistent with the model data rather than the merged in-situ products which are SST and T(2m) over land. Note the difference between product is small compared to model errors.
9-2355	9	174				Figure 9.8: What is at the top of figure A added? Is this really helpful. Skip it or describe it properly in the figure caption. [Farahnaz Khosrawi, Sweden]	Taken into account - The top inset shows 1961-1990 mean temperature, we have adjusted the panels and made it clearer, refer to the responses to comments 9-2351.
9-2356	9	174				It would be useful to know the magnitude of aerosol forcing (direct and indirect) that results from the prescribed (or modeled) aerosol conc. To what degree have the models been tuned (using aerosols) to agree with the 20th century warming? If the CMIP5 models had been run with aerosol input giving as a low aerosol forcing as given in this draft I doubt that the agreement between models and observations would have been as	Taken into account -- evidence that the CMIP5 models have not been explicitly tuned via their aerosol loading is discussed in section 9.4.6.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						good as shown in Figure 9.8, [Henning Rodhe, Sweden]	
9-2357	9	175	1	175	1	Figure 9.8 caption: "dotted line"? Isn't that shown by a solid line? Add that the black solid and dotted lines are the observations. [Farahnaz Khosrawi, Sweden]	Taken into account - The figure has been redrawn and the caption has been revised.
9-2358	9	176				Fig. 9.9 uncertainty bars on the observation values are needed in this figure to understand whether the model relationship agrees or not within observational uncertainty. The ERA Interim and UAH values are very different from models or each other. One problem may be the short length of the data record given natural variability: if this is an issue, the uncertainty in the regressions will give some measure of it. Or are the observations just not sufficiently accurate over decade time scales? [Government of United States of America]	Rejected. This plot is based on a peer reviewed publication.
9-2359	9	178	1			SST (anomaly) is not "surface air temperature" or even close to it, in regions where sea-ice is present (eg especially N Atlantic). This must surely be well-known, but if you would like a reference, Fig 2 of Hargreaves et al Clim. Past, 7, 917–933, 2011 shows it clearly for the PMIP2 models. In panel b, is there a reason for only using such a small region of the high latitudes? When talking about polar amplification, maybe the whole high latitude region would be more appropriate. [James Annan, Japan]	taken into account. More precise language is now used.
9-2360	9	178		178		Figure 9.11b: The axis labels and the legend cannot be read, the font is too small [European Union]	Accepted
9-2361	9	178				Figure 9.11, panel b: Increase the font size. The legend and axis description can hardly be read in the present form. The font size in panel c could also be increased. [Farahnaz Khosrawi, Sweden]	Accepted
9-2362	9	179	13	179	13	Figure 9.12 is very interesting and the first time there is a synthetic metric used for paleo. I have however some difficulties to understand the colour legend. If as in Fig 9.7, the reference is taken with the mean of the models but should not then the mean appear in white ? you may also consider if possible to reduce a bit the number of metrics or detail more the legend. [SYLVIE JOUSSAUME, France]	the figure has been simplified to improve the readability and the normalisation better explained in the figure caption
9-2363	9	179				Figure 9.12: Increase font size. [Farahnaz Khosrawi, Sweden]	Accepted
9-2364	9	179				Fig 9-12 caption: I suggest switching LGM and Mid-Holocene in the first sentence of the caption to be consistent with the order in which they appear in the figure. [RYM MSADEK, United States of America]	Accepted
9-2365	9	179				Figure 9.12: Second panel, the median bias and Euclidean distance are exactly the same? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	N they are not, because the normalisation is done on the median of the PMIP3/CMIP5 simulations only. The figure has been revised as well as the figure caption
9-2366	9	180	7	180	8	If observations are available from 1874 to 2008, why not comparing models and observations during the same period (1975-2005)? If there is a trend in SST and SSS since 1874, this could change the model biases against these observations, couldn't it? [Marie-Estelle Demory, United Kingdom of Great Britain & Northern Ireland]	Taken into account. The observations most closely resemble a mean for the period 1980-2010, while warming has occurred during this period, the 30-year period selected for the models (1975-2005) most closely represents the period coincident to the available observations. We have updated the caption to more accurately convey this information. See also Chapter 3
9-2367	9	181				Figure 9.14: which HADISST version is used here? The most recent one? Specify in Table 9.2 as I suggested. [Alexander Loew, Germany]	Taken into account. Version clarified
9-2368	9	181				Fig 9-14: To be consistent with previous figures, I would suggest using red for CMIP5 mean instead of black, as black was used mainly for observations in other figures. [RYM MSADEK, United States of America]	Taken into account. Figure has been completely revised.
9-2369	9	181				Fig 9.14. This is a valuable figure. It should be plotted vs sin(lat) to avoid distortion of perception (equal area representation). Suggest means for each model be shown at right of graph. It is difficult to distinguish the different models by the colors; perhaps a line code device can be used as well. [Stephen E Schwartz, United States of America]	Noted. Latitude is used for internal consistency with the other report figures. Figure has been completely revised
9-2370	9	182				Fig. 9.15: Given the length of the chapter I'm not sure this figure is essential. Part of the ocean discussion is interesting but I can't see much in terms of conclusions that would carry over to the projections chapter. [Reto Knutti, Switzerland]	Taken into account. Figure has been replaced.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2371	9	182				Fig 9-15 caption: Although it is specified in the figure description p.33 it could be added in the caption here that the models shown are CMIP3 models. [RYM MSADEK, United States of America]	Taken into account. Figure has been replaced.
9-2372	9	183	5	183	5	Figure 9.16 caption: Observed spatial pattern. What is meant here with spatial pattern? [Farahnaz Khosrawi, Sweden]	Taken into account. Figure has been completely revised.
9-2373	9	183	8	183	8	Figure 9.16 caption: Why global ocean if it is only between 50 N and 50S? [Farahnaz Khosrawi, Sweden]	Taken into account. Figure has been completely revised.
9-2374	9	183				Figure 9.16: add at the x-axes "observation" and at the y-axes "modelled" in parantheses so that it is clearer which standard deviation it is. [Farahnaz Khosrawi, Sweden]	Taken into account. Figure has been completely revised.
9-2375	9	183				Figure 9.16: Is this figure really helpful? The figure is somewhat difficult to understand. [Farahnaz Khosrawi, Sweden]	Taken into account. Figure has been completely revised.
9-2376	9	184				Figure 9.17: Include some space between figure frame and legend frame. Position the legend either at the right top or right bottom of the figure. [Farahnaz Khosrawi, Sweden]	Taken into account. Figure has been completely revised.
9-2377	9	184				Figure 9.17,Difficult to identify the different model tiem seris by similar kind of colours. It will be nice if some legend is also added in the line, or some other means to make it more distinguishable. [HASIBUR RAHAMAN, India]	Taken into account. Figure has been completely revised.
9-2378	9	184				Suggest auxilliary vertical axis in units W yr m-2 and lines of constant slope corresponding to different heating rates in W m-2, to guide the eye. [Stephen E Schwartz, United States of America]	Taken into account. Figure has been completely revised.
9-2379	9	184				Sorry to be a pain but I find Fig. 9.8 quite striking: the growing departure between models and obs global mean temperature during the last decade or so is nearly unprecedented and makes the models look "alarmist" (to take the contrarian's phrase), or else indicates something has gone wrong with forcings toward the end of the dataset. But then I see Fig. 9.17 which seems to contradict that (maybe there is a time lag issue). These points are important enough to highlight in the summary, which is currently rather bland and avoids the tough questions that many are asking about our models. If they are ignored, critics will have some justification if (when) they call the chapter a whitewash. Better to confront these things and explain why they are very unlikely to mean we have everything all wrong. [Steven Sherwood, Australia]	Taken into account. Figure has been completely revised. See the Hiatus Box.
9-2380	9	184				Figure 9.17: This is an anomaly plot like 9.8. As in 9.8, it would be useful to include the absolute values for reference. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Rejected. Energy content by itself is not a very informative quantity; it is the change that matters.
9-2381	9	185	11	185	14	Figure 9.17 caption: The sentence should be moved to the main text. [Farahnaz Khosrawi, Sweden]	Taken into account. It is Figure 9.18 caption. Note, however, that unlike in a paper, figure and caption in an IPCC report should be more self-contained, implying more interpretation in a caption than one would find in a paper.
9-2382	9	186	7	186	9	Figure 9.18 caption: move this sentence to the main text. [Farahnaz Khosrawi, Sweden]	Accepted
9-2383	9	186				Figure 9.18: Use only "CMIP3" and "CMIP5" in the plot title. Add a legend and put the rest of the information there. [Farahnaz Khosrawi, Sweden]	Rejected. It is important to compare with previous simulations (PMIP2). However the layout of the figure has been revised, so that only the results of the PMIP3 simulations areplotted in colours.
9-2384	9	186				Figure 9.19. In figure caption please change "NCPE1" to "NCEP1" [HASIBUR RAHAMAN, India]	Editorial. Corrected
9-2385	9	187	1	187	2	Lower frame title wrong -- should be CMIP5. Also identification wrong, should be (b). [Government of Australia]	Editorial. Corrected
9-2386	9	187	8	187	10	Figure 9.20 caption: This sentence should be included in the main text. [Farahnaz Khosrawi, Sweden]	Taken into account
9-2387	9	187		187		Figure 9.20: Title of the lower panel should be: "(b) CMIP5..." [European Union]	Editorial. Corrected
9-2388	9	187				Fig. 9.20 panel b is mislabeled. Should be CMIP5? [Government of United States of America]	Editorial. Corrected

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
9-2389	9	187				Figure 9.20: The figure title in the bottom figure reads (a) instead of (b) and there is also written "CMIP3" instead of "CMIP5". Leave only "CMIP3" and "CMIP5" in the figure title and put all other information into a legend. [Farahnaz Khosrawi, Sweden]	Editorial. Corrected
9-2390	9	187				Fig 9-20:The lower figure should have panel b as title and I presume it shows CMIP5 results while the upper panel shows CMIP3 results unlike what the current title says. [RYM MSADEK, United States of America]	Editorial. Corrected
9-2391	9	187				Figure 9.20, both the figure caption are same. The lower panel should be (b) CMIP5 [HASIBUR RAHAMAN, India]	Editorial. Corrected
9-2392	9	187				Figure 9.20, bottom panel: Panel title is incorrect, should be "(b) CMIP5...". And both panels say "NCPE1" instead of "NCEP1". [Andrew Wittenberg, United States of America]	Editorial. Corrected
9-2393	9	188	20	188	20	Figure 9.21 (caption): What magnitude (percentage) is the correction? 0.1%? 10%? [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account. Caption clarified
9-2394	9	188				Figure 9.21 caption: The text is too long and parts of it should be included in the main text. [Farahnaz Khosrawi, Sweden]	Taken into account
9-2395	9	189	1	191	12	That would be good if (1) The reference periods for average sea ice extent and concentration in Fig. 9.22 and 9.23 were the same, i.e. if the 1980-1999 period was used for Figure 9.23.A. [François Massonnet, Belgium]	Rejected. Wherever possible the new AR5 baseline (1986-2005) is used. Unfortunately it is technically difficult when CMIP3 is compared with CMIP5 - then the AR4 baseline (1980--1999) is used..
9-2396	9	189	5	189	5	Figure 9.22 caption: skip the text in parantheses. [Farahnaz Khosrawi, Sweden]	Rejected. This is important information.
9-2397	9	189	5	189	9	The satellite data shown in this figure should be consistent with those used in Chapters 4, 11 and 12. Since the model distribution is not Gaussian around the mean, consider showing the 5-95% range of the ensembles (as in Chapter 12) instead of the +/- 1 std. [Thierry Fichet, Belgium]	Taken into account. Figure modified.STD is still used here following the Pavlova et al., 2011. However 5-95% range is used further at fig.9.24
9-2398	9	189				Figure 9.22: Increase font size. [Farahnaz Khosrawi, Sweden]	Taken into account. Figure modified.
9-2399	9	190	10	190	11	The data plotted in this figure should be consistent with those used in Chapters 4, 11 and 12. [Thierry Fichet, Belgium]	Noted. But the comment is not concrete.
9-2400	9	190				Fig. 9.23. The figure is hard to read. How about a redesign: panel a: CMIP5 minus obs; b) CMIP5 (March) and Obs (March) 15% lines; c) CMIP5 (Sept) and Obs (Sept) 15% lines (in different colors). [Government of United States of America]	Rejected. No differences with observations are supposed to be shown at 9.23
9-2401	9	191	1	191	12	Fig. 9.24. As far as I understand, there is redundancy between the two insets and the figure they are respectively nested in. Is it necessary to show explicitly all models names and extents, since this is usually not done in the chapter for other variables? I would only consider the insets, that I think contain enough information and are much less noisy. [François Massonnet, Belgium]	Rejected. Insets are not redundant. Showing individual models is found helpful.
9-2402	9	191	1	191	12	Fig. 9.24. Why are 34 CMIP5 models shown here, and 40 in Fig. 9.23? Wouldn't it be better, for consistency, to show the same number of models? [François Massonnet, Belgium]	Noted. The intention is to show as many models as possible. However for 9.24 the ""historical" runs are extended by using "scenario" runs which at some point of time have been not available for all 40 models.
9-2403	9	191	5	191	11	The satellite data shown in this figure should be consistent with those used in Chapters 4, 11 and 12. Since the model distribution is not Gaussian around the mean, consider showing the 5-95% range of the ensembles (as in Chapter 12) instead of the +/- 1 std. Also, consider including the 2012 data in the figure. On the other hand, the caption should mention that only one ensemble member per model is displayed. Finally, is there a particular reason for not showing the times series of the winter sea ice areal coverages? [Thierry Fichet, Belgium]	Taken into account. Figure modified.Not showing winter time series is due to space limits.
9-2404	9	191				Fig. 9:24: Here and in many other spaghetti figures, I don't see the value of using individual colors. It is impossible to see which line corresponds to which label. The figures would be easier with all individual	Rejected. Of a number of versions of this figure the current version has been found optimal.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						simulations having the same color. I also think the panels embedded in the bigger panels are not attractive. If they are needed, they should be separate. Further, grid lines should be avoided. [Reto Knutti, Switzerland]	
9-2405	9	192				Figure 9.25: Add a figure title reading "February 1986-2005". [Farahnaz Khosrawi, Sweden]	Accepted. Caption modified.
9-2406	9	192				Figure 9.25: What is denoted by the grey areas? [Farahnaz Khosrawi, Sweden]	Noted. All 24 models considered have snow (see the colour bar)
9-2407	9	192				Figure 9.25: What is the unit of the color bar and color coding showing? [Farahnaz Khosrawi, Sweden]	Taken into account. Figure modified. The colour bar shows the number of models with snow in a particular grid cell.
9-2408	9	193	5	193	9	The representation of uncertainty is not clear in this diagram and caption. The figure appears that the confidence level increases as the range of value decrease, which is not making sense to me. [Takamitsu Ito, United States of America]	Accepted and taken into account - revised figure has a corrected (i.e. reversed) uncertainty scale.
9-2409	9	193	5	193	9	Figure 9.26: Add to the caption text that the vertical lines are denoting volcanic eruptions. [Farahnaz Khosrawi, Sweden]	For clarity we have decided to remove the lines representing volcanic eruptions.
9-2410	9	193				Fig. 9.26 We don't understand the confidence limits here. (Scale/legend, etc) [Government of United States of America]	See response to 9-2408.
9-2411	9	193				Figure 9.26: Why is one of the volcanic eruptions given by a red line? What is emphasized here? [Farahnaz Khosrawi, Sweden]	See response to 9-2409.
9-2412	9	195	8	195	11	Figure 9.28 caption: Aren't these all CMIP models. Is it really necessary to have this list in the caption? [Farahnaz Khosrawi, Sweden]	Rejected -- since not all CMIP5 models are included, and since the ones that are included are combined into a single error metric and hence cannot be readily distinguished or identified, the names of the models included in the figure have been retained in the caption.
9-2413	9	195				Figure 9.28: add levels between 0-1 [Massimo Bollasina, Italy]	Accepted -- levels have been added between 0 and 1.
9-2414	9	195				Figure 9.28: This figure uses an OLD version of MODIS data for a limited timeperiod. Why is no recent MODIS collection data used? Where are differences significant? Why is only limited timeperiod used? [Alexander Loew, Germany]	Accepted. Data over oceans has been updated to latest MODIS collection and data over land has been replaced with MISR retrievals
9-2415	9	195				The relevant quantity would be error in AOD, not relative error in AOD. A relative error of 100% is much more important for high AOD than low AOD because forcing scales with AOD, not relative error. Also at low AOD measurement error is itself very large, giving a distorted picture. It would be valuable to have this for each of the models (supplementary material), not just the median. [Stephen E Schwartz, United States of America]	Accepted -- the error is now shown in absolute units
9-2416	9	195				Quantity should be denoted AOD aerosol optical depth (vertical integral of extinction coeff), not AOT, aerosol optical thickness, which is integral along some path. Same for page 196. [Stephen E Schwartz, United States of America]	Accepted -- terminology has been switched to AOD.
9-2417	9	196	7	196	8	Figure 9.29: can't this be added to the legend? [Farahnaz Khosrawi, Sweden]	Accepted -- labeling of models has been harmonized with other figures.
9-2418	9	196	7	196	8	Figure 9.29: Are the ACCMIP models not included in CMIP? [Farahnaz Khosrawi, Sweden]	Rejected -- No, the ACCMIP models are not included in the CMIP5 archive. For example, the NCAR-CAM3.5 model in the ACCMIP ensemble that is plotted on this figure is not included in CMIP5.
9-2419	9	196				Figure 9.29: Why are the ACCMIP models additionally plotted? What does the reader gain here? [Farahnaz Khosrawi, Sweden]	Rejected -- the ACCMIP results show that the spread among the climate models extends to carefully controlled intercomparisons among chemical transport models and the parent chemical parameterizations

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							used in the CMIP5 models.
9-2420	9	196				<p>The spread of AOD among the several models should be viewed, in my opinion, not as a pat on the back for the modelers for somehow matching trends in observations, but rather as a cause for concern. For a sensitivity of TOA forcing to AOD of roughly 30 W m⁻² per tau (24 hr avg, equinox, low and mid latitudes) in cloud free conditions (Mccomiskey et al 08); derate by a factor of 2 for cloud cover to get 15 W m⁻² per tau. So the error in tau of ~0.1 indicated in the figure corresponds to 1.5 W m⁻². This is substantial in terms of forcing over the industrial period. Note this is a measure of uncertainty in direct forcing that greatly exceeds estimate of such uncertainty given in the forcing chapter.</p> <p>McComiskey, A., S. E. Schwartz, B. Schmid, H. Guan, E. R. Lewis, P. Ricchiazzi, and J. A. Ogren (2008), Direct aerosol forcing: Calculation from observables and sensitivities to inputs, J. Geophys. Res., 113, D09202, doi:10.1029/2007JD009170 [Stephen E Schwartz, United States of America]</p>	Noted, and accepted -- a finding has been added to ES regarding the spread in AOD.
9-2421	9	198	4	198	14	Figure 9.31 caption: Add (a), (b) and c in the figure caption text so that it becomes clearer which test part refers to which panel. [Farahnaz Khosrawi, Sweden]	Accepted
9-2422	9	198				Figure 9.31: What is exactly shown in panel a and b? What is shown by the solid lines and what by the dashed lines? [Farahnaz Khosrawi, Sweden]	references to pannel a, b and c has been added in the figure caption
9-2423	9	199				Fig. 9.32 the CMIP3 results in this figure need to be replaced with CMIP5 results. [Government of United States of America]	Done
9-2424	9	200	6	200	6	Figure 9.33 caption: Which line refers to which data set? Are the orange lines the CMIP5 models and the colored curves the observations? [Farahnaz Khosrawi, Sweden]	The legend has been revised to clarify this point
9-2425	9	200				On page 208, Fig. 9.38 is almost too small to be useful. Consideration of a different presentation would be advisable, perhaps it could be split across two pages and turned sideways with one page for each of temp and precipitation. (Note that correct "from page" number could not be input for this comment due to limit of comment form) [Government of Canada]	Take into account - This figure has been redrawn.
9-2426	9	200				On page 218, FAQ 9.1 Fig 1. Why is there no temperature data for CMP5? It is acknowledged that the figure will be redrawn when more complete results are available and perhaps this is the reason for the omission. [Government of Canada]	Taken into account. CMIP5 temperature data have been added.
9-2427	9	200				On page 216, Line 40. 'Southern Annual Mode' should be 'Southern Annualar Mode". [Government of Canada]	Editorial
9-2428	9	200				Fig 9.45 (page 215, but the spreadsheet won't accept that): I really like this figure as a summary of the relative performance of different components. Should NBP and fgco2 be added to the "trends" panel? Figure 9.26 shows trends of both of these and evaluates the models' performance against them [Chris Jones, United Kingdom of Great Britain & Northern Ireland]	Accepted - NBP-t and fgco2-t added.
9-2429	9	200				Figure 9.33 "BCC-csm2-1" in the legend of bottom panel should be "BCC-CSM1-1". It's better to unify the model name in the whole chapter even the whole report. [Weiping Li, China]	Taken into account - CMIP5 names are used as in other figures
9-2430	9	200				Figure 9.33: Top plot no observations? Middle plot ideally needs something like "consistency bars" as although the model mean looks to be in reasonably the right place, none of the features of the obs are represented. Bottom plot would be better with thinner lines so that the range of the reconstructions is still visible (current version overlays the model data onto the recons and the grey lines are not very visible). [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account - Figure has been revised.
9-2431	9	202		202		The maps given on Figure 9.34 mark the Sino-Indian border with "McMahon Line" and no intermittent lines in the South China Sea. It is recommended to delete these maps or replace them with a border free one. [Government of China]	Rejected. This is a figure from a published paper that cannot be changed. Incidentally, China is not on the map.
9-2432	9	202				Fig. 9.34 The different metrics use different AMOC baseline time series for construction which is confusing. [Government of United States of America]	Noted. This is a figure from a published paper that cannot be changed. We are not sure what is

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
							confusing here.
9-2433	9	205	2			Fig. 9.36: X labels need to be resized. [Indrani Roy, U.K.]	Editorial. Corrected
9-2434	9	205				Figure 9.36: What is shown on the x-axes and what on the y-axes? Improve the labelling on the x-axis. [Farahnaz Khosrawi, Sweden]	Editorial. Corrected
9-2435	9	205				Figure 9.36: Is the figure really useful? [Farahnaz Khosrawi, Sweden]	We decided it was and did not get scientific comments otherwise.
9-2436	9	206	4	207	10	Figure 9.37 caption: Sort the figure caption text. In its current form it is hard to follow which text refers to which figure panel. It seems there are too many details discussed. [Farahnaz Khosrawi, Sweden]	Taken into account - caption revised. However, please note that the caption should be self-contained and detailed description is inevitable to some extent.
9-2437	9	206	6	206	6	Figure 9.37: please add the words in the figure caption "...errors for an individual model relative to the median model with respect to the respective reference dataset." [Jana Sillmann, Canada]	Accepted (in paraphrased form).
9-2438	9	206	9	206	9	Figure 9.37 caption: Skip top, left right and bottom, this is confusing. Use a, b, c, d as it is given in the figure itself. [Farahnaz Khosrawi, Sweden]	Rejected - note that top, left, etc. refers to the positions in each box in the panel (a), instead of different panels.
9-2439	9	206	10	206	10	Figure 9.37 caption: Which indices? [Farahnaz Khosrawi, Sweden]	Indices are clearly indicated in the caption.
9-2440	9	206				Figure 9.37: It would be more concise if this figure would be splitted into two figures. [Farahnaz Khosrawi, Sweden]	Noted - though that would be true, we decide to keep them in one figure as they are coming from the same paper.
9-2441	9	206				Figure 9.37 More available model results should be included to increase the confidence of the relevant analyses. [Weiping Li, China]	The updated figure includes as many CMIP5 models as possible at the cutoff date.
9-2442	9	207	4	207	4	newer results as in Sillmann et al. 2012 are based on HadEX2 [Jana Sillmann, Canada]	Noted - the figure is updated.
9-2443	9	208	5	208	6	Figure 9.38: Write that each red line is for a specific model. Is the thick line the model mean? [Farahnaz Khosrawi, Sweden]	Noted - the information can be found in the figure caption.
9-2444	9	208				Figure 9.38: Most of the individual plots are unreadable, especially the upper panel. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Taken into account - the figure has been redrawn to make it more legible.
9-2445	9	213	1			Fig. 9.43 (A) : Unfilled circles are not labeled [Indrani Roy, U.K.]	Taken into account - the figure has been redrawn and all the symbols can be found in legend.
9-2446	9	213		213		Fig. 9.43 have meaningfully different estimates for GFDL models than Winton et al. 2012. [Timothy Merlis, United States of America]	Accepted -- values for GFDL models in Figure 9-43 have been checked and reconciled against the Soden et al (2013) paper. Note that the GFDL model is not part of the sub-ensemble evaluated by Vial et al (2013) now plotted in this figure and table 9-4.
9-2447	9	213				fig. 9.43 Table 9.4 lists the ECS of GFDL AM2.1 (via. ESM2G or ESM2M) as 2.4C, while the published value based on a slab model is 3.4C (Stouffer et al J Clim, 2006). This is a 30% difference! Is this a typo, a calculation error somewhere, or is the approximation method not so good. This affects Fig. 9.43b also. If 3.4 C were used for the GFDL models ESM2M and ESM2G the correlation in the figure would be somewhat worse. [Government of United States of America]	Taken into account --combined with comment 9-2446
9-2448	9	213				Fig. 9.43: Please add units on the axes of both panels. [Government of United States of America]	Accepted -- Units have been added to both axes.
9-2449	9	214	4	214	4	Figure 9.44: Increase font size. [Farahnaz Khosrawi, Sweden]	Accepted -- Font size in the figure has been increased.
9-2450	9	214		214		Figure 9.44 : Estimates from Vial et al. (Clim. Dyn., 2012) for CMIP5 models could also be included. Vial et al. (2012) also translate the spread of feedback parameters into a spread of climate sensitivity (in K). [Sandrine	Accepted -- Estimates of feedbacks from Vial et al (2012) have been added.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						BONY, France]	
9-2451	9	214		214		Fig. 9.44 has the unit "W" lowercase. [Timothy Merlis, United States of America]	Accepted -- Unit of "W" has been capitalized.
9-2452	9	214				FAQ 9.1: To answer question about "how do would we know climate models are getting better", we may need to mention that recent observations show climate models are improving, e.g. Benka (2012). Reference: Benka, S.G., "Cloud simulations improving in climate models", Physics Today, Volume 65, Issue 10, October 2012. (http://www.physicstoday.org/daily_edition/physics_update/cloud_simulations_improving_in_climate_models) [Government of United States of America]	Taken into account. FAQ figure does show observationally-based performance metric.
9-2453	9	214				(Left side) Why are the colors for a given feedback sometime different between the two symbols? Am I missing something here? If not, I suggest using the same symbol for all the CMIP5 points and another symbol for all the CMIP3 points, and then use one color per feedback. [Karen Shell, United States of America]	Taken into account - the figure has been completely revised using grey circle for CMIP3 models and different color circle for CMIP5 models.
9-2454	9	214				Also, is this really adapted from Soden and Held, 2006, rather than Soden et al., 2008? Note that the cloud feedback is treated differently in the two papers. From the text, I think this is probably Soden et al., 2008. [Karen Shell, United States of America]	Noted -- reference to Soden and Held (2006) was correct and has been left unchanged.
9-2455	9	215	1	215	1	Did any fields get worse? [Gavin Schmidt, United States of America]	Taken into account - the revised version has a color also for degradation from CMIP3
9-2456	9	215	6	215	6	Interesting figure. I would suggest to include "stratospheric variability" instead of QBO, to generalize the evaluation of the stratospheric component. Results from Charlton-Perez et al JGR (fig 1) 2012 (submitted) can be used to this purpose (QBO is included in Charlton-Perez et al intercomparison). [Elisa Manzini, Germany]	Rejected - QBO is more specific.
9-2457	9	215	13	215	50	Figure 9.45 caption: Give this information in a table instead of the figure caption. [Farahnaz Khosrawi, Sweden]	Rejected - the author team believes the figure is clearer if the list is in the caption.
9-2458	9	215				Chapter 9. Figure 9.45 (page 215). It is recommended to complete the list of findings presented in this excellent summary figure and include aspects that are lacking, like the assessment of sea level rise (see results in 9.4.2.2) [Government of Spain]	Accepted - a more complete list of findings is presented in the final figure.
9-2459	9	215				Figure 9.45 What's the criteria for assessing model quality? No details about this are provided in the text. If multi-model mean is assumed in model quality assessment, many "sister" models come from similar origins (refer to Table 9.1) can't be counted as an independent sample in CMIP5. There were much less models in CMIP3. [Weiping Li, China]	Noted - This section provides a synthesis of model performance. Details on the criteria of assessing model quality and the model assessment are given throughout the chapter, with a particular reference to the subsections given in the caption of Figure 9.45.
9-2460	9	215				Figure 9.45: I was expecting this to show robustness and agreement on the axes and would like to see that graph as well. "Model quality" does not seem to be well defined. There is nothing in the "very low confidence" category and this could be removed? It is not immediately obvious that the interpretation of each box should be "there is medium confidence that the model quality of the mean state of TrlnOcean is high" rather than "there is medium confidence in the mean state of the TrlnOcean based on a high quality model" so this may be worth clarifying. [Erica Thompson, United Kingdom of Great Britain & Northern Ireland]	Rejected - there was a more complex figure that included that included robustness and agreement in the FOD but it was simplified in response to a reviewer comment. The author team believes this figure is clear and would be overloaded if more details are added. A reference to the section that backs up the assessment for each entry shown is given in the figure caption.
9-2461	9	218	7			Golaz et al 2011 (doi:10.1175/2010JCLI3945.1) also constructed plausible alternate model configurations by modifying cloud parameters. They found the present-day climate to be very similar from one configuration to another. However, they found that total aerosol forcing (AFaci) to vary by as much as 1 W/m2. [Jean-Christophe Golaz, United States of America]	Noted. Relationship to FAQ figure not apparent.
9-2462	9	218	9			Jackson et al 2008 (doi:10.1175/2008JCLI2112.1) performed a systematic parameter estimation using atmospheric GCM. [Jean-Christophe Golaz, United States of America]	Noted. Relationship to FAQ figure not apparent.
9-2463	9	219				Figure Caption cites incorrect Booth 2012 paper. The Nature paper is correctly cited else where in this capter, however this figure caption should be citing: "High sensitivity of future global warming to land carbon cycle processes, Ben B B Booth, Chris D Jones, Mat Collins, Ian J Totterdell, Peter M Cox, Stephen Sitch, Chris Huntingford, Richard A Betts, Glen R Harris and Jon Lloyd, Environmental Research Letters, 7, 2, 24002,	Taken into account. Reference corrected.

Expert and Government Review Comments on the IPCC WGI AR5 Second Order Draft – Chapter 9

Comment No	Chapter	From Page	From Line	To Page	To Line	Comment	Response
						2012 " [ben booth, United Kingdom]	