Chapter 9: Evaluation of Climate Models

Coordinating Lead Authors: Gregory Flato (Canada), Jochem Marotzke (Germany)

Lead Authors: Babatunde Abiodun (South Africa), Pascale Braconnot (France), Sin Chan Chou (Brazil), William Collins (USA), Peter Cox (UK), Fatima Driouech (Morocco), Seita Emori (Japan), Veronika Eyring (Germany), Chris Forest (USA), Peter Gleckler (USA), Eric Guilyardi (France), Christian Jakob (Australia), Vladimir Kattsov (Russia), Chris Reason (South Africa), Markku Rummukainen (Sweden)

Contributing Authors: Krishna AchutaRao, Alessandro Anav, Timothy Andrews, Johanna Baehr, Alejandro Bodas-Salcedo, Jennifer Catto, Ping Chang, Aiguo Dai, Clara Deser, Paul Durack, Michael Eby, Ramon de Elia, Thierry Fichefet, Bode Gbobaniyi, Clare Goodess, Fidel Gonzales-Rouco, Stephen Griffies, Alex Hall, Elizabeth Hunke, Tatiana Ilyina, Masa Kageyama, Viatcheslav Kharin, Stephen A. Klein, Jeff Knight, Reto Knutti, Felix Landerer, Tong Lee, Natalie Mahowald, Gerald Meehl, David Neelin, Tatiana Pavlova, Judith Perlwitz, Florian Rauser, Jouni Räisänen, Jeffrey Reid, Mark Rodwell, Adam A. Scaife, John Scinocca, David Sexton , Rym M'Sadek, Hideo Shiogama, Jana Sillmann, Ken Sperber, David Stephenson, Bjorn Stevens, Mark Webb, Keith Williams, Tim Woollings, Shang-Ping Xie, Jangling Zhang

Review Editors: Isaac Held (USA), Andy Pitman (Australia), Serge Planton (France), Zong-Ci Zhao (China)

Date of Draft: 5 October 2012

Notes: TSU Compiled Version

Table of Contents

Executive Summary

This chapter assesses 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). The diversity of models assessed has increased since the AR4 and now includes:

- coupled Atmosphere-Ocean General Circulation Models (AOGCMs), which are used in long-term climate projection and shorter-term (seasonal to decadal) climate prediction;

- the extension of AOGCMs with biogeochemical cycles into 'Earth System' Models (ESMs);

- Regional Climate Models (RCMs), used extensively in downscaling global climate results to particular regions; also statistical downscaling (SD) is considered;

- Earth System Models of Intermediate Complexity (EMICs) that are used to undertake very long (e.g., millennial) climate simulations, or to explore parameter uncertainty with very large ensembles.

The evaluation of climate models requires the availability of high-quality observational data, the uncertainty of which also needs to be understood and quantified. Observational data are described in Chapters 2–5. A particular advance in model evaluation since the AR4 has been in the area of model 'performance metrics' – that is, quantitative measures of model performance reflecting the difference between a model and a corresponding observational estimate that allows more quantitative evaluation and facilitates the synthesis of model performance. This chapter will make extensive use of such performance metrics, along with more traditional presentation of maps and time series.

The availability of coordinated multi-model experiments, notably the Coupled Model Intercomparison Projects (CMIP3 and CMIP5) allow for increasingly in-depth analysis of model results at different space and time scales using simulations designed specifically for this purpose. Results from CMIP5 are available in larger numbers; however comprehensive analysis is just getting started. The multi-model ensemble provides standardized output from different models, run with the similiar external forcing (greenhouse gases, aerosols, etc) over the historical period. This allows for direct comparison against historical observations, but also allows for some assessment of uncertainty in cases where suitable observations are not available (by examining inter-model spread). Investigating the connection between model errors/biases and characteristics or process parameterisations in a model is aided by extensive documentation of the models by the research community, something in progress and substantially improved since the AR4.

Simulations of Historical Climate

The ability of climate models to simulate historical climate, its variability, and its change, has improved in many, though not all, important respects relative to the previous generation of models featured in the AR4. Figure 9.45 provides an overview of models' capabilities as assessed in this chapter, including improvements from CMIP3 to CMIP5. Examples include:

- There continues to be very high confidence that models simulate realistically the surface temperature on continental and larger scales (Section 9.4.1.1, Figure 9.2). There is likewise very high confidence that models simulate realistically the global-scale surface temperature increase over the historical period, especially the last fifty years (Section 9.4.1.3.1, Figure 9.8). Together with the fact that climate models are based on fundamental physical and biogeochemical principles, these assessments lead to the very high confidence that models generally respond correctly to external forcing like changing greenhouse gases (Box 9.1, Chapter 10–12, 14).

- The simulation of large-scale patterns of precipitation has improved since the AR4, but there is robust evidence that models perform less well on precipitation than on surface temperature (Section 9.4.1.1, Figure 9.6). There is, however, medium confidence that models correctly simulate precipitation increases in wet areas and precipitation decreases in dry areas on large spatial scales in a warming climate, based on high agreement among models but only limited evidence that this is been detected in observed trends (Section 9.4.1.3.4) (Chapters 10–12 and 14).

- On regional scales (sub-continental and smaller), the confidence in model capability to simulate surface temperature and precipitation is less than for the larger scales. Nevertheless, there is high confidence that regional-scale surface temperature is now simulated with some improvement since AR4. Regional-scale precipitation continues, however, not to be simulated equally well, and the assessment remains difficult owing to observational uncertainties (Section 9.6.1.1).

- There is very high confidence that CMIP5 models realistically simulate the annual cycle of Arctic sea-ice extent, and there is high confidence that they realistically simulate the trend in Arctic sea-ice extent over the past decades. This is a clear improvement since AR4. Most models simulate a decreasing trend in Antarctic sea-ice extent, albeit with larger intermodal spread, as compared to a small increasing trend in observations (Section 9.4.3).

- There is high confidence that many models simulate realistically the observed trend in ocean heat content, although some models show substantial deviations from the observations (Section 9.4.2.2, Figure 9.17). This gives confidence in using climate models to assess the global energy budget (e.g., Box 13.2).

- There is robust evidence that the simulation of the El Niño-Southern Oscillation (ENSO) has improved from CMIP3 to CMIP5, with several models now simulating realistically the ENSO frequency spectrum and amplitude in sea surface temperature (Section 9.5.3.4.1, Figures 9.35 and 9.36). However, this improvement is not universal across the CMIP5 ensemble. Other important modes of variability, such as the North Atlantic Oscillation and the Quasi-Biennial Oscillation, are also simulated realistically, although the confidence in this assessment is only medium (Section 9.5.3). CMIP5 models are able to realistically simulate northern-hemisphere surface temperature variability on timescales of decades to centuries (Figure 9.33). The realistic simulation of patterns and levels of variability gives confidence to using climate models to separate signal and noise in detection and attribution studies (Chapter 10).

- There has been substantial progress since AR4 in the assessment of model simulations of extreme events (Section 9.5.4). There is medium evidence and high agreement that the global distribution of temperature extremes is represented well by models. The observed warming trend of temperature extremes in the second half of the 20th century is well captured, but there is medium evidence and medium agreement that models tend to overestimate the warming of warm temperature extremes and underestimate the warming of cold temperature extremes. There is medium evidence and medium agreement that CMIP5 models tend to simulate more intense and thus more realistic precipitation extremes than CMIP3. Some high-resolution atmospheric models have realistically simulated tracks and counts of tropical cyclones.

- There is high confidence that the trends in stratospheric ozone, whether prescribed or calculated interactively, is well represented in CMIP5 models, although deviations for individual models exist. This constitutes a significant improvement since the AR4, and there is high confidence that the representation of associated impacts on high latitude climate has improved compared to CMIP3.

- Most CMIP5 ESMs produce global land and ocean carbon sinks over the latter part of the 20th century that fall within the range of observational estimates (Section 9.4.5.3). The models also reproduce aspects of interannual variability and regional patterns of carbon uptake and release.

- 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). In some cases, model results are in general agreement with observations, but the observational uncertainty precludes definitive statements about model quality.

Climate Sensitivity in the CMIP5 Ensemble

- The CMIP5 model spread in equilibrium climate sensitivity ranges from 2.1 to 4.7°C and is very similar to the assessment in AR4. No correlation is found between biases in global-mean surface temperature and equilibrium climate sensitivity, which enhances confidence in model simulations of a warming climate even in the presence of errors in the time-mean (Section 9.7.2.1, Figure 9.43).

- There is very high confidence that the primary factor contributing to the spread in equilibrium climate sensitivity continues to be the cloud feedback. This applies to both the modern climate and the last glacial maximum. There is likewise very high confidence that there is a strong positive correlation between SST and water vapour on regional to global scales, implying realistic and positive water-vapour feedback in the models (Sections 9.4.1.3.1 and 9.7.2.2; Figures 9.9 and 9.44).

Relating Model Performance to Credibility of Model Applications

- There is emerging evidence that large inter-model differences in seasonal and interannual variability or past trends are well correlated with comparably large inter-model differences in the model projections for some particular quantities like Arctic summer-time sea-ice trends, snow albedo feedback, and interannual sensitivity of the CO_2 growth-rate to tropical temperature. These relationships provide a means to transform an observable quantity into a constraint for future projections (Section 9.8.3, Figure 9.46).

- While there has been substantial progress since the AR4 in the sophistication of assessing the reliability of the ensembles and evaluating whole ensembles of climate models against observations, there is to date no robust strategy for how to weight different models based on their performance within an ensemble (Section 9.8.3).

Regional Downscaling

- There is high confidence that regional downscaling methods do offer a means of providing credible, physically-based climate information at the smaller scales needed for many regional climate and impact studies. Although biases in global models propagate to RCMs, and the latter share many of the same resolution and process parameterisation issues with AOGCMs, the overall finding is that downscaling adds value both in regions with highly variable topography and for various small-scale phenomena. In addition, progress is being made on understanding and correcting regional biases. (Section 9.6).

Global climate models are able to reproduce many important aspects of the observed climate, its variability and change. There have been a number of improvements in model quality relative to the generation of models assessed in the AR4, although for many climate quantities, there has been no significant change. This indicates that the CMIP3 multi-model ensemble remains useful for many applications. Models still exhibit systematic errors and biases that are being actively researched, but overall, models provide a physically-sound basis for predictions and projections of future climate.

9.1 Climate Models and their Characteristics

9.1.1 Scope and Overview of this Chapter

Climate models constitute the primary tools available for investigating the response of the climate system to various forcings, for making climate predictions on seasonal to decadal time scales, and for making projections of future climate over the coming century and beyond. It is crucial therefore to critically evaluate the performance of these models, both individually and collectively. The focus of this chapter is particularly on the models whose results will be used in the detection and attribution chapter and the projections chapters 10 through 12, and so this is necessarily an incomplete evaluation. In particular, we will draw heavily on model results collected as part of the Coupled Model Intercomparison Projects (CMIP3 and CMIP5 (Meehl et al., 2007a); (Taylor et al., 2012) as this constitutes a set of well-controlled and increasingly well-documented climate model experiments. Other intercomparison efforts, such as those dealing with regional climate models (RCMs) and those dealing with Earth System Models of intermediate complexity (EMICs) are also used. It should be noted that the CMIP3 model archive has been extensively evaluated, and much of that evaluation has taken place subsequent to the AR4. By comparison, the CMIP5 models are only now being evaluated and so there is less published literature available. Where possible we show results from both CMIP3 and CMIP5 models so as to illustrate changes in model performance over time; however, where only CMIP3 results are available, they still constitute a useful evaluation of model performance in that for many quantities, the CMIP3 and CMIP5 model performance is broadly similar.

The direct approach to model evaluation is to compare observations with model output and analyze the resulting difference. This requires knowledge of the errors and uncertainties in the observations, which have been discussed in Chapters 2 through 6. Where possible, averages over the same time period in both models and observations will be compared, although for many quantities the observational record is rather short, or only observationally-based estimates of the climatological mean are available. In cases where observations are lacking, we will resort to intercomparison of model results to provide some quantification of model ‘uncertainty’.

After a more thorough discussion of the climate models and methods for evaluation in Sections 9.1 and 9.2, we describe the basic characterization of climate model experiments in Section 9.3, evaluate recent and longer-term records as simulated by climate models in Section 9.4, variability and extremes in Section 9.5, and regional-scale climate simulation including downscaling in Section 9.6. We conclude with a discussion of model performance and climate sensitivity in Section 9.7 and the relation between model performance and the credibility of future climate projections in Section 9.8.

9.1.2 Overview of Model Types to be Evaluated

The models used in climate research range from simple energy balance models to complex Earth System Models using state of the art high-performance computing. The choice of model depends directly on the scientific question being addressed (Collins et al., 2006c; Held, 2005). Applications include simulating palaeo or historical climate, predicting near-term climate variability and change on seasonal to decadal time scales, making projections of future climate change over the coming century or more, and downscaling such projections to provide more detail at the regional and local scale. Computational cost is a factor in all of these, and so simplified models (with reduced complexity or spatial resolution) can be used when larger ensembles or longer integrations are required. Examples of the latter include exploration of parameter sensitivity or simulations of climate change on the millennial or longer time scale. Here, we provide a brief overview of the climate models evaluated in this chapter.

9.1.2.1 Atmosphere-Ocean General Circulation Models (AOGCMs)

AOGCMs were the “standard” climate models assessed in the AR4. Their primary function is to understand the dynamics of the physical components of the climate system (atmosphere, ocean, land, and sea-ice), and for making projections based on future greenhouse gas and aerosol forcing. These models continue to be extensively used, and in particular are run (often at higher resolution) for seasonal to decadal climate prediction applications. In addition, high-resolution or variable-resolution AOGCMs are often used in process studies or applications with a focus on a particular region. Details of the AOGCMs assessed in this

Chapter can be found in Table 9.1. For some specific applications, the atmospheric component of such models is used on its own. This is referred to as an Atmospheric General Circulation Model (AGCM).

9.1.2.2 Earth System Models (ESMs)

ESMs are the current state-of-the-art climate models in the CMIP5, in terms of the extent to which the overall Earth system is represented. Compared to AOGCMs, ESMs include representation of various biogeochemical cycles such as those involved in the carbon cycle, the sulphur cycle, or stratospheric ozone (Flato, 2011). These models provide the most comprehensive tools available for simulating past and future response of the climate system to external forcing, in which biogeochemical feedbacks play a potentially important role. Details of the ESMs assessed in this Chapter can be found in Table 9.1.

[INSERT TABLE 9.1 HERE]

Table 9.1: Salient features of the AOGCMs and ESMs participating in CMIP5. Column 1: identification (Model ID) along with the calendar year ('vintage') of the first publication for each model; Column 2: sponsoring institution(s), main reference(s) and flux correction implementation (*not yet described*); Subsequent Columns for each of the 8 CMIP5 realms: component name, code independence and main component reference(s). Additionally, there are standard entries for the Atmosphere realm: horizontal grid resolution, number of vertical levels, grid top (low or high top); and for the Ocean realm: horizontal grid resolution, number of vertical levels, top level, vertical coordinate type, ocean free surface type ("Top BC"). This table information was automatically extracted from the CMIP5 online questionnaire (http://q.cmip5.ceda.ac.uk/) as of 12 November 2011.

9.1.2.3 Earth-System Models of Intermediate Complexity (EMICs)

EMICs attempt to include relevant components of the earth-system, but often in an idealized manner or at lower resolution than the models described above. These models are applied to certain scientific questions such as understanding climate feedbacks on millennial time scales or exploring sensitivities in which long model integrations or large ensembles are required (Claussen et al., 2002; Petoukhov et al., 2005). This class of models often includes Earth system components not yet included in all ESMs (e.g., ice sheets). As computing power increases, this model class has continued to advance in terms of resolution and complexity.

9.1.2.4 Regional Climate Models (RCMs)

Regional climate models (RCMs) are limited area models with representations of climate processes comparable to those in the atmospheric and land surface components of AOGCMs. Most RCMs are uncoupled and thus without interactive ocean and sea ice. The typical application of an RCM is to 'downscale' AOGCM simulations for some particular geographical region to provide more detailed information (Laprise, 2008; Rummukainen, 2010). Empirical and statistical downscaling methods constitute a range of techniques to provide similar regional or local detail.

9.1.3 Model Improvements

The climate models assessed in this report have seen a number of improvements since AR4. Model development is a complex and iterative task. Improved physical process descriptions are developed and included in the various model components,,entirely new model components, representing new insight into the physical, biological and chemical interactions of the climate system are introduced, and the spatial and temporal resolution of the models is improved. Finally, after the assembly of all model components into an AOGCM or ESM, model parameters are adjusted, or tuned, to provide a stable model climate. The overall approach to model development and tuning is summarized in Box 9.1.

[START BOX 9.1 HERE]

Box 9.1: Climate Model Development and Tuning

The AOGCMs, ESMs and RCMs evaluated here are based on fundamental laws of nature (e.g., energy, mass and momentum conservation). The *development* of climate models involves several principal steps:

i) Expressing the system's physical laws in mathematical terms. This requires theoretical and observational work in deriving and simplifying mathematical expressions that best describe the system;

ii) Implementing these mathematical expressions on a computer. This requires developing numerical methods that allow solution of the discretized mathematical expressions, usually implemented on some form of grid such as the latitude-longitude-height grid for atmospheric models.

iii) Building and implementing conceptual models (usually referred to as parameterisations) for those processes that cannot be represented explicitly, either because of their complexity (e.g., biogeochemical processes in vegetation) or because the spatial scales on which they occur are not resolved by the discretized model equations (e.g., cloud processes and turbulence). The development of parameterisations has become very complex (Jakob, 2010) and is often achieved by developing conceptual models of the process of interest in isolation using observations and comprehensive process-models. The complexity of each process representation is constrained by observations, computational resources and current knowledge (e.g., (Randall et al., 2007)).

The application of complex climate models requires significant supercomputing resources. Limitations in those resources lead to additional constraints. Even when using the most powerful computers today, compromises need to be made in three main areas:

i) Numerical implementations allow for a choice of grid spacing and time step, often referred to as "model resolution". Higher model resolution generally leads to mathematically more accurate models (although not necessarily more reliable simulations) but also to higher computational costs. Currently affordable climate model resolutions imply that the effects of certain processes must be represented through parameterisations (e.g., the carbon cycle or cloud and precipitation processes, see Chapters 6 and 7).

ii) The climate system contains many processes, the relative importance of which varies with the time-scale of interest. 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)

iii) Due to uncertainties in both the model formulation and the initial model state, a single model simulation only represents one of the possible pathways the climate system might follow. To allow some evaluation of these uncertainties it is necessary to carry out a number of simulations either with several models or by using an ensemble of simulations with a single model, either of which increases computational cost.

Trade-offs amongst the various considerations outlined above are guided by the intended model application and lead to the several classes of models introduced in Section 9.1.2.

Individual model components (e.g., the atmosphere, the ocean, etc.) are typically evaluated in isolation as part of the model development process. For instance, the atmospheric component can be evaluated by prescribing sea surface temperature (Gates et al., 1999) or the ocean and land components by prescribing the atmospheric input (Barnier et al., 2006); (Griffies et al., 2009). Subsequently, the various components are assembled into a comprehensive model, and a small subset of model parameters remain to be adjusted so that the model adheres to large-scale observational constraints. This final parameter adjustment procedure is usually referred to as *model tuning*. As model tuning aims to match observed climate system behaviour, it is connected to judgments as to what constitutes a skilful representation of the Earth's climate. For instance, maintaining the top-of-the-atmosphere energy balance in a simulation of observed historical climate is essential to prevent the climate system from drifting to an unrealistic state. The models used in this report almost universally contain small adjustments to parameters in their treatment of clouds to fulfil this important observed constraint of the climate system (Donner et al., 2011; Gent et al., 2011; Hazeleger et al., 2011; Martin et al., 2011; Mauritsen et al., 2012; Watanabe et al., 2010)
.

With very few exceptions (Mauritsen et al., 2012) modelling centres do not routinely describe in detail how they tune their models. Therefore the complete list of observational constraints toward which a particular model is tuned is generally not known. However, it is clear that tuning involves trade-offs, and this keeps the number of constraints that can be used small, and usually focused on global measures of skill related to budgets of energy, mass and momentum. It has been shown for at least one model that the tuning process does not necessarily lead to a single, unique set of parameters for a given model, but that different combinations of parameters can yield equally plausible models (Mauritsen et al., 2012). There have been recent efforts to develop systematic parameter optimization methods, but due to model complexity they have not been applied to fully coupled climate models yet (Neelin et al., 2010).

Model tuning directly influences the evaluation of climate models, as the quantities that are tuned cannot be used in model evaluation. Quantities closely related to those tuned will only provide weak tests of model performance. The use of data is integral to the model development process, and this complicates the construction of critical tests. Nonetheless, by focusing on those quantities not generally involved in model tuning while discounting metrics clearly related to it, it is possible to gain insight into model skill. The concurrent use of many model quantities, evaluation techniques, and performance metrics that together cover a wide range of emergent (or un-tuned) model behaviour ensures a stringent test of model quality.

The requirement for model tuning raises the question of whether climate models are a reliable for future climate projections. 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. What emerges is that the models that plausibly reproduce the past,universally produce significant warming under increasing greenhouse gas concentrations. It is this fact that underlies the broad consensus behind the results presented in this report.

[END BOX 9.1 HERE]

9.1.3.1 Parameterisations

Parameterisations are included in all model components to represent processes that cannot be explicitly resolved; they are evaluated both in isolation and in the context of the full model. The purpose of this section is to highlight recent developments in the parameterisations employed in each model component.

9.1.3.1.1 Atmosphere

Atmospheric models must parameterise a wide range of processes, including those associated with atmospheric convection and clouds, cloud-microphysical and aerosol processes and their interaction, boundary layer processes, as well as radiation and the treatment of unresolved gravity waves. Advances made in cloud processes and atmospheric convection are described in Chapter 7 (Section 7.2.3).

Improvements in representing the atmospheric boundary layer since the AR4 have focussed on basic boundary-layer processes, the representation of the stable boundary layer, and boundary layer clouds (Teixeira et al., 2008). Several global models have successfully adopted new approaches to the parameterization of shallow cumulus convection and moist boundary layer turbulence that acknowledge their close mutual coupling. One new development is the Eddy-Diffusivity-Mass-Flux (EDMF) approach (Neggers, 2009; Neggers et al., 2009; Siebesma et al., 2007). This approach, like the shallow cumulus scheme of (Park and Bretherton, 2009), determines the cumulus-base mass flux from the statistical distribution of boundary layer updraft properties, a conceptual advance over the ad-hoc closure assumptions used in the past. The realistic treatment of the stable boundary layer remains difficult (Beare et al., 2006; Cuxart et al., 2006; Svensson and Holtslag, 2009) with implications for the modelling of the diurnal cycle of temperature even in clear skies (Svensson et al., 2011).

The influence of internal gravity waves on the general circulation and mass distribution of the troposphere and lower stratosphere has been well established by the success of early efforts to parameterise unresolved orographic gravity-wave drag (GWD) (e.g., (Palmer et al., 1986); (McFarlane, 1987)). More recent parameterisations include sources of low-level drag such as blocking, lee vortices, downslope windstorm flow, and trapped lee waves (e.g., (Lott and Miller, 1997); (Gregory et al., 1998); (Scinocca and McFarlane, 2000). These efforts have been aided by satellite and ground-based observations of gravity-wave momentum fluxes, high-resolution numerical modelling, and focused process studies (Alexander et al., 2010).

The parameterisation of drag due to non-orographic gravity waves is becoming a common feature of GCMs that include the middle atmosphere (i.e., stratosphere and mesosphere). The basic wind and temperature structure of the middle atmosphere arises largely from a balance between radiative driving and (primarily non-orographic) GWD (Holton, 1983). The term non-orographic refers to the fact that the sources of these waves (e.g., convection and frontal dynamics) are non-stationary. In the stratosphere, such GWD is essential to the driving of both the quasi-biennial oscillation in the Tropics (Dunkerton, 1997) and the equator-to-pole residual circulation in the summer hemisphere (Alexander and Rosenlof, 1996).

9.1.3.1.2 Ocean

Mesoscale and submesoscale eddy parameterisations

Ocean components in contemporary climate models generally have horizontal resolution that is too coarse to admit mesoscale eddies. Consequently, such models typically employ some version of the (Redi, 1982) neutral diffusion and (Gent and McWilliams, 1990) eddy advection parameterisation (see also (Gent et al., 1995; McDougall and McIntosh, 2001). Since the AR4, the main focus has been on how parameterised mesoscale eddy fluxes in the ocean interior interact with boundary-layer turbulence; some CMIP5 models have implemented such features (Gnanadesikan et al., 2007) (Ferrari et al., 2008; Ferrari et al., 2010) and (Danabasoglu et al., 2008). Another focus concerns eddy diffusivity, with many CMIP5 models employing flow dependent schemes. Both of these refinements are important for the mean state and the response to changing forcing, especially in the Southern Ocean (Boning et al., 2008; Farneti and Gent, 2011; Farneti et al., 2010; Gent and Danabasoglu, 2011; Hallberg and Gnanadesikan, 2006; Hofmann and Morales Maqueda, 2011).

In addition to mesoscale eddies, there has been a growing awareness of the role that submesoscale eddies and fronts play in restratifying the mixed layer ((Boccaletti et al., 2007), (Fox-Kemper et al., 2008), and (Klein and Lapeyre, 2009)), and the parameterisation of (Fox-Kemper et al., 2011) is now used in some CMIP5 models.

Parameterisations of dianeutral transformation

There is an active research effort on the representation of dianeutral mixing associated with breaking gravity waves (MacKinnon et al., 2009), with this work adding rigor to the prototype abyssal tidal mixing parameterisation of (Simmons et al., 2004) now used in several climate models (e.g., (Jayne, 2009). The transport of dense water down-slope with gravity currents (e.g., (Legg et al., 2008; Legg et al., 2009) has also been the subject of focused efforts, with associated parameterizations making their way into some CMIP5 models (Danabasoglu et al., 2010; Jackson et al., 2008b; Legg et al., 2009).

Ocean biogeochemical (OBGC) models

Oceanic uptake of CO_2 is highly variable in space and time, and is determined by the interplay between the biogeochemical and physical processes in the ocean. About half ofCMIP5 OBGC models are based on so-called NPZD-type models that partition marine ecosystems into nutrients, plankton, zooplankton, and detritus. These models allow simulation of some of the important feedbacks between climate and oceanic CO_2 uptake, but are limited by the lack of marine ecosystem dynamics. Some efforts have been made to include more plankton groups or plankton functional types in the models (PFPs; (Le Quere et al., 2005) with as-yet uncertain implications for Earth system response.

Ocean acidification and the associated decrease in calcification in many marine organisms provides a negative feedback on atmospheric CO_2 increase (Ridgwell et al., 2007a), although the increase in the ocean-surface partial pressure of CO_2 (pCO2) which leads to ocean acidification also reduces the abiotic ocean uptake of CO_2. New-generation OBGC models therefore include various parameterisations of calcium carbonate ($CaCO_3$) production as a function of the saturation state of seawater with respect to calcite (Gehlen et al., 2007; Ilyina et al., 2009; Ridgwell et al., 2007a) or pCO2 (Heinze, 2004). On centennial scales, deep-sea carbonate sediments neutralize atmospheric CO_2. A growing number of CMIP5 models include the sediment carbon reservoir, and progress has been made towards refined sediment representation in the models (Heinze et al., 2009).

9.1.3.1.3 Land

Land-surface properties such as vegetation, soil type, and the amount of water stored on the land as soil moisture, snow, and groundwater, all strongly influence climate, particularly through their effects on surface albedo and evapotranspiration. These climatic effects can be profound; for example, it has been suggested that changes in the state of the land-surface may have played an important part in the severity and length of the 2003 European drought (Fischer et al., 2007b), and that 60% of summer temperature variability in the Mediterranean region is due to soil moisture-temperature feedbacks (Seneviratne et al., 2006).

The land-surface schemes employed in GCMs calculate the fluxes of heat, water, and momentum between the land and the atmosphere, and update the surface state variables such as soil moisture, soil temperature and snow-cover, that influence these fluxes. There has been a steady increase in the complexity of land-surface components on GCMs from the first generation soil "bucket" models employed in the 1970s (Manabe, 1969) to fourth-generation schemes that attempt to model vegetation controls on transpiration through stomatal pores on their leaves (Cox et al., 1999a; Sellers et al., 1996). However, even the more complex land surface schemes used in the AR4 suffer from obvious simplifications, such as the need to prescribe rather than simulate the vegetation cover, and a tendency to ignore lateral flows of water and sub-gridscale heterogeneity in soil moisture (Pitman, 2003).

Since the AR4, a number of climate models have included some representation of vegetation dynamics (see 9.1.3.2.4) and sub-gridscale hydrology (Gedney and Cox, 2003; Oleson et al., 2008c) and most also include a large-scale river network model to route runoff to the appropriate ocean outflow points (Oki et al., 1999).

The evaluation of land-surface schemes is in principle more straightforward than for other components of climate models because they can be tested easily in "offline" mode. The meteorological data required to drive land models is generally available and there is a growing amount of local data for evaluations (Baldocchi et al., 2001), although large-scale evaluation is still a challenge owing to the patchy coverage of flux towers.

9.1.3.1.4 Sea ice

Most large-scale sea-ice processes are well understood and well represented in models (Hunke et al., 2011a). For example, the basic thermodynamic description has been available for 40 years (Maykut and Untersteiner, 1971), and a relatively straightforward representation of sea-ice dynamics is nearly 35 years old (Hibler, 1979). Schemes like this capture the first-order behaviour of sea ice in the climate system, but may not capture important details of sea ice dynamics and deformation, especially at small scales (Coon et al., 2007; Girard et al., 2009; Hutchings et al., 2011). Since the AR4, in which there was a major advance associated with inclusion of sea-ice dynamics in most CMIP3 AOGCMs, progress in improving sea-ice components in climate models has apparently slowed. Sea ice model development is currently focused on: (1) more precise descriptions of physical processes such as microstructure evolution and anisotropy; and (2) extensions for "Earth system" simulations by including biological and chemical species.

Sea ice dynamics

The Arctic and Antarctic sea ice packs are mixtures of open water, thin first-year ice, thicker multiyear ice, and even thicker pressure ridges. An essential aspect of sea ice thermodynamics is the variation of growth and melting rates for different ice thicknesses: thin ice grows and melts more quickly than thicker ice. Similarly, thinner ice is more likely to undergo mechanical deformation than thicker ice. Most early sea ice models neglected sub-grid-scale thickness variations, but many models now include some representation of the thickness distribution, and a description of mechanical redistribution that converts thinner ice to thicker ice under convergence and shear (Hunke et al., 2010).

Sea ice thermodynamics

Sea ice albedo has long been recognized as a critical aspect of the global heat balance. The average surface albedo on the scale of a climate model grid cell is (as on land) the result of a mixture of surface types: bare ice, melting ice, snow-covered ice, open water, etc. The parameterisation of surface albedo remains uncertain and may be tuned to produce a realistic simulation of sea ice extent, compensating for deficiencies in both atmosphere and ocean forcing, (e.g., (Losch et al., 2011). Many sea ice models use a relatively simple albedo parameterisation that specifies four albedo values: cold snow; warm, melting snow; cold, bare ice; and warm, melting ice. Other models use more complex formulations that take into account the ice and snow thickness, spectral band, and surface melt (e.g., (Pedersen et al., 2009; Vancoppenolle et al., 2009b).. Solar radiation may be distributed within the ice column assuming exponential decay or via a more complex multiple-scattering radiative-transfer scheme (Briegleb and Light, 2007).

Salinity affects thermodynamic properties of sea ice, and is used in the calculation of fresh water and salt exchanges at the ice-ocean interface (Hunke et al., 2011b). Some models allow the salinity to vary in time (Schramm et al., 1997), while others assume a salinity profile that is constant, (e.g., (Bitz and Lipscomb, 1999). Vancoppenolle et al. (2009a) developed a simplified approach to simulate the desalination of Arctic sea ice as it grows and then transitions from first-year to multi-year ice. Another new thrust is the inclusion of chemistry and biogeochemistry (Hunke et al., 2011b; Piot and von Glasow, 2008a, 2008b; Vancoppenolle et al., 2010; Zhao et al., 2008), with dependencies on the ice microstructure and salinity profile. Related work involves the vertical transport and cycling of quantities such as aerosols (Bailey et al., 2010) and gases (Nomura et al., 2010) that pass gradually through the ice and can modify oceanic or atmospheric chemistry.

One of the difficulties in evaluating the sea-ice component of a climate model is that errors arise from not only the sea-ice component itself, but also from errors in the atmosphere above and the ocean below (e.g., (Bitz et al., 2002) and, because of the strong feedbacks in the system, these errors are amplified.

9.1.3.2 New Components and Couplings: Emergence of Earth System Modelling

9.1.3.2.1 Carbon cycle

The omission of internally-consistent feedbacks between the physical, chemical, and biogeochemical processes in the climate system is an inherent feature of AOGCMs. The conceptual issue is that the physical climate controls natural sources and sinks of CO_2 and CH_4, the two most important long-lived greenhouse gases. ESMs incorporate many of the important biogeochemical processes, making it possible to simulate the evolution of radiatively active species based upon their emissions from natural and anthropogenic sources together with their interactions with the rest of the Earth system. Alternatively, when forced with specified concentrations, one can diagnose these sources with feedbacks included (Hibbard et al., 2007). Given the large natural sources and sinks of CO_2 relative to its anthropogenic emissions, and given the primacy of CO_2 among anthropogenic GHGs, one of the most important enhancements is the addition of terrestrial and oceanic carbon cycles. These cycles have been incorporated into many models (Christian et al., 2010; Tjiputra et al., 2010) used to study the long-term evolution of the coupled Earth system under anthropogenic climate change (Jungclaus et al., 2010; Schurgers et al., 2008).

9.1.3.2.2 Aerosols

The treatment of aerosols has advanced since the AR4. Many ESMs now include the basic features of the sulphur cycle and so represent both the direct effect of sulphate aerosols, along with some of the more complex indirect effects involving cloud droplet number and size. Further, several ESMs are currently capable of simulating the mass, number, size distribution, and mixing state of interacting multicomponent aerosols (Bauer et al., 2008b). The incorporation of more physically complete representations of aerosols often improves the simulated climate under historical and present-day conditions, including the mean pattern and interannual variability in continental rainfall (Rotstayn et al., 2010). However, the representation of aerosols and their interaction with clouds and radiative transfer remains an important source of uncertainty (see Chapter 7). Additional aerosol related topics that have received attention include the connection between dust aerosols and ocean biogeochemistry, the production of oceanic dimethyl sulfide (DMS) (a natural source of sulphate aerosol), and vegetation interactions with organic atmospheric chemistry (Collins et al., 2011b).

9.1.3.2.3 Methane cycle and permafrost

In addition to carbon dioxide, an increasing number of ESMs are also incorporating prognostic methane to quantify the feedbacks from changes in methane sources and sinks under a warming climate. Some models now simulate the evolution of permafrost carbon stock (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).

9.1.3.2.4 Dynamic global vegetation models and wildland fires

One of the potentially more significant effects of climate change is the alteration of the distribution, speciation and life cycle of vegetated ecosystems (Bergengren et al., 2011; Bergengren et al., 2001). Vegetation has a significant influence on the surface energy balance, exchanges of non-CO_2 greenhouse gases, and the terrestrial carbon sink. Systematic shifts in vegetation, for example northward migration of boreal forests, would therefore introduce biogeophysical and biogeophysical feedbacks on the physical climate system (Clark et al., 2011). In order to include these effects in projections of climate change, several dynamic global vegetation models (DGVMs) have been developed and deployed in ESMs (Ostle et al., 2009; (Cramer et al., 2001; Sitch et al., 2008). DGVMs can simulate the interactions among natural and anthropogenic drivers of global warming, the state of terrestrial ecosystems, and ecological feedbacks on further climate change. The incorporation of DGVMs has required considerable enhancement and improvement in coupled models to produce stable and realistic distributions of flora (Oleson et al., 2008c). The improvements include better treatments of surface, subsurface, and soil hydrological processes, the exchange of water with the atmosphere, and the discharge of water into rivers and streams. While the first DGVMs have been primarily coupled to the carbon cycle, the current generation of DGVMs are being extended to include ecological sources and sinks of other non-CO_2 trace gases including CH_4, N_2O, biogenic volatile organic compounds (BVOCs), and nitrogen oxides collectively known as NO_x (Arneth et al., 2010). BVOCs and NO_x can alter the lifetime of some GHGs and act as precursors for secondary organic aerosols (SOAs) and ozone. Disturbance of the natural landscape by fire has significant climatic effects through its impact on vegetation and air quality and through its emissions of greenhouse gases, aerosols, and aerosol precursors. Since the frequency of wildland fires increases rapidly with increases in ambient temperature, the effects of fires are projected to grow over the 21st century. The interactions of fires with the rest of the climate system are now being introduced into ESMs (Arora and Boer, 2005; Pechony and Shindell, 2009).

9.1.3.2.5 Land-use / land-cover change

The impacts of land-use and land-cover change on the environment and climate are explicitly included as part of the representative concentration pathways (RCPs) used for climate projections to be assessed in later chapters (Moss et al., 2010). Several important types of land-use and land-cover change include effects of agriculture and changing agricultural practices, including the potential for widespread introduction of biofuel crops; the management of forests for preservation, wood harvest, and production of woody biofuel stock; and the global trends toward greater urbanization. ESMs include increasingly detailed treatments of crops and their interaction with the landscape (Arora and Boer, 2010; Smith et al., 2010a; Smith et al., 2010b), forest management (Bellassen et al., 2010; Bellassen et al., 2011), and the interactions between urban areas and the surrounding climate systems (Oleson et al., 2008a).

9.1.3.2.6 Chemistry-climate interactions

Ozone recovery projected for the 21st century will likely affect the climate system, for example introducing equatorward trends in the positions of the subtropical jets and strengthening the Brewer-Dobson circulation (SPARC-CCMVal, 2010; WMO, 2011). Chemistry-climate models (CCMs) are three-dimensional atmospheric climate models with fully coupled chemistry, developed specifically to explore interactions in which chemical reactions drive changes in atmospheric composition and this in turn changes the atmospheric radiative balance and hence dynamics (SPARC-CCMVal, 2010). Several stratospheric chemistry-climate modules have been incorporated into ESMs and are part of the CMIP5 ensemble (see Table 9.1). Important chemistry-climate interactions have also been identified in tropospheric ozone projections for the 21st century. For example, tropospheric ozone may increase due to positive climate feedbacks such as an increased influx of ozone from the stratosphere, which then increases radiative forcing. Several of the CMIP5 models currently simulate tropospheric chemistry interactively, and those that include tropospheric and stratospheric chemistry can be used for internally consistent simulations of the interactions among stratospheric cooling, ozone recovery, and the rest of the climate system.

9.1.3.2.7 High-top/low-top global models

It is now widely accepted that in addition to the well-known influence of tropospheric circulation and climate change on the stratosphere, stratospheric dynamics can in turn influence the tropospheric circulation and its variability (SPARC-CCMVal, 2010; WMO, 2011). As a result, many climate models now have the ability to include a fully resolved stratosphere with a model top above the stratopause (see Table 9.1). The subset of CMIP5 models with this so-called high-top configuration allows a comparison to the standard set of low-top models (i.e., those with a model top in the middle stratosphere) (Charlton-Perez and al, *2012;* Manzini and al., 2012; Wilcox et al., 2012).

9.1.3.2.8 Land ice sheets

The amount of melt water that could be released from the Greenland and Antarctic ice sheets in response to climate change remains a major source of uncertainty in projections of sea-level rise. Until recently, the long-term response of these ice sheets to alterations in the surrounding atmosphere and ocean has been simulated using offline models. Several ESMs currently have the capability to have ice-sheet component models coupled to the rest of the climate system (Vizcaino et al., 2008) although these capabilities are not exercised for CMIP5. Idealized experiments suggest that the coupling leads to accelerated melting of the Greenland ice sheet (Vizcaino et al., 2010). Uncertainties remain regarding small-scale ice-acceleration processes at the ice-sheet peripheries, particularly those involving ocean-ice interactions (see Chapter 4).

9.1.3.2.9 New features in ocean-atmosphere coupling

Several new features have arisen in the coupling between the ocean and the atmosphere since AR4. The bulk formulae used to compute the turbulent fluxes of heat, water, and momentum at the air-sea interface have been revised. A number of models now consider the surface current when calculating wind stress (e.g., (Jungclaus et al., 2006; Luo et al., 2005). The coupling frequency has been increased in some cases to include the diurnal cycle, which was shown to improve SST bias in the tropical Pacific (Bernie et al., 2008); (Ham et al., 2010). Several models now represent the coupling between the penetration of the solar radiation into the ocean and light-absorbing chlorophyll, with some implications on the representation of the mean climate and climate variability (Murtugudde et al., 2002; Wetzel et al., 2006). This coupling is achieved either by prescribing the chlorophyll distribution from observations, or by computing the chlorophyll distribution with an ocean biogeochemical model (e.g., (Arora et al., 2009).

9.1.3.3 Resolution

Improved resolution in climate models (i.e., adopting a finer grid in the modelled atmosphere, ocean and other components) is expected to improve some aspects of model performance. Generally, improved resolution leads to better representation of finer scale physical processes, as well as effects of details in topography, land-sea distribution and land cover. Model resolution needs to be developed in concert with parameterisations and their scaling with resolution in order to realize the expected improvements.

The typical horizontal resolution (defined here as horizontal grid spacing) for current AOGCMs and ESMs is roughly 1 to 2 degrees for the atmospheric component and around 1 degree for the ocean (Table 9.1). Associated with increases in computational capacity, there has been some modest increase in model resolution since the AR4, especially for the near-term simulations (e.g., around 0.5 degree for the atmosphere in some cases). On the other hand, for the models used for the long-term simulations with interactive biogeochemistry, the resolution has not increased significantly due to the trade-off against higher complexity in such models. It has been suggested that higher vertical resolution is needed to realize the full benefits of increased horizontal resolution (e.g., Roeckner et al., 2006). Since the AR4, typical regional climate model resolution has increased from around 50 km to around 25 km (see Section 9.6.2.2), and the impact of this has been explored with multi-decadal regional simulations (e.g., Christensen et al., 2010). In some cases, RCMs are being run at 10 km resolution or higher (e.g., Kanada et al., 2008; Kendon et al., 2012; Kusaka et al., 2010; van Roosmalen et al., 2010).

Higher resolution may lead to a stepwise, rather than incremental, improvement in model performance (e.g., Roberts et al., 2004; Shaffrey et al., 2009). For example, ocean models undergo a transition from laminar to eddy-permitting when the computational grid contains more than one or two grid points per first baroclinic Rossby radius (i.e., finer than 50 km at low latitudes and 10 km at high latitudes) (McWilliams, 2008; Smith et al., 2000). Such mesoscale eddy-permitting ocean models better capture the large amount of energy contained in fronts, boundary currents, and time dependent eddy features (e.g., (McClean et al., 2006b). Models run at such resolution have been used for simulations of climate time-scales (decadal to centennial) and found to be promising; though much work remains before they are as mature as the coarser models currently in use (Bryan et al., 2007; Bryan et al., 2010; Farneti et al., 2010; McClean et al., 2011).

Similarly, atmospheric models with grids that allow the explicit representation of convective cloud systems (i.e., finer than a few km) avoid employing a parameterisation of their effects -- a longstanding source of uncertainty in climate models. For example, Kendon et al. (2012) simulated the climate of the UK region over a 20-year period at 1.5 km resolution, and demonstrated several improvements of errors typical of coarser resolution models. Further discussion on this is provided in Chapter 7.

9.2 Techniques for Assessing Model Performance

9.2.1 Objectives and Limitations

Systematic evaluation of models through comparisons with observations is a prerequisite to their confident application. The objective is to understand model strengths and weaknesses in order to guide model development and to inform judgements regarding the credibility of the model-based predictions and projections.

In the AR4, the evaluation of climate models was mainly done qualitatively by comparing simulated and observed fields (e.g., time series or spatial maps). Since the AR4, performance metrics, which are statistical measures of agreement between a simulated and observed quantity (or covariability between quantities), have been more extensively used. Performance metrics derived from a variety of observationally-based diagnostics provide an objective synthesis and visualization of model performance (Cadule et al., 2010; Gleckler et al., 2008; Pincus et al., 2008a; Sahany et al., 2012; Waugh and Eyring, 2008) and enable quantitative assessment of model improvements (Reichler and Kim, 2008a). These metrics can also be used to explore the value of weighting projections based on model performance, although for this purpose the need for process-oriented evaluation, especially for those processes that are related to climatically important feedbacks, has been emphasized (Eyring et al., 2005; Knutti et al., 2010b; Neelin et al., 2010).

Despite these developments, quantitative assessment of climate model skill is still limited for a number of reasons (Hargreaves, 2010; Knutti et al., 2010b). Unlike weather prediction models, which make specific predictions for specific times and can be tested against subsequent observations, climate models are concerned with climatological distributions, and so require long records to obtain suitable samples. Climate model evaluation therefore requires the availability of long-term, consistent, error-characterized global and regional Earth observations (satellite and in situ) as well as accurate globally gridded reanalyses in the atmosphere, the ocean, or, ultimately, the coupled system. Since the AR4, the Earth Observation community has undertaken a large effort to develop data sets of selected Essential Climate Variables (ECVs). If available, observational uncertainty can be included in model evaluation either by using error estimates provided with the observational data set, or by using more than one data set. In many cases, the lack of long-term observations, observations suitable for the evaluation of important processes, or observations in particular regions (e.g., polar areas, the upper troposphere / lower stratosphere (UTLS), and the deep ocean) remains an impediment.

9.2.2 New Developments in Model Evaluation Approaches

In this section we provide a brief overview of progress in the methods used to evaluate models as this is a rapidly evolving field. In many cases we later show results from these new methods, but in some cases applications of these new methods are too preliminary to report on.

9.2.2.1 Evaluating the Overall Model Results

The most straightforward approach to evaluate models is to compare simulated quantities (e.g., global distributions of temperature, precipitation, radiation etc.) with corresponding observations (e.g., Gleckler et al., 2008; Pincus et al., 2008a; Reichler and Kim, 2008b). This approach can be extended to evaluate models in terms of variability and extreme events by comparing simulated spatial structure of variability modes or extreme indices with observationally based counterparts (e.g., AchutaRao and Sperber, 2006; Oshima and Tanimoto, 2009; Sillmann et al., 2012; see Section 9.5 for further discussion). For quantitative comparison, statistical measures are often used as performance metrics (e.g., root-mean square error, centred and uncentred pattern correlations). Some studies aggregate a number of metrics to form a single metric (Gleckler et al., 2008; Murphy et al., 2004; Pierce et al., 2009; Reichler and Kim, 2008b; Santer et al., 2009a), with more recent work aimed at reducing redundancy of multiple metrics through methods such as cluster analysis (Nishii et al., 2011; Yokoi et al., 2011).

9.2.2.2 Isolating Processes

It is often illuminating to evaluate the representation of key processes both in the context of the full model, and in isolation. A number of evaluation techniques to achieve both process and component isolation have been developed. One involves the so-called “regime-oriented” approach to process-evaluation. Instead of averaging model results in time (e.g., seasonal averages) or space (e.g., global averages), results are averaged within categories that describe physically distinct regimes of the system under study. Applications of this approach since AR4 include the use of circulation regimes (Bellucci et al., 2010; Bony and Dufresne, 2005; Brown et al., 2010b), cloud regimes (Chen and Del Genio, 2009; Williams and Webb, 2009; Williams and Brooks, 2008), or thermodynamics states (Sahany et al., 2012; Su et al., 2011). The importance of the regime-oriented approach lies in its ability to isolate processes that might be responsible for particular errors (Jakob, 2010).

Another approach involves either the removal of a particular model component or process parameterisation for use in off-line simulations. Results of such simulations are compared to measurements from detailed field studies or to results from more sophisticated process models (Randall et al., 2003). Numerous important process-related data sets to support such evaluations have been collected since the AR4 (Illingworth et al., 2007; May et al., 2008; Redelsperger et al., 2006; Verlinde et al., 2007; Wood et al., 2011) and have been applied to the evaluation of climate model processes (Boone et al., 2009; Boyle and Klein, 2010; Hourdin et al., 2010; Xie et al., 2008). These studies are crucial to test the realism of the process formulations that underpin climate models.

9.2.2.3 Instrument Simulators

Satellites provide nearly global coverage, sampling across many meteorological conditions. This makes them powerful tools for model evaluation. The conventional approach has been to convert satellite-observed radiation information to *model-equivalents* through so-called retrievals (Stephens and Kummerow, 2007). Retrieved properties have been used in numerous studies to evaluate simulations of clouds and precipitation (Allan et al., 2007; Gleckler et al., 2008; Jiang et al., 2012b; Li et al., 2011a; Pincus et al., 2008b). The main challenge is that modelled and retrieved variables are difficult to define consistently due to limitations of the satellite sensors and the assumptions used in the retrievals. These limitations vary across different satellite instruments.

An alternative approach is to calculate *observation-equivalents* from models using radiative transfer calculations to 'simulate' what the satellite would provide if the satellite system were 'observing' the model. This approach is usually referred to as an “instrument simulator”. While not always a true description of the specific tool used, this term is used to describe the “observation-equivalent” approach to model evaluation. Microphysical assumptions (which differ from model to model) can be included in the simulators, avoiding retrieval inconsistencies. A simulator for cloud properties from the International Cloud Satellite Climatology Project (ISCCP) (Klein and Jakob, 1999; Webb et al., 2001; Yu et al., 1996) has been widely used for model evaluation since AR4 ((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). New simulators for other satellite products have also been developed and are increasingly applied for model evaluation (Bodas-Salcedo et al., 2011). While often focussed on clouds and precipitation, the simulator approach has also been used successfully for other variables such as upper tropospheric humidity (Allan et al., 2003; Brogniez et al., 2005; Iacono et al., 2003; Ringer et al., 2003; Zhang et al., 2008b) (Bodas-Salcedo et al., 2011; Brogniez and Pierrehumbert, 2007).

9.2.2.4 Paleoclimate Studies

Past climates offer a wide range of climatic states that can be used to test a model’s response to different forcing (see Chapter 5); however this can be achieved only for periods with sufficient data coverage. Such data sets have been developed for the Last Glacial Maximum (21,000 years BP) and the mid-Holocene (6000 years BP), as part of the global ocean reconstruction from marine data (CLIMAP, 1981; Waelbroeck et al., 2009) and the Biome 6000 project (Prentice et al., 1998).

Paleo proxies, such as pollen or $\delta^{18}O$ in ice cores, are indirect measurements of climatic conditions, and so care must be taken to compare a modelled climate variable that best characterizes the major fluctuations of the proxy indicator. Recent work on marine proxies suggests that, depending on the region, the same proxy is not necessarily dominated by the same aspect of climate (Jungclaus et al., 2010).

An alternative 'forward modelling' approach consists of simulating the proxy indicators themselves. Some ESMs now include a dynamical vegetation module in their land surface scheme, and so simulated results can be compared directly to past vegetation reconstructions (Braconnot et al., 2007c). Some models can be run with a representation of water isotopes which allows direct comparison of model output with isotopic measurements (LeGrande et al., 2006; Tagliabue et al., 2009).

9.2.2.5 Use of Data Assimilation and Initial Value Techniques

To be able to forecast the weather a few days ahead, knowledge of the present state of the atmosphere is of primary importance. In contrast, climate predictions and projections simulate the *statistics* of weather seasons to centuries in advance. Despite their differences, both weather predictions and projetions of future climate are performed with very similar atmospheric model components. The atmospheric component of climate models can be integrated essentially as a weather prediction model if initialised appropriately (Phillips et al., 2004). This allows testing some parameterised sub-grid scale processes without the complication of feedbacks substantially altering the underlying state of the atmosphere.

The application of these techniques since AR4 has led to some new insights. For example, many of the systematic errors in the modelled climate develop within a few days of simulation, highlighting the important role of fast, parameterised processes in contributing to these errors (Boyle et al., 2008; Klein et al., 2006; Xie et al., 2012). Errors in cloud properties for example were shown to be present from very early on in a forecast in at least some models (Williams and Brooks, 2008), although this was not the case in another model (Boyle and Klein, 2010; Zhang et al., 2010b). Other studies have highlighted the advantage of such methodologies for the detailed evaluation of model processes using observations that are only available for limited locations and times (Bodas-Salcedo et al., 2008; Boyle and Klein, 2010; Hannay et al., 2009; Williamson and Olson, 2007; Xie et al., 2008), an approach that is difficult to apply to long-term climate simulations. As with atmospheric data assimilation, it is evident that ocean data assimilation will provide a useful opportunity for the assessment of ocean processes at their characteristic timescales (Balmaseda et al., 2008; Bell et al., 2004).

9.2.2.6 Evaluation Techniques for RCMs

Evaluation of RCMs may involve the same global data sets used for AOGCM evaluation, or more regionally-specific climatologies. A complication is that biases in an RCM arise from both biases in the boundary conditions and the representation of regional processes in the RCM itself (Deque et al., 2012; Déqué et al., 2007). Evaluation of RCMs therefore often involves simulations with global reanalyses as boundary conditions to minimize boundary condition errors (Christensen et al., 1997). In addition, reanalysis-forced simulations produce sequences of climate variability that can be directly compared to observed time series, in addition to the more usual summary statistics over some longer period. Since the AR4, multi-decadal reanalysis-based RCM evaluation runs have become common (e.g., Christensen et al., 2010).

9.2.2.7 Evaluation Techniques for Ensemble Approaches

Ensemble methods are used to explore the uncertainty in climate model simulations that arise from internal variability, boundary conditions, parameter values for a given model structure, or structural uncertainty due to different model formulations (Hawkins and Sutton, 2009; Knutti et al., 2010a; Tebaldi and Knutti, 2007b). Since the AR4, techniques have been designed to specifically evaluate model performance of individual ensemble members. While this is typically done to better constrain the uncertainties explored by these ensembles, the methods and insights are applicable to model evaluation in general.

The ensembles are generally of two types: Multi-model Ensembles (MMEs) and Perturbed Parameter (or sometimes Physics) Ensembles (PPEs). The MME is created from existing model simulations from multiple climate modelling centres. MMEs sample structural uncertainty and internal variability. However, the sample size of MMEs is small and is confounded because some climate models have been developed by sharing model components leading to shared biases (Masson and Knutti, 2011b). Thus, MME members cannot be treated as purely independent and this implies a reduction in the effective number of independent models (Jun et al., 2008b; Knutti, 2010; Knutti et al., 2010a; Pennell and Reichler, 2011a; Tebaldi and Knutti, 2007b). In contrast, PPEs are created to assess uncertainty based on a single model, are typically larger than MMEs, and benefit from the explicit control on parameter perturbations. This allows statistical methods to determine which parameters are the main drivers of uncertainty across the ensemble (e.g., (Rougier et al., 2009a)). PPEs have been used frequently in simpler models such as EMICs, (Forest et al., 2006, 2008; Forest et al., 2002; Knutti and Tomassini, 2008; Loutre et al., 2011; Sokolov et al., 2009; Stott and Forest, 2007; Xiao et al., 1998) and are now being applied to more complex models (Annan et al., 2005; Brierley et al., 2010; Collins et al., 2007; Collins et al., 2006a; Jackson et al., 2008a; Klocke et al., 2011b; Lambert et al., 2012; Murphy et al., 2004; Stainforth et al., 2005b). The disadvantage of PPEs is that they do not explore structural uncertainty. Also, the extent to which a PPE explores uncertainty can depend on the underlying model that is perturbed (Yokohata et al., 2010a). Recognising the importance of sampling both parametric uncertainty and structural uncertainty, (Sanderson, 2012) and (Sexton et al., 2012) both combine information from MMEs and PPEs. However even these approaches cannot account for the effect on uncertainty of systematic errors that are common to all of the current generation of climate models.

Bayesian methods have been developed for both MMEs (Furrer et al., 2007a; Greene et al., 2006; Tebaldi and Sanso, 2009; Tebaldi et al., 2004) and PPEs (Sexton et al., 2012; Rougier, 2007) which allow different ensemble members or different variants of a particular model to be weighted by a value between 0 and 1 inclusive, according to their ability to simulate some aspects of historical climate. Rougier (2007) presents a rigorous Bayesian framework to produce probabilistic climate projections for a given set of boundary conditions that is underpinned by a PPE and that accounts for the key uncertainties: internal variability, parametric uncertainty, and uncertainty due to structural differences. Using the Rougier (2007) approach, Sexton et al. (2012) apply the framework to estimate ECS and demonstrate that accounting for model imperfections weakens the observational constraint and avoids over-confident projections. Many metrics that aim to weight climate models or simply compare their performance do not account for this effect of structural errors.

The Bayesian methods offer insights into how to account for model inadequacies and combine information from several metrics (Sexton and Murphy, 2012; Sexton et al., 2012) but they are complex. A simpler strategy of screening out some model variants on the basis of some observational comparison has been used with some PPEs (Lambert et al., 2012; Shiogama et al., 2012). Edwards et al. (2011) provided a statistical framework for "pre-calibrating" out such poor model variants using an Implausibility metric and thus avoids deceptively large uncertainty ranges as shown in Williamson et al. (2012). Screening techniques have also been used with MMEs (Santer et al., 2009b).

Additional Bayesian methods are applied to the MMEs so that past model performance is combined with prior distributions to estimate uncertainty from the MME (Furrer et al., 2007b; Milliff et al., 2011; Tebaldi and Knutti, 2007b). Similar to Bayesian PPE methods, common biases can be assessed within the MME to determine effective independence of the climate models (Knutti et al., 2012) (see Section 12.2.2 for a discussion of the assumptions in the Bayesian approaches.)

9.2.3 Overall Summary of Model Evaluation Approach in this Chapter

Exploitation of the most comprehensive set of observations necessitates an emphasis on recent decades, although older 20th century records and paleo data also play an important role. In some circumstances valuable insight into a model’s behaviour can be achieved without observations (via analysis of inter-model differences), but we will use this approach sparingly.

A rational progression of such a broad scope evaluation begins with an examination of the large-scale features of the mean state in each of the model components (Section 9.4). This is followed by an evaluation of the ability of models to capture the dominant features of natural variability on observable time scales, including extremes (Section 9.5). This path of increasing focus takes us to more regional evaluation of model performance, including approaches to augment regional information with downscaling techniques (Section 9.6). Throughout our evaluation, we rely on routine diagnostic methods to compare model simulations with observations, such as spatial maps and space or time decompositions (e.g., zonal means or anomaly time series). As the evaluation focuses on increasing detail, a sampling of more sophisticated diagnostic approaches will also be exploited. To complement these diagnostics, we also rely on performance metrics to quantify the level of agreement between models and observations. Performance metrics provide an approach to succinctly summarize model performance and to quantify changes in model performance over time.

Some factors contributing to the quality of a simulation, such as uncertainties in specified external forcing,,are beyond the scope of this chapter, but some of these are discussed in Chapter 10. The assessment in this chapter focuses on a multi-model perspective (e.g., CMIP3, CMIP5) in which the inter-model spread provides at least some estimate of model uncertainty. However, the error structure of model behaviour is extremely complex (e.g., (Santer et al., 2009a), and it must be emphasized that the relative performance of individual models can vary widely from one diagnostic/metric to another. The prospects for synthesizing this information to gauge the reliability of projections are addressed in Section 9.8.

9.3 Experimental Strategies in Support of Climate Model Evaluation

9.3.1 The Role of Model Intercomparisons

Gauging the extent to which climate models realistically simulate the Earth's climate and capture fundamental processes requires extensive comparisons with observations on a range of space and time scales. Organized model intercomparison projects (MIPs) serve a variety of purposes for the climate research community and typically include standard or "benchmark" experiments that represent critical tests of a model's ability to simulate the observed climate. When modelling centres perform a common experiment, it offers the possibility to compare their results not just with observations, but with other models as well. This "intercomparison" enables researchers to explore the various strengths and weakness of different models in a controlled setting. Model evaluation is a necessary step towards identification of model error. Benchmark MIP experiments offer a way to distinguish between errors particular to an individual model and those which might be more universal. The resulting multi-model perspective provides the context for much of what follows.

9.3.2 Experimental Strategy for CMIP5

9.3.2.1 Structure of the Historical Experiments

The fifth Coupled Model Intercomparison Project (CMIP5) includes a much more comprehensive suite of model experiments than was available in the preceding CMIP3 results assessed in AR4 (Meehl et al., 2007a). In addition to a better constrained specification of historical forcing, the CMIP5 collection also includes initialized decadal-length projections and long-term experiments using ESMs and AOGCMs (Taylor et al., 2012) (Figure 9.1). The observable properties of the basic mean states from these experiments are evaluated against the historical data record in the next Section. This assessment addresses two principal requirements that climate models must satisfy in order to provide useful projections of climate change. First, it is necessary for climate models to reproduce the observed state as accurately as possible to minimize the effects of state-related errors on projections of future climate. Second, many relationships among climatic forcing, feedback, and response manifested in projections of future climate change can be tested using the observational record (Soden and Held, 2006). However, agreement with the observational record is a necessary but not sufficient condition to narrow the range of uncertainty in projections due, e.g., to remaining uncertainties in historical forcing, recent trends in oceanic heat storage, and the internal variability of the climate system (Klocke et al., 2011c).

[INSERT FIGURE 9.1 HERE]

Figure 9.1: Left: Schematic summary of CMIP5 short-term experiments with tier 1 experiments (yellow background) organized around a central core (pink background). From (Taylor et al., 2012), their Figure 2. Right: Schematic summary of CMIP5 long-term experiments with tier 1 and tier 2 experiments organized around a central core. Green

font indicates simulations to be performed only by models with carbon cycle representations, and "E-driven" means "emission-driven". Experiments in the upper hemisphere either are suitable for comparison with observations or provide projections, whereas those in the lower hemisphere are either idealized or diagnostic in nature, and aim to provide better understanding of the climate system and model behaviour. From (Taylor et al., 2012), their Figure 3.

9.3.2.2 Forcing of the Historical Experiments

Under the protocols adopted for CMIP5 and previous assessments, the transient climate experiments are conducted in three phases. The first phase covers the start of the modern industrial period through to the present-day corresponding to years 1850 to 2005 (van Vuuren et al., 2011). The second phase covers the future, 2006 to 2100, and is described by a collection of Representative Concentration Pathways (Moss et al., 2010). The third phase is described by a corresponding collection of Extension Concentration Pathways (Meinshausen et al., 2011). The forcings for the first phase are relevant to the historical simulations evaluated in this Section and are described briefly here (with more details in Annex II).

In the CMIP3 20th century experiments experiments, the forcings from radiatively-active species other than long-lived greenhouse gases and sulphate aerosols were left to the discretion of the individual modelling groups (IPCC, 2007). By contrast, a comprehensive set of historical anthropogenic emissions and land-use and land-cover change data have been assembled for the CMIP5 experiments in order to produce a relatively homogeneous ensemble of historical simulations with common time-series of forcing agents.

For AOGCMs without chemical and biogeochemical cycles, the forcing agents are prescribed as a set of concentrations. The concentrations for GHGs and related compounds include CO_2, CH_4, N_2O, all fluorinated gases controlled under the Kyoto Protocol (HFCs, PFCs, and SF_6), and ozone depleting substances controlled under the Montreal Protocol (CFCs, HCFCs, Halons, CCl_4, CH_3Br, CH_3Cl). The concentrations for aerosol species include sulphate (SO_4), ammonium nitrate (NH_4NO_3), hydrophobic and hydrophilic black carbon, hydrophobic and hydrophilic organic carbon, secondary organic aerosols (SOA), and four size categories of dust and sea salt. For ESMs that include chemical and biogeochemical cycles, the forcing agents are prescribed both as a set of concentrations and as a set of emissions with provisions to separate the forcing by natural and anthropogenic CO_2 (Hibbard et al., 2007) The emissions include time-dependent spatially-resolved fluxes of CH_4, NO_X, CO, NH_3, black and organic carbon, and volatile organic carbon (VOCs). For models that treat the chemical processes associated with biomass burning, emissions of additional species such as C_2H_4O (acetaldehyde), C_2H_5OH (ethanol), C_2H_6S (dimethyl sulphide), and C_3H_6O (acetone) are also prescribed. Historical land-use and land-cover change is described in terms of the time-evolving partitioning of land-surface area among cropland, pasture, primary land and secondary (recovering) land, including the effects of wood harvest and shifting cultivation, as well as land-use changes and transitions from/to urban land (Hurtt et al., 2009). These emissions data are aggregated from empirical reconstructions of grassland and forest fires (Mieville et al., 2010; Schultz et al., 2008), international shipping (Eyring et al., 2010), aviation (Lee et al., 2009), sulphur (Smith et al., 2011c), black and organic carbon (Bond et al., 2007), and NO_X, CO, CH_4 and NMVOCs (Lamarque et al., 2010) contributed by all other sectors.

9.3.2.3 Relationship of Observational Initialization and Decadal Predictive Uncertainty

The CMIP5 archive also includes a new class of decadal-prediction experiments (Meehl et al., 2009) (Figure 9.1). The goal is to understand the relative roles of forced changes and internal variability in historical and near-term climate variables, and to assess the predictability that might be realized on decadal time scales. These experiments are comprised of two sets of hindcast and prediction ensembles with initial conditions spanning 1960 through 2005. The set of 10-year ensembles are initialized starting at 1960 in 5-year increments through the year 2005 while the 30-year ensembles are initialized at 1960, 1980, and 2005. Results from these experiments will be described in detail in Chapter 11; here we focus on evaluation of the models used in such predictions.

9.4 Simulation of Recent and Longer-Term Records in Global Models

9.4.1 Atmosphere

Many aspects of the atmosphere have been more extensively evaluated than other climate model components. One reason is the availability of near-global observationally-based data for energy fluxes at the top of the atmosphere, cloud cover and cloud condensate, temperature, winds, moisture, total column ozone, and other important properties. As discussed in Chapter 3, atmospheric reanalysis have also been instrumental in integrating independent observations in a physically consistent manner. In this section we use this diversity of data to evaluate the large-scale atmospheric behaviour with emphasis on key physical processes (for example clouds and water vapour, which are critical in determining climate sensitivity).

9.4.1.1 Spatial Patterns of the Mean State

As discussed in Section 9.1, all component models of the Earth system are built upon fundamental principles such as the conservation of energy, momentum, and mass. For the atmospheric component, realistic simulation of the energy and water cycles is particularly important. Surface temperature is perhaps the most routinely examined quantity in atmospheric models. The surface is influenced by many factors that must be adequately represented in order for a model to realistically capture the observed temperature distribution. The dominating external influence is incoming sunlight, but many aspects of the simulated climate play an important role in modulating regional temperature such as the presence of clouds and the complex interactions between the atmosphere and the underlying land, ocean, snow, ice, and biosphere.

The annual mean distribution of surface air temperature (at 2 meters) is shown in Figure 9.2 for the multi-model average of CMIP5 models presently available. The meridional temperature gradient is qualitatively similar to the observations, which although not shown is evident from the CMIP5 climatology bias. The maximum annual mean temperatures of the western tropical Pacific and tropical Indian Ocean are also well represented by the models. A comparison of the multi-model mean with an observationally-based reanalysis (Dee et al., 2011) shows that in most areas the models agree with the observations to within 2°C, but there are several locations were the biases are much larger, particularly at elevations over the Himalayas and parts of both Greenland and Antarctica. Arctic surface inversions in models have been compared with radiosonde data and identify significant model bias (Zhang et al., 2011). The mean absolute error of the individual CMIP5 models provides a similar picture as the mean bias, implying that compensating errors across models is not a major factor (Figure 9.2). Furthermore, the inconsistency across reanalyses is smaller than the mean absolute bias in almost all regions, further enhancing confidence in the mean bias as a measure of model quality (Figure 9.2).

[INSERT FIGURE 9.2 HERE]
Figure 9.2: Annual mean surface (2 meter) air temperature (°C) for the period (1980–2005). Top left: Multi-model (ensemble) mean constructed with all available models used in the CMIP5 historical experiment. Top right: Multi-model mean bias as the difference between the CMIP5 multi-model mean and the climatology from ERA-Interim (1990–2005, (Dee et al., 2011)). Bottom left: Mean absolute model error with respect to the climatology from ERA-Interim. Bottom right: Mean inconsistency between three different reanalysis products as the mean of the absolute pair-wise differences between those fields.

A first look of the seasonal performance of models can be obtained by examining the difference between extreme seasons (DJF and JJA). Figure 9.3 shows multi-model mean seasonal cycle amplitude (as measured by the difference between the DJF and JJA surface air temperature). This clearly demonstrates the much larger seasonal cycle over land, particularly at higher latitudes. The bottom panels of Figure 9.3 shows the multi-model mean bias of the seasonal cycle relative to the reference data (Dee et al., 2011).

[INSERT FIGURE 9.3 HERE]
Figure 9.3: Annual surface (2 meter) air temperature (°C) range (DJF-JJA) for the period (1980–2005). Top left: multi-model mean of the DJF-JJA seasonality, calculated from all available CMIP5 models for the historical experiment. Top right: the mean absolute multi-model seasonality. Bottom left: the difference between the multi-model-mean and the ERA-Interim seasonality. Bottom right: the difference between the multi-model-mean and the ERA-Interim absolute seasonality.

Simulation of precipitation is a much tougher test for models as it depends heavily on processes that are not explicitly resolved, and must be parameterised. Figure 9.4 shows precipitation simulated by the CMIP5 multi-model ensemble, along with measures of error relative to observations. Many studies make use of several estimates of precipitation to illustrate the large uncertainties that exist in certain regions. Known large-scale features are reproduced by the multi-model mean, such as a maximum precipitation denoting the ITCZ just north of the equator in the central and eastern tropical Pacific, dry areas over the eastern subtropical ocean basins, and the minimum rainfall in Northern Africa (Dai, 2006). While many large-scale features of the tropical circulation are reasonably well simulated, there are persistent biases. These biases contribute to the low pattern correlation for precipitation climatology (Figure 9.4) and aspects may be seen in annual average precipitation field error maps in Figure 9.4 such as low precipitation along the equator in the Western Pacific associated with ocean-atmosphere feedbacks maintaining the equatorial cold tongue (Collins et al., 2010) and excessive precipitation in tropical convergence zones south of the equator in the Atlantic and the Eastern Pacific (Lin, 2007; Pincus et al., 2008b). Regional-scale precipitation simulation has strong parameter dependence (Chen et al., 2010; Neelin et al., 2010; Rougier et al., 2009b), and in some models substantial improvements have been shown by improvements of resolution and of representation of subgrid scale processes, in particular of convection (Delworth et al., 2011; Neale et al., 2008).

[INSERT FIGURE 9.4 HERE]

Figure 9.4: Annual mean precipitation for the period (1980–2005). Top left: Multi-model mean constructed with all available AOGCMs used in the CMIP5 historical experiment. Top right: difference between multi-model mean with and observations (Adler et al., 2003). Bottom left: Mean absolute model error with respect to observations. Bottom right: multi-model mean error relative to the multi model mean precipitation itself.

9.4.1.1.2 Atmospheric moisture, clouds, and radiation

The global annual-mean precipitable water is the measure of the total moisture content of the atmosphere. For the CMIP3 ensemble, the values of precipitable water agreed with one another and with multiple estimates from the NCEP/NCAR and ERA meteorological reanalyses to within approximately 10% (Waliser et al., 2007). Modeling the vertical partitioning of water vapor is subject to greater uncertainty since the humidity profile is governed by a variety of hydrological processes, sub-grid scale vertical transport, and coupling between the boundary layer and free troposphere. In general, the CMIP3 models exhibit a significant dry bias of up to 25% in the boundary layer and a significant moist bias in individual layers of the free troposphere of up to 100% (John and Soden, 2007). Upper tropospheric water vapor varies by a factor of three across the multi-model ensemble (Su et al., 2006). The CMIP3 models reproduce the gradients in free-tropospheric humidity between ascending and descending dynamical regimes and between convective-cloud-covered and cloud-free regions of the tropics to within 10% (Brogniez and Pierrehumbert, 2007), and the relationship between tropospheric moisture and externally forced warming in the 20th century is consistent across the ensemble and uncorrelated with the biases in the individual models (John and Soden, 2007). Comparable analyses for the CMIP5 models are not yet available.

The spatial patterns and annual cycle of the radiative fluxes at the top of the atmosphere represent some of the most important observable properties of the Earth system, and CMIP3 models reproduce these patterns with considerable fidelity relative to the NASA CERES data sets (Pincus et al., 2008a). This level of agreement is as expected since the spatial patterns and the annual cycle of radiative fluxes are governed primarily by the meridional gradient and seasonal cycle in solar insolation, both of which are reasonably reproduced by all the models in the CMIP3 ensemble. However, the models systematically underestimate the albedo during boreal summer at 60°N and austral summer at 60°S (Bender et al., 2006). The CMIP3 models also exhibit much less skill in reproducing either the spatial correlations or spatial variance in shortwave and longwave cloud radiative effects (see below), particularly in the subtropical stratus-cloud regions west of Africa and North and South America (Bender et al., 2006).

Comparisons against surface components of radiative fluxes show that, on average, the CMIP5 models overestimate the downward all-sky shortwave flux at the surface by 2 ± 6 W m^{-2} (1 ± 3%) (Stephens et al., 2012a) and underestimate the downward longwave flux by 6 ± 9 W m^{-2} (2 ± 2%) (Stephens et al., 2012b). The resulting average underestimate in the total downwelling radiant flux is 4 W m^{-2}. While in tropical regions between 1 to 3 W m^{-2} of the bias may be due to systematic omission of precipitating ice hydrometeors (Waliser et al., 2011b), the correlation between the biases in the all-sky and clear-sky downwelling fluxes suggests that systematic errors in clear-sky radiative transfer calculations may be a primary cause for these biases. This is consistent with an analysis of the global annual-mean estimates of clear-sky atmospheric absorption from the CMIP3 ensemble. The underestimation of absorption can be attributed to the omission or underestimation of absorbing aerosols, in particular carbonaceous species, and to the omission of weak-line absorption by water vapour, the predominant absorbing gas for shortwave radiation in the current climate (Wild et al., 2006). The net shortwave energy absorbed by the surface is set by the downwelling flux and the surface albedo. The difference between the mean surface albedo of 0.351 from the CMIP3 ensemble and the observationally derived albedo of 0.334 from the International Satellite Cloud Climatology Project (ISCCP) is statistically significant since the difference is approximately three times the standard deviation in surface albedo among the models (Donohoe and Battisti, 2011).

One of the major influences on radiative fluxes in the atmosphere is the presence of clouds and their radiative properties. To measure the influence of clouds on model deficiencies in the top of the atmosphere radiation budget, Figure 9.5 shows maps of deviations from observations in annual mean shortwave (top left), longwave (middle left) and net (bottom left) cloud radiative effect (CRE) for the CMIP5 multi-model mean. The Figure also shows zonal averages of the same quantities from two sets of observations (thick solid and dashed black line), the individual CMIP5 models (grey lines), and the multi-model average (thick red line). The definition of CRE and observed mean fields for these quantities can be found in Chapter 7 (Section 7.2.1.2).

[INSERT FIGURE 9.5 HERE]

Figure 9.5: Annual mean errors in shortwave (top left), longwave (middle left) and net (bottom left) cloud radiative effect of the CMIP3 multi-model mean. Shown on the right are zonal averages of the absolute values of the same quantities from observations (solid black: CERES EBAF 2.6; dashed black: CERES ES-4), individual models (thin grey lines), and the multi-model mean (thick red line). For a definition of cloud radiative effect and maps of its absolute values, see Chapter 7.

Models show large regional biases in CRE in the shortwave component, and these are particularly pronounced in the subtropics with too weak an effect of the model clouds on shortwave radiation in the stratocumulus regions and too strong an effect in the trade cumulus regions. A too weak cloud influence on shortwave radiation is also evident over the sub-polar oceans of both hemispheres and the Nothern Hemisphere land areas. It is evident in the zonal mean graphs that there is a wide range of errors in both longwave and shortwave CRE between individual models. As is also evident, a significant reduction in the difference between models and observations has resulted from changes in the observational estimates of CRE, in particular at polar and sub-polar as well as sub-tropical latitudes (Loeb et al., 2009).

Understanding the errors in CRE in models requires a more in-depth analysis of the errors in cloud properties, including the fractional coverage of clouds as well as their liquid water and ice content. Major progress in this area has resulted from both the availability of new observational data sets and improved diagnostic techniques, including the increased use of instrument simulators. Several studies have identified significant progress in the simulation of clouds in the CMIP5 models compared to their CMIP3 counterparts (Jiang et al., 2012b; Klein et al., 2012; Li et al., 2012c). Particular examples include the improved simulation of vertically integrated ice water path (Jiang et al., 2012b; Li et al., 2012c) and cloud liquid water path (Jiang et al., 2012a) as well as a reduction of overabundant optically thick clouds in the mid-latitudes (Klein et al., 2012). In many of the above variables models are now falling within the often-large observational uncertainties.

Despite demonstrable overall progress, errors in the simulations of clouds in individual models remain. 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). The global mean fraction of clouds that can be detected with confidence from satellites (optical thickness >1.3, (Pincus et al., 2012)) is underestimated by 5 to 10 % (Klein et al., 2012). With the emergence of vertically better resolved cloud information from satellite observations it has become apparent the models have difficulties in simulating clouds and their associated water vapour fields in the tropical upper troposphere (Jiang et al., 2012b). Some of the above errors in clouds presumably compensate to provide the global balance in radiation required by model tuning (Box 9.1).

The overall improvement in the cloud simulations from CMIP3 to CMIP5 is the result of numerous studies aimed at identifying the causes for particular model deficiencies. Subtropical clouds have been shown to be of great importance to a model's climate sensitivity (Bony and Dufresne, 2005; Dufresne and Bony, 2008; Williams and Webb, 2009). In-depth analysis of several global and regional models (Karlsson et al., 2008; Teixeira et al., 2011) has shown that the interaction of boundary layer and cloud processes with the larger scale circulation systems that ultimately drive the observed subtropical cloud distribution remain poorly simulated. Large errors in subtropical clouds have been shown to negatively affect SST patterns in coupled model simulations (Hu et al., 2011; Wahl et al., 2011).

Several studies have highlighted the potential importance and poor simulation of sub-polar clouds in the Arctic and Southern Oceans (Haynes et al., 2011; Karlsson and Svensson, 2010; Trenberth and Fasullo, 2010b; Tsushima et al., 2006), and (Karlsson and Svensson, 2010) showed that the CMIP3 models have great difficulties in simulating Arctic cloud properties. 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). Regime-oriented approaches to the evaluation of model clouds (Teixeira et al., 2011; Williams and Webb, 2009; Williams and Tselioudis, 2007; Williams and Brooks, 2008) are beginning to provide deeper insight into model errors and strategies for model improvement.

In summary, there remain significant errors in the model simulation of clouds. It is very likely that these errors contribute significantly to the uncertainties in estimates of cloud feedbacks (see Section 9.7.4 and Chapter 7) and consequently in the climate change projections reported in Chapter 12.

9.4.1.3 Quantifying Model Performance with Metrics

Performance metrics can be constructed to quantify what models simulate well and to demonstrate model performance deficiencies. The have been used to some extent in the TAR and AR4, and are expanded upon here because of their increased appearance in the recent literature. As a simple example, Figure 9.6 illustrates how the pattern correlation between the observed and simulated climatological annual mean depends very much on the quantity examined. All CMIP3 models capture the mean surface temperature distribution quite well, with correlations of 0.95 and higher, which is largely dictated by the dominant meridional temperature gradient. Correlations for outgoing longware radiation are somewhat lower. For precipitation and the TOA shortwave radiation however, the typical correlation between models and observations is below 0.8, with considerable scatter. This example illustrates how fields associated with the large-scale atmospheric circulation (e.g., temperature) agree more closely with observations than fields directly related to parameterisations (e.g., precipitation and clouds and their radiative effects). Challenges associated with simulated precipitation are compounded by the link to surface fields (topography, coastline, vegetation) that lead to much greater spatial heterogeniety at regional scales, and makes it both harder to model and validate against point-based observations. Incremental improvement in the multi-model mean is also evident in each field, with the CMIP5 ensemble having slightly smaller errors than CMIP3.

[INSERT FIGURE 9.6 HERE]

Figure 9.6: Global annual mean climatology (1980–1999) centred pattern correlations between models and observations. Results are shown for individual models from CMIP3 and CMIP5 (black dashes) and the average result for each (red for CMIP3, blue for CMIP5). The four variables shown are surface air temperature (TAS), top-of-atmosphere (TOA) outgoing longwave radiation (RLUT), precipitation (PR), and TOA shortwave cloud radiative effect (SW CRE). The observations used for each variable are the default products and climatological periods identified in Table 9.2. The centred pattern correlations are computed at a resolution of 5 degrees in longitude and latitude. Only one realization is used from each model from the CMIP3 20C3M and CMIP5 historical simulations.

Several studies have used more sophisticated performance metrics to compare the mean state of multiple fields with available observations (e.g., Gleckler et al., 2008; Reichler and Kim, 2008a; Pincus et al., 2008a; Yokoi et al., 2011). Figure 9.7 (following Gleckler et al. (2008)), depicts the space-time root mean square error (RMSE) for the 1980–2005 climatological annual cycle of the historically forced CMIP3 simulations. For each of the fields examined, this “portrait plot” depicts model performance relative to the median of all model errors, with blue shading indicating a model’s performance being better, and red shading worse, than the median of all model results. In each case, two observational estimates are used to demonstrate the impact of the selection of reference data on the results. The results in this figure are illuminating. Some models consistently compare better with observations than others, some exhibit mixed performance, and some stand out as relatively poor performers. For most fields, the choice of the observational dataset does not substantially change the result for global error measures, indicating that inter-model differences are substantially larger than the differences between the two reference datasets. Nevertheless, it is important to recognize that different data sets often rely on the same source of measurements, and that the results in this figure can have some sensitivity to a variety of factors such as instrument uncertainty, sampling errors (e.g., limited record length of observations), the spatial scale of comparison, the domain considered, and the choice of metric.

Another notable feature of Figure 9.7 is that in most cases the multi-model mean (and median) agree more favourably with the observations than any individual model. This has been long recognized to hold for surface temperature and precipitation (e.g., Lambert and Boer, 2001), but it is now clear that this characteristic of the multi-model mean holds for a broad range of climatological fields and characteristics. Recent work is helping to improve the understanding of why the multi-model mean compares so well with observations (e.g., Annan and Hargreaves, 2011; Pincus et al., 2008b),

[INSERT FIGURE 9.7 HERE]

Figure 9.7: Relative error measures of CMIP5 model performance, based on the global annual cycle climatology (1980–2005) computed from the historical experiments. Rows and columns represent individual variables and models, respectively. The error measure is a global annual cycle space-time root-mean square error (RMSE), which, treating each variable separately, is portrayed as a relative error by normalizing the result by the median error of all model results (Gleckler at al., 2008). For example, a value of 0.20 indicates that a model's RMSE is 20% larger than the typical CMIP5 error for that variable, whereas a value of –0.20 means the error is 20% smaller than the typical error. No colour (white) denotes that data are currently unavailable. A diagonal splits each grid square, showing the relative error with respect to both the default (upper left triangle) and the alternate (lower right triangle) reference data sets. The relative errors are calculated independently for the default and alternate data sets. All reference data used in the diagram are summarized in Table 9.2.

Correlations between the relative error results for different fields in Figure 9.7 are known to exist, reflecting physical relationships in the model formulations (and in the real world) (Gleckler et al., 2008; Yokoi et al., 2011). Cluster analysis methods have recently been used in an attempt to reduce this redundancy (e.g., Yokoi et al., 2011; Nishii et al., 2012). Starting from 43 multivariate RMSE and bias metrics, 7 independent clusters were identified for the CMIP3 simulations (Yokoi et al., 2011). Approaches such as these may to lead to more succinct summaries of model performance.

Some studies have made use of an overall skill score of the mean climate by averaging together the results from multiple metrics such as those in Figure 9.7. Although this averaging process is arbitrary (e.g., with temperature and radiative fluxes having equal influence), it does reduce the chance that a poorer performing model will get the right answer for the wrong reasons. Using a multivariate skill score, Reichler and Kim (2008b) demonstrated how errors were reduced in CMIP3 when compared to earlier model generations. Gleckler et al. (2008) illustrated how an average error of each CMIP3 model was a residual of a rather large spread of variable-specific model performance, suggesting that the average error could mask substantial individual model errors. However, Nishii et al. (2012) demonstrated that different methods of producing a multi-variate skill measure for the CMIP3 models did not substantially alter the conclusions about the better and lesser performing models.

In summary, large scale performance metrics such as those discussed above are a typical first-step toward quantifying model agreement with observations and succinctly summarizing selected aspects of model performance. These metrics are well suited for identifying outliers in various aspects of model performance which, once identified, can be further investigated with more in-depth analysis. Confidence in metrics-based model evaluation is greatest when the metrics are relatively simple, statistically robust, and the results are not strongly dependent upon various analysis choices (Knutti et al., 2010b).

9.4.1.3 Long-Term Global-Scale Changes

The comparison of observed and simulated change is complicated by the fact that the simulation results depend on both model formulation and the time-varying external forcings imposed on the models (Allen et al., 2000; Santer et al., 2007). De-convolving the importance of model and forcing differences (e.g., indirect aerosol effects) in the historical simulations is an important topic that is addressed in Chapter 10.

9.4.1.3.1 Global surface temperature and humidity

Figure 9.8 compares the observational record of 20th century changes in global surface temperature to that simulated by each CMIP5 model. The frequency and magnitude of the interannual variability in most of these simulations is generally similar to that of the observations although there are several exceptions. The gradual warming evident in the observational record, particularly in the more recent decades, is also evident in the simulations, although again there are some important differences among models. The interannual variations in the observations are noticeably larger than the multi-model ensemble because the averaging of individual model results acts to filter much of the variability simulated by the models. On the other hand, the episodic volcanic forcing that is applied to many of the models is evident in the multi-model agreement with the observed cooling particularly after the 1991 Pinatubo eruption. Because the interpretation of differences in model behaviour can be confounded by internal variability and forcing, some studies have attempted to identify and remove dominant factors such as ENSO and the impacts of volcanic eruptions (Fyfe et al., 2010). Efforts such as these can reduce trend uncertainties and thereby improve our ability to evaluate simulated changes with observations. In summary, models broadly capture the observed historical changes in global surface temperature, and in particular the warming of recent decades. Both model formulation and the applied external forcings (see Chapter 10) influence this level of agreement.

[INSERT FIGURE 9.8 HERE]

Figure 9.8: Observed and simulated annual mean global average anomaly time series of surface air temperature. All anomalies are differences from the 1961–1990 time-mean of each individual time series. Top: single simulations currently available for CMIP5 (thin lines); multi-model mean (thick red line); different observations (thick black lines). Vertical dotted brown lines represent times of major volcanic eruptions. Observational data are HadCRUT4 (Morice et al., 2012), GISTEMP (Hansen et al., 2010), and NCDC (Smith et al., 2008b) and are merged surface temperature (2 m height over land and surface temperature over the ocean). Top, inset: the absolute global mean surface temperature for the reference period 1961–1990, for each individual model (colours) and the observations (black, (Jones et al., 1999). Bottom: single simulations from a variety of EMICs (thin lines). Vertical dotted brown bars represent times of major volcanic eruptions. Observational data are the same as for the top panel.

Simulated changes in near surface specific humidity over land have been examined and found to be broadly consistent with observational estimates for the period 1973–1999 (Willett et al., 2010). In the Northern Hemisphere, the extratropical trend in most models is slightly smaller than the observed positive trend, whereas in the tropics the picture is less clear because of substantial inter-model differences, which can at least in part be attributed to large interannual variability. In the extratropics of the Southern Hemisphere there is no significant trend in the observations whereas most of the models have Southern Hemisphere trends similar to their northern counterparts. Given the sparse data network in the Southern Hemisphere this discrepancy may result from a combination of model errors and observational sampling uncertainty.

9.4.1.3.2 Upper tropospheric temperature trends

Most climate model simulations show a larger warming in the tropical troposphere than is found in observational datasets (e.g., (McKitrick et al., 2010) (Santer et al., 2012)). There has been an extensive and sometimes controversial debate in the published literature as to whether the difference between models and observations is statistically significant, once observational uncertainties and natural variability are taken into account (e.g., Douglass et al., 2008; Santer et al., 2008; Christy et al., 2010; McKitrick et al., 2010; Bengtsson and Hodges, 2011; Fu et al., 2011; Santer et al., 2012; Thorne et al., 2011). For the thirty-year period 1979 to 2009 (sometimes updated through 2010 or 2011), the various observational datasets find, in the tropical lower troposphere (LT, see Chapter 2 for definition), an average warming trend ranging from 0.07°C to 0.15°C per decade. In the tropical middle troposphere (MT, see Chapter 2 for definition) the average warming trend ranges from 0.02°C to 0.15°C per decade (e.g., Chapter 2, Figure 2.15; McKitrick et al., 2010). Uncertainty in these trend values arises from different methodological choices made by the groups deriving satellite products (Mears et al., 2011) and radiosonde compilations (Thorne et al., 2011), and from fitting a linear trend to a time series containing substantial interannual and decadal variability (Santer et al., 2008; McKitrick et al., 2010). Although there have been substantial methodological debates about the calculation of trends and their uncertainty, a 95% confidence interval of around ±0.1°C per decade has been obtained consistently for both LT and MT (e.g., Chapter 2; McKitrick et al., 2010). Hence, a trend of zero is, with 95% confidence, consistent with some observational trend estimates but not with others.

For the thirty-year period 1979 to 2009 (sometimes updated through 2010 or 2011), the CMIP3 models simulate a tropical warming trend ranging from 0.1°C to somewhat above 0.4°C per decade for both LT and MT (McKitrick et al., 2010), while the CMIP5 models simulate a tropical warming trend ranging from slightly below 0.15°C to somewhat above 0.4°C per decade for both LT and MT (Santer et al., 2012; see also Po-Chedley and Fu, 2012) who, however, considered the period 1979–2005). Both model ensembles show trends that are higher on average than the observational estimates, although both model ensembles overlap the observational ensemble. Because the differences between the various observational estimates are largely systematic and structural (Chapter 2; Mears et al., 2011), the uncertainty in the observed trends cannot be reduced by averaging the observations as if the differences between the datasets were purely random. Likewise, to properly represent internal variability, the full model ensemble spread must be used in a comparison against the observations, as is well known from ensemble weather forecasting (e.g., Raftery et al., 2005). The very high significance levels of model-observation discrepancies in LT and MT trends that were obtained in some studies (e.g., Douglass et al., 2008; McKitrick et al., 2010) thus arose to a substantial degree from using the standard error of the model ensemble mean as a measure of uncertainty, instead of the standard deviation or some other appropriate measure of ensemble spread. Nevertheless, almost all model ensemble members show a warming trend in both LT and MT larger than observational estimates (McKitrick et al., 2010; Po-Chedley and Fu, 2012; Santer et al., 2012).

It is unclear whether the tropospheric model-trend bias is primarily related to internal atmospheric processes or to coupled ocean-atmosphere processes. The CMIP3 models show a 1979–2010 tropical SST trend of 0.19°C per decade in the multi-model mean, much larger than the various observational trend estimates ranging from 0.10°C to 0.14°C per decade (including the 95% confidence interval, (Fu et al., 2011)). This SST trend bias would cause a trend bias also in TL and TM even if the models' atmospheric components were perfectly realistic. The influence of SST trend errors on the analysis can be reduced by considering changes in tropospheric static stability, measured either by the difference between MT and LT changes or by the amplification of MT changes against LT changes; another approach is to consider the amplification of tropospheric changes against SST changes. For month-to-month variations there is consistency between observations and CMIP3 models concerning amplification aloft against SST variations (Santer et al., 2005), and between observations and CMIP5 models concerning amplication of TM against TL variations (Po-Chedley and Fu, 2012). The 30-year trend in tropical static stability, however, is larger than in the observations for almost all ensemble members in both CMIP3 (Fu et al., 2011) and CMIP5 (Po-Chedley and Fu, 2012). For two CMIP3 models, ECHAM5/MPI-OM and GFDL-CM2.1, this trend bias in static stability lies outside each model's internal variability and is hence highly statistically significant The bias persists even when the models are forced with the observed SST, as was found in the CMIP3 model ECHAM5 (Bengtsson and Hodges, 2011) and the CMIP5 ensemble (Po-Chedley and Fu, 2012).

In summary, there is high confidence (robust evidence although only medium agreement) that most, though not all, CMIP3 and CMIP5 models overestimate the warming trend in the tropical troposphere during the satellite period 1979–2011. The cause of this bias remains elusive.

9.4.1.3.3 Extra-tropical circulation

The AR4 concluded that models, when forced with observed SSTs, are capable of producing the spatial distribution of storm tracks, but generally show deficiencies in the numbers of cyclones and exact locations of the storm tracks. The ability to represent extratropical cyclones in climate models was found to be improving, partly due to increases in the horizontal resolution of the models.

Most studies evaluating extratropical cyclones and storm tracks in climate models to date have looked at individual models, but there are two studies assessing the CMIP5 multi-model ensemble. A regional study by (Zappa, 2012) finds that storm track biases over the North Atlantic have decreased in CMIP5 models compared to CMIP3, although models still produce too zonal a storm track in this region and most models underestimate cyclone intensity. (Chang et al., 2012) also find the storm tracks in the CMIP5 models to be too weak and too equatorwards in their position, similar to the CMIP3 models. Studies based on individual models typically find that models capture the general characteristics of storm tracks and extratropical cyclones (Catto et al., 2010; Ulbrich et al., 2008), and show improvements over earlier model versions (Loptien et al., 2008). However, some models have deficiencies in capturing the location of storm tracks (Catto et al., 2011; Greeves et al., 2007), in part due to problems related to the location of warm waters such as the Gulf Stream and Kuroshio Current (Greeves et al., 2007; Keeley et al., 2012). This is an important issue since future projections of storm tracks are sensitive to changes in SSTs (Catto et al., 2011; Laine et al., 2011; McDonald, 2011). Some studies find that storm track and cyclone biases are more strongly related to atmospheric processes and parameterisations (Bauer et al., 2008a; Boer and Lambert, 2008; Zappa, 2012). Representation of the Mediterranean storm track has been shown to be particularly dependent on model resolution (Bengtsson et al., 2009; Pinto et al., 2006; Raible et al., 2007; Ulbrich et al., 2009), as is the representation of storm intensity and associated extremes (Champion et al., 2011). Most studies have focussed on Northern Hemisphere storm tracks and so there is a lack of information available regarding Southern Hemisphere storm track evaluation.

Westerly jet streams occur in both hemispheres and are associated with the storm tracks and variability in annular modes. Earlier studies noted an increase in the strength of the westerly jets and a poleward shift in their location under increasing greenhouse gases due to the increase in meridional temperature gradient between the rapidly warming upper troposphere in the tropics and cooling lower stratosphere in the extratropics. Indeed, observations suggest that recent decades saw a large poleward shift in the subtropical jets associated with widening of the tropics (Seidel et al., 2008) and latitudinal shifts can occur in response to other forcings (e.g., Simpson et al., 2009). 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) 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).

9.4.1.3.4 Tropical circulation

Trends in tropical atmospheric circulation as estimated from short observational records must be treated with due caution owing to decadal variability. For instance, initial assessments of a weakening Walker circulation (DiNezio et al., 2009; Vecchi and Soden, 2007; Vecchi et al., 2006a) from models and reanalysis products (Yu and Zwiers, 2010) have been tempered by subsequent evidence that tropical Pacific Trade winds may have strengthened since the early 1990s (e.g., Merrifield and Maltrud, 2011). Models suggest that the width of the Hadley cell should increase (Frierson et al., 2007; Lu et al., 2007), and there are indications that this has been observed over the past 25 years (Seidel et al., 2008) but at an apparent rate (2 to 5 degrees of latitude since 1979) that is faster than in the CMIP3 models (Johanson and Fu, 2009).

The tendency in a warming climate for wet areas to receive more precipitation and subtropical dry areas to receive less, often termed the "rich-get richer" mechanism (Chou et al., 2006; Held and Soden, 2006) is simulated in CMIP3 models (Chou and Tu, 2008), and observational support for this is found from ocean salinity observations (Durack et al., 2012) precipitation gauge data over land (Zhang et al., 2007). There is medium confidence that models correctly simulate precipitation increases in wet areas and decreases in dry areas on broad spatial scales in a warming climate based on high agreement among models and limited evidence that this is been detected in observed trends

Several recent studies have examined the co-variability of tropical climate variables as a further means of evaluating climate models. Specifically, there are observed relationships between lower tropospheric temperature and total column precipitable water (Mears et al., 2007), and between surface temperature and relative humidity (Willett et al., 2010). Figure 9.9 (updated from Mears et al., 2007) shows the relationship between 24-year (1988–2011) linear trends in tropical precipitable water and lower tropospheric temperature for individual historical simulations (extended by appending RCP8.5 simulations after 2005, see Santer et al., 2012 for a description of this process) observations, and reanalysis output. As described by Mears et al. (2007) the ratio between changes in these two quantities is fairly tightly constrained in the model simulations and similar across a range of time scales. In the updated figure, the RSS observations are in fairly good agreement with model expectations, and the UAH observations less so. The points associated with two of the reanalyses are far from the line, indicating that these results are inconsistent with model physics. 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).

[INSERT FIGURE 9.9 HERE]

Figure 9.9: Scatter plot of decadal trends in tropical (20ºS to 20ºN) precipitable water as a function of trends in lower tropospheric temperature (TLT) over the world's oceans. Open symbols are from 19 CMIP5 models, filled squares are from satellite observations, and filled triangles are from reanalysis output. Trends are calculated over the 1988-2011 period, so CMIP5 historical runs, which typically end in December 2005, were extended using RCP8.5 simulations initialized using these historical runs. Figure updated from (Mears et al., 2007).

9.4.1.3.5 Ozone and lower stratospheric temperature trends

Ozone has been subject to a major perturbation in the stratosphere since the late 1970s due to anthropogenic emissions of ozone-depleting substances, now successfully controlled under the Montreal Protocol and its Amendments and Adjustments (WMO, 2011). Since the AR4, there is increasing observational and modelling evidence that Antarctic stratospheric ozone loss has contributed to changes in southern high-latitude climate (Thompson et al., 2011). Together with increasing GHG concentrations, the ozone hole has led to a poleward shift and strengthening of the Southern Hemisphere midlatitude tropospheric jet during summer, which has in turn contributed to robust summertime trends in surface winds, observed warming over the Antarctic Peninsula, and cooling over the high plateau (McLandress et al., 2011; Perlwitz et al., 2008; Polvani et al., 2011; Son et al., 2008; Son et al., 2010; SPARC-CCMVal, 2010; Swart and Fyfe, 2012; WMO, 2011). These trends are well captured in chemistry-climate models (CCMs) with interactive stratospheric chemistry and in CMIP3 models with prescribed time-varying ozone (Son et al., 2010; SPARC-CCMVal, 2010). However, around half of the CMIP3 models prescribe ozone as a fixed climatologically, and these models are not able to simulate trends in surface climate correctly (Fogt et al., 2009; Karpechko et al., 2008; Son et al., 2008; Son et al., 2010). To address this, a new ozone dataset (Cionni et al., 2011), based on observations in the past (Randel and Wu, 2007) and CCM projections in the future (SPARC-CCMVal, 2010), was developed and prescribed as zonal mean field in the majority of the CMIP5 models without interactive chemistry. It should be noted that single model studies have shown prescribing zonal mean ozone results in an underestimation of Antarctic temperature trend and related tropospheric circulation changes, as compared to a model with interactive chemistry (Crook et al., 2008; Waugh et al., 2009). Nine of the CMIP5 models include interactive chemistry. Figure 9.10 shows that total column ozone trends for CMIP5 models with prescribed and with interactive chemistry agree well with observations, although some models show substantial deviations (Eyring et al., 2012b).

[INSERT FIGURE 9.10 HERE]

Figure 9.10: Time series of area-weighted total column ozone from 1960 to 2005 for (a) annual mean global mean (90°S–90°N) and (b) Antarctic October mean (60°S–90°S). The multi-model mean and individual CMIP5 models with interactive or semi-offline chemistry (CHEM, red solid and coloured lines) and standard deviation (blue shaded area) are compared to the multi-model mean of the CMIP5 models that prescribe ozone (NOCHEM, green solid line), the IGAC/SPARC ozone database (black dotted line), the CCMVal-2 multi-model mean (yellow solid line) and observations from five different sources (symbols). The observations include ground-based measurements (updated from Fioletov et al. (2002)), NASA TOMS/OMI/SBUV(/2) merged satellite data (Stolarski and Frith, 2006), the NIWA combined total column ozone database (Bodeker et al., 2005), Solar Backscatter Ultraviolet (SBUV, SBUV/2) retrievals (updated from Miller et al. (2002)), and DLR GOME/SCIA/GOME-2 (Loyola and Coldewey-Egbers, 2012; Loyola et al., 2009). Ozone depletion increased after 1960 as equivalent stratospheric chlorine values steadily increased throughout the stratosphere. Adapted from Figure 3 of Eyring et al. (2012b).

Overall, the good representation of past stratospheric ozone changes in the CMIP5 models seen in Figure 9.10 has led to a more realistic representation of the effects of anthropogenic forcings on stratospheric temperatures and subsequent impacts on tropospheric climate (Eyring et al., 2012b) . Lower stratosphere temperature changes since 1958 are characterized by a long-term global cooling trend interrupted by three two-year warming episodes following large volcanic eruptions (Chapter 2, Figure 2.12). During the satellite era (since 1979) the cooling occurred mainly in two step-like transitions in the aftermath of the El Chichón eruption in 1982 and the Pinatubo eruption in 1991, with each cooling transition followed by a period of relatively steady temperatures (Randel et al., 2009; Seidel et al., 2011). This specific evolution of global lower stratosphere temperatures since 1979 is well captured in the CMIP5 models when forced with both natural and anthropogenic climate forcings, although the models tend to underestimate the long-term cooling trend (Charlton-Perez and Coauthors, 2012; Eyring et al., 2012b; Santer and Co-authors, 2012) (see Chapter 10). In general, however, CMIP5 models represent better the overall observed cooling trend compared to CMIP3 models because of the improved representation of stratospheric ozone forcing discussed above. This improvement is independent of the height of the model top (Charlton-Perez and Coauthors, 2012). On a more regional level, Young et al. (2012a) find that CMIP3 and CMIP5 models are in agreement with observations in austral spring over the southern hemisphere polar region, although observational uncertainties are large. Potential causes for biases in lower stratosphere temperature trends are forcing errors, related to prescribed stratospheric aerosol loadings, and stratospheric ozone changes affecting the tropical lower stratosphere (Free and Lanzante, 2009; Santer and Co-authors, 2012; Solomon et al., 2012a). It is found that in CMIP5 models with a model top above the stratopause (high-top models) the variability of lower stratospheric climate is generally well simulated while in low-top models including the CMIP3 model ensemble it is generally underestimated (Charlton-Perez and Coauthors, 2012; Cordero and Forster, 2006).

Tropospheric ozone in the historical period has increased due to increases in ozone precursor emissions from anthropogenic activities. Since the AR4, a new emission dataset has been developed (Lamarque et al., 2010), which has led to some differences in tropospheric ozone burden compared to previous studies, mainly due to biomass burning emissions (Cionni et al., 2011; Lamarque et al., 2010; Young et al., 2012b). Climatological mean tropospheric ozone in the CMIP5 simulations agrees generally well with satellite observations and ozonesondes, although as in the stratosphere, biases exist for individual models (Eyring et al., 2012b; Young et al., 2012b). Details on tropospheric ozone and other radiative forcings agents such as stratospheric water vapour are discussed in Chapter 8.

9.4.1.4 Model Simulations of the Last Glacial Maximum and the Mid-Holocene

Paleoclimate simulations offer a means of evaluating models when confronted with larger forcing changes than those of the 20th century. Some evaluation of millennial trends and variability was provided in Chapter 5; here we focus on simulations of the Last Glacial Maximum (LGM, 21000 years BP) and mid-Holocene (6000 years BP) as two particular benchmark periods. The LGM allows testing of the modelled climate response to the presence of a large ice-sheet in the northern hemisphere and to lower concentration of radiatively active trace gases, whereas the mid-Holocene tests the response to changes in seasonality of insolation in the northern Hemisphere (see Chapter 5). Independent data syntheses over land and ocean, that include estimates of different sources of uncertainties, are available for quantitative model assessment (Braconnot et al., 2012). The CMIP5 protocol included these paleoclimate simulations and so there are results obtained with the same model versions as those used for future climate projections (Taylor et al., 2012), and these can be compared to previous phases of the Paleoclimate Modelling Intercomparison Project (PMIP, (Joussaume and Taylor, 1995), (Braconnot et al., 2007c).

Figure 9.11 shows the most recent update of surface temperature over land (Bartlein et al., 2010a), ocean (Waelbroeck et al., 2009) and ice sheets (Braconnot et al., 2012), and continental precipitation (Bartlein et al., 2010a). Although Figure 9.11b shows that for most models the simulated cooling is within the range of the climate reconstuctions for the LGM, Hargreaves et al (Hargreaves et al., 2011) note a global mean model warm bias over the ocean of about 1°C at the LGM. Figure 9.11b also highlights that models tend to overestimate the tropical cooling and underestimate the mid-latitude cooling, thereby confirming previous conclusions from Kageyama et al. (2006) for the North Atlantic and from Otto-Bliesner et al. (2009) for the tropical oceans. They thus underestimate the polar amplification, which is a feature also found for the mid-Holocene (Masson-Delmotte et al., 2006; Zhang et al., 2010a) and other climatic contexts (Masson-Delmotte et al., 2010). Part of this can be attributed to uncertainties in the representation of sea-ice and vegetation feedbacks that have been shown to amplify the response at the LGM and the MH in these latitudes (Braconnot et al., 2007b; O'ishi and Abe-Ouchi, 2011; Otto et al., 2009). Biases in the representation of the coupling between vegetation and soil moisture are also responsible for excessive continental drying at the LGM (Wohlfahrt et al., 2008) and uncertainties in vegetation feedback in monsoon regions (Dallmeyer et al., 2010; Wang et al., 2008). Nevertheless, the ratio (1.5) between the LGM change in temperature over land and over the ocean (Figure 9.11) is rather similar in different models, resulting mainly from differences in the hydrological cycle over land and ocean (Laine et al., 2009; Sutton et al., 2007), and is consistent with the LGM reconstructions.

At a regional scale, models tend to underestimate the changes in the north-south temperature gradient over Europe both at the LGM (Ramstein et al., 2007) and at the mid-Holocene (Brewer et al., 2007; Davis and Brewer, 2009). In the southern hemisphere, the simulated change in atmospheric circulation is consistent with precipitation records in Patagonia and New Zealand, even though the differences between model results are large and data have large uncertainties (Rojas and Moreno, 2011; Rojas et al., 2009).

An overall assessment of the ability of climate models to reproduce the LGM and MH climates is provided by (Harrison et al., 2012) who considered several criteria (Figure 9.12). Their measures differentiate the magnitude and pattern of the change as well as the way models represent changes in different bioclimatic variables (temperature of the coldest and of the warmest month, growing degree days, moisture index) over land in the different regions. They find that models generally capture large-scale gradients of climate change but show more limited ability to reproduce spatial patterns. Hargreaves et al. (2012) found that current models do not reliably reproduce regional patterns of MH change. Figure 9.12 also shows that more recent models (CMIP5) do not perform better than earlier versions (PMIP2) despite higher resolution and sophistication, suggesting important systematic biases remain that are not connected to representation of vegetation or carbon cycle feedbacks (Braconnot et al., 2007c; Hargreaves et al., 2011). Despite biases common to most models, some consistently reproduce past changes better than others (Figure 9.12).

[INSERT FIGURE 9.11 HERE]

Figure 9.11: Ability of climate models to reproduce surface temperature during the Last Glacial Maximum (LGM, 21 ka BP, top) and precipitation during the mid-Holocene (6 ka BP, bottom), as shown by palaeo-environmental climate reconstructions from pollen and macrofossils over land (Bartlein et al., 2010b) and ice cores in a and c, and from different type of marine records for the 21 ka ocean (Waelbroeck et al., 2009) in a. In a and c, the size of the dots is proportional to the uncertainties at the different sites as provided in the reconstructions. Panel b provides a synthetic view of the ability of climate models to reproduce the relationship between changes in annual-mean temperature over ocean and land in the tropics (red) and changes in annual-mean temperature over the North Atlantic and in Europe (cyan). The squares show the mean value of the reconstructions, with the range shown in black. The empty symbols are the results of individual coupled ocean–atmosphere general circulation models from the second phase of the Paleoclimate Modeling Intercomparison Project (PMIP2) (Braconnot et al., 2007c), and the filled symbols are the results from CMIP5. Panel d shows how model skill in reproducing changes in annual-mean precipitation in different data rich regions has evolved during the different phases of PMIP. Box plots for reconstructions provide estimates of the scatter between different sites, whereas for models they represent the spread of model results. The limits of the boxes are as follows: W Europe (40°N–50°N, 10°W–30°E); North East America (35°N–60°N, 95°W–60°W); North Africa (10°N–25°N, 20°W–30°) and E Asia (25°N–40°N, 75°E–105°E). Adapted from Braconnot et al. (2012).

[INSERT FIGURE 9.12 HERE]

Figure 9.12: Summary of benchmark metrics for the Last Glacial Maximum (LGM, ca 21,000 yr BP) and the mid-Holocene (MH, ca 6000 yr BP). For each variable, six different metrics are considered. The median and the inter-quartile range (IQR) characterize the global distribution of the values considering only grid cells where there are observations. The agreement between the simulated and reconstruction maps is then characterised by the difference between the median values (median bias) and the ratio of the simulated and reconstructed IQR (IQR ratio). Other metrics consider the Euclidian distance between the simulated and reconstructed maps and a variant based on fuzzy logic that takes into account an estimate of simulated and reconstructed field uncertainties. The Kendall rank correlation (1-Tau) measures the similarities or differences in the spatial patterns without regard for magnitude and is used as an alternative to the Normalized Mean Square Error (NMSE) based on standardised variables. In this graph all the values have been normalised following Gleckler et al. (2008), so as to highlight model spread as in Figure 9.7. For the mid-Holocene we only plotted the values for all the CMIP5 simulations and consider only the PMIP2 results as ensembles to show how the performances of the model evolved. For the LGM there are fewer simulations and results of both PMIP2 and CMIP5 simulations are included. Adapted from Harrison et al. (2012).

9.4.1.5 Summary

From a global perspective, there is high confidence that large-scale patterns of surface temperature are well simulated by the CMIP5 models. In certain regions this agreement is limited, particularly at elevations over the Himalayas and parts of both Greenland and Antarctica. There is also high confidence that the broad-scale features of precipitation as simulated by the CMIP5 models are in modest agreement with observations, with the overall quality assessment being influenced by systematic errors in the tropics and other notable deficiencies. There remain significant errors in the model simulation of clouds (and their radiative effects) in CMIP5, and it is very likely that these errors contribute to the uncertainties in cloud feedbacks. Gauging model quality with performance metrics, there is robust evidence that some models generally simulate the mean state more realistically than others, and limited evidence that the CMIP5 models perform better than their CMIP3 counterparts.

The CMIP5 models broadly capture the observed historical changes in global surface temperature, and in particular the warming of recent decades. Both model formulation and the applied external forcings influence this medium level of agreement. There is robust evidence that most, though not all, CMIP3 and CMIP5 models overestimate the warming trend in the tropical troposphere during the satellite period 1979–2011. The cause of this bias has remained elusive.

There is high confidence that the trends in stratospheric ozone, whether prescribed or calculated interactively, are generally in good agreement with observations, although deviations for individual models exist. There is robust evidence that this constitutes a significant improvement over CMIP3, where all models prescribed stratospheric ozone and half of them used a fixed ozone climatology. Correspondingly, there is high confidence that the representation of associated impacts on high latitude surface climate and lower stratospheric cooling trends has improved compared to CMIP3.

The evaluation of temperature and humidity indices in simulations of the mid-Holocene shows robust evidence that models generally capture large-scale gradients of climate change but show medium to poor performance when considering patterns and magnitude at a regional scale. More recent models (CMIP5) do not perform better than earlier versions (PMIP2) despite higher resolution and sophistication.

9.4.2 Ocean

Accurate simulation of the ocean in climate models is essential for the correct estimation of transient ocean heat uptake and transient climate response, ocean CO_2 uptake, sea level rise, and coupled climate modes such as ENSO. In this Section we focus on the evaluation of model performance in simulating the mean state of ocean properties, surface fluxes and their impact on the simulation of ocean heat content and sea level, and tropical features of importance for climate variability. Simulations of both the recent (20th century mean and evolution) and more distant past are evaluated against available data. Ocean reanalysis are not used for model evaluation as many of their properties depend on the model used to build the reanalysis (this is especially true near the equator).

9.4.2.1 Simulation of Mean Temperature and Salinity Structure

Potential temperature and salinity are the main ocean state variables and strongly interact with the ocean circulation and the surface forcing of heat, fresh water and momentum. Their zonal distribution offers an initial evaluation of the performance of climate models in simulating the different regions of the ocean (upper ocean, thermocline, deep ocean). Over most latitudes, at depths ranging from 200 m to 2000 m, the CMIP5 multi-model mean zonally averaged ocean temperature is too warm (Figure 9.13a), albeit with a cooler deep ocean. Similar biases were evident in the CMIP3 multi-model mean. Above 200 m, however, the CMIP5 (and CMIP3) multi-model mean is too cold, with maximum cold bias (more than 1°C) near the surface at mid-latitudes of the NH and near 200 m at 15°S. The zonal salinity errors (Figure 9.13b) exhibit a different pattern from those of the potential temperature indicating that most do not occur via density compensation. Some near surface structures in the tropics and in the northern mid-latitude are indicative of density compensation and are most likely due to surface fluxes errors. At intermediate depths, errors in water mass formation translate into errors in both salinity and potential temperature [see below].

[INSERT FIGURE 9.13 HERE]

Figure 9.13: Time-mean differences between CMIP5 multi-model ensemble-mean and observed (A) potential temperature (°C) and (B) salinity (PSS-78) (colour). The observed climatological values (Antonov et al., 2010; Levitus et al., 2009) are zonally averaged for the global ocean (excluding marginal and regional seas) and are shown as labelled black contours. The simulations cover the period 1975 to 2005 from available historical simulations, whereas the observations are from 1874 to 2008. Multiple realizations from individual models are first averaged to form a single model climatology, before the construction of the multi-model ensemble mean. 21 available CMIP5 models have contributed to the temperature panel (A) and 20 models to the salinity panel (B).

AR4 noted that the largest errors in the simulation of sea surface temperature (SST) in CMIP3 were found in mid and high latitudes. While this is still the case in CMIP5 (Figure 9.14), there is marginal improvement with fewer individuals models exhibiting serious bias. The sea surface salinity (SSS) is more challenging to observe, even though the last decade has seen substantial improvements in the development of global salinity observations, such as those from the ARGO network (see Chapter 3). Whereas SST is strongly constrained by air-sea interactions, the sources of SSS variations (surface forcing via evaporation minus precipitation, sea-ice formation/melt and river runoff) are only loosely related to the SSS itself, allowing errors to develop unchecked in coupled models. An analysis of twelve CMIP3 models that did not use flux adjustments showed that the near-global (60°N–60°S) mean SSS bias across the models was between –0.8 and +0.3 psu (Waliser et al., 2011a), while regional SSS biases are as high as ±2.5 psu (Terray et al., 2012). Comparisons of modeled versus observed estimates of evaporation minus precipitation suggest that model biases in surface freshwater flux play a role in some regions (e.g., double ITCZ in the East Pacific, (Lin, 2007)).

[INSERT FIGURE 9.14 HERE]

Figure 9.14: Annual mean, zonally averaged SST error, simulated minus observed climatology in (a) CMIP5 and (b) CMIP3. The Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) (Rayner et al., 2003a) observational climatology for 1850 to 2010 is the reference used here, and the model results are for the same period in the historical simulations. The multi-model mean of CMIP5 and CMIP3 are shown in both panels (bold black and blue respectively). [PLACEHOLDER FOR FINAL DRAFT: legend that identifies the individual models to be included.]

Detailed assessments of the performance of coupled climate models in simulating hydrographic structure and variability are still relatively sparse. Two important regions, the Labrador and Irminger Seas and the Southern Ocean, have been investigated to some extent (de Jong et al., 2009) and (Sloyan and Kamenkovich, 2007). Eight CMIP3 models produce simulations of the intermediate and deep layers in the Labrador and Irminger Seas that are generally too warm and saline, with biases up to 0.7 psu and 2.9°C. The biases arise because the convective regime is restricted to the upper 500 m; thus, intermediate water that in reality is formed by convection is, in the models, partly replaced by warmer water from the south. In the Southern Ocean, Subantarctic Mode Water (SAMW) and Antarctic Intermediate Water (AAIW), two water masses indicating very efficient ocean ventilation, are found to be well simulated in some models but not in others (Sloyan and Kamenkovich, 2007). (McClean and Carman, 2011) found biases in the properties of the North Atlantic mode waters and their formation rates in the CMIP3 models. Errors in Subtropical Mode Water (STMW) formation rate and volume produce a turnover time of 1–2 years, approximately half of that observed (Figure 9.15), implying errors in in ocean heat input and hence, storage.

Few studies have assessed the performance of models in simulating Mixed Layer Depth (MLD). In the North East Pacific region, Jang et al. (2011) found that the CMIP3 models exhibit the observed deep MLD in the Kuroshio Extension, though with a deep bias and only one large deep MLD region, rather than the observed two localized maxima. Other studies have noted MLD biases near sea-ice edges (Capotondi et al., 2012).

[INSERT FIGURE 9.15 HERE]

Figure 9.15: Sub-Tropical Mode Water (STMW) turnover time for various models compared with (Kwon and Riser, 2004); time is calculated by annual maximum volume divided by annual production. Values are means; error bars give ranges of one standard deviation. Square data symbols indicate those models with a distinct (if small) secondary water mass transformation rate peak corresponding to STMW formation. Triangular data symbols indicate those models with broad, diffuse, or indiscernible STMW formation peak (from McClean and Carman, 2011).

9.4.2.2 Simulation of Sea Level and Ocean Heat Content

Steric and dynamical components of the sea surface height (SSH) pattern are simulated by the current generation of climate models and can be evaluated with high quality near-global satellite altimetry measurements (Ducet et al., 2000). Performance metrics have been used to evaluate the time-mean spatial distribution of SSH in the CMIP3 simulations (Yin et al., 2010) and the annual cycle and interannual variability in CMIP5 (Landerer et al., submitted). A Taylor-diagram (Figure 9.16) of the climatological annual cycle of SSH in CMIP3 and CMIP5 models shows that the correlations with observations are relatively low (0.5 and 0.8) compared to some well-observed atmospheric quantities (cf. Figure 9.6), likely because of the combination of processes that determine SSH (surface wind stress and heating, ocean dynamics, etc.). Many models do however have a spatio-temporal variability (standard deviation) that agrees fairly well with the observations. A few models perform less well; these models also have a larger mean RMSE for the time-mean dynamic topography (Yin et al., 2010). The familiar result of the multi-model ensemble mean outperforming any individual model in atmospheric measures (Section 9.4.1.2) is also evident in Figure 9.16. While performance differences between the CMIP3 and CMIP5 models are limited for the annual cycle, the representation of the time-mean SSH pattern has markedly improved in CMIP5 over CMIP3 (Landerer et al., submitted). Improved simulation of SSH has also been demonstrated in high resolution eddy resolving ocean models when compared to coarser resolution versions (McClean et al., 2006a). Chapter 13 provides a more extensive assessment of sea level changes in the CMIP3 and CMIP5 simulations including comparisons with century-scale historical records.

[INSERT FIGURE 9.16 HERE]

Figure 9.16: Taylor diagram of the dynamic sea-level height seasonal cycle climatology (1987–2000). The radial coordinate shows the standard deviation of the spatial pattern, normalised by the observed standard deviation. The azimuthal variable shows the correlation of the modelled spatial pattern with the observed spatial pattern. The root-

mean square error is indicated by the dashed grey circles about the observational point. Analysis is for the global ocean, 50°S–50°N. The reference dataset is AVISO, a merged satellite product (Ducet et al., 2000), which is described in Chapter 3. Figure currently shows results for the CMIP3 models and the CMIP5 data currently available.

Ocean heat content (OHC) depends only on ocean temperature, whereas absolute changes in sea level are also influenced by processes that are only now being incorporated into global models (e.g., mass loss from large ice sheets). It is worth noting, however, that global scale changes in OHC are highly correlated with the thermosteric contribution to global SSH changes (Domingues et al., 2008). Figure 9.17 shows observed and simulated global 0–700 m OHC changes during the overlap period of the observational record and the CMIP5 historical experiment (1960–2005). The differences between the three observational estimates suggest that substantial uncertainties remain, however, there is no statistical difference in their 1960-2005 trends (cf., Gleckler et al., 2012b). Approximately half of the historical CMIP3 simulations included the effects of volcanic eruptions. Without these eruptions, the models take up too much heat during the late 20th century (Domingues et al., 2008; Gleckler et al., 2006). In the CMIP5 models shown in Figure 9.17, all but the IPSL-CM5A-LR and CSIRO-MK3-6-0 models include volcanic forcings (Santer et al., 2012). In idealized CMIP5 experiments (CO_2 increasing 1% yr^{-1}), the heat uptake efficiency of the CMIP5 models varies by a factor of two, explaining about 50% of the model spread (Till Kuhlbrodt and Gregory, 2012). Despite observational uncertainties, this recent work also provides limited evidence that in the upper 2000 m, most CMIP5 models are less stratified (in the global mean) than is observed, which is suggestive that these models transport heat downwards more efficiently than the real ocean. These results are consistent with earlier studies (Boe et al., 2009b; Forest et al., 2006, 2008; Sokolov et al., 2010) that concluded that CMIP3 models may overestimate oceanic mixing efficiency and therefore underestimate TCR and its impact on future surface warming. Eby et al. (submitted) find that most EMICs also overestimate OHC changes for the upper 2000 m which has implications for temperature effects on the ocean carbon-cycle.

The interannual variability of simulated OHC by the CMIP3 models agrees better with observations when the model data is sampled using the observational data mask (AchutaRao et al., 2007). Independent evidence suggests that the differences between the inter-annual changes in observed global net TOA heating and OHC are not statistically significant (when accounting for uncertainties) from which it can be inferred that CMIP3 simulated OHC may sometimes be too large (Loeb et al., 2012). On decadal time scales, there is limited evidence that basin scale space-time variability structure of CMIP3 models is approximately 25% lower than the (poorly constrained) observations (Gleckler et al., 2012a).

[INSERT FIGURE 9.17 HERE]
Figure 9.17: Time series of observed and simulated (CMIP5 historical) global 0–700 m ocean heat content anomalies (with respect to 1957–1990 climatologies), with units of 10^{22} Joules. The three observational estimates (thick lines) are discussed in Chapter 3. Individual simulations (one per model) are shown for the historical period. [PLACEHOLDER FOR FINAL DRAFT: to be updated; model simulations will be corrected for simulation drift, which for some models may be significant (Sen Gupta et al., 2012).]

9.4.2.3 Simulation of Circulation Features Important for Climate Response

9.4.2.3.1 Simulation of recent ocean circulation

Atlantic Meridional Overturning Circulation

The Atlantic Meridional Overturning Circulation (AMOC) consists of northward transport of shallow warm water overlying a southward transport of deep cold water and is responsible for a considerable part of the northward oceanic heat transport. In the AR4, models showed considerable spread in the time-mean strength of the AMOC, though the majority of models showed an AMOC strength that is within observational uncertainty. Long-term AMOC estimates have had to be inferred from hydrographic measurements sporadically available over the last decades (e.g., Bryden et al., 2005; Lumpkin et al., 2008). At 26°N, these indicate a time-mean value of about 18 Sv with an observational uncertainty of ±6 Sv. Continuous AMOC monitoring at 26°N was started in 2004 (Cunningham et al., 2007) and provide a four-year mean value of 18.7 Sv with an error of ±2.1 Sv (Kanzow et al., 2010). The ability of models to simulate this important circulation feature is tied to the credibility of simulated AMOC weakening during the 21st century because the strength of the weakening is correlated with the initial AMOC strength (Gregory et al., 2005). The CMIP5 mean AMOC strength ranges from 15 to 30 Sv for the historical period which is comparable to that of the CMIP3 models, and so suggests no notable improvement (see Figure 12.37 in Chapter 12).

Western boundary currents

The relatively low horizontal resolution of the ocean component of current AOGCMs leads to mis-located western boundary currents that are too weak and diffuse. This contributes to biases in heat transport, SST, SSS and subtropical mode water formation (Kwon et al., 2010). In the Southern Hemisphere, Gupta et al. (2009) found considerable spread in the ability of CMIP3 AOGCMs to represent both the meridional variation in the transports of the Agulhas, Brazil and East Australian Currents as well as in the latitude of maximum transport.

Southern Ocean circulation

The Southern Ocean is an important driver for the meridional overturning circulation and is closely linked to the zonally continuous Antarctic Circumpolar Current (ACC). The ACC has a typical transport through the Drake Passage of about 135 Sv (e.g., Cunningham et al., 2003). The ability of CMIP3 models to adequately represent Southern Ocean circulation and water masses seems to be affected by several factors (Russell et al., 2006). The most important appear to be the strength of the westerlies at the latitude of the Drake Passage, the heat flux gradient over this region, and the change in salinity with depth across the ACC. Gupta et al. (2009) noted that relatively small deficiencies in the position of the ACC lead to more obvious biases in the SST in the models. The models show considerable variations in the strength of the circumpolar circulation. A comparison of CMIP5 models (Meijers et al., 2012) shows that, firstly, the ACC transport through Drake Passage is improved as compared to the CMIP3 models, and secondly, that the intermodal range in the zonal mean ACC position is smaller than in the CMIP3 case (in CMIP5, the mean transport is 148 Sv and the standard deviation is 50 Sv across an ensemble of 21 models).

9.4.2.3.2 Simulation of glacial ocean conditions

Reconstructions of the last glacial maximum from sediment cores indicate that the regions of deep water formation in the North Atlantic were shifted southward, that the boundary between North Atlantic Deep Water (NADW) and Antarctic Bottom Water (AABW) was substantially shallower than today, and that NADW formation was less intense (Curry and Oppo, 2005; Dokken and Jansen, 1999; Duplessy et al., 1988; McManus et al., 2004). This signal, although estimated from a limited number of sites, is robust (see chapter 5). The AR4 reported that model simulations showed a wide range of AMOC response to LGM forcing (Weber et al., 2007), with some models reducing the strength of the AMOC and its extension at depth and other showing no change or an increase. Figure 9.18 provides an update of the diagnosis proposed by (Otto-Bliesner et al., 2007a) to compared model results with the deep ocean data from (Adkins et al., 2002) using PMIP2 and CMIP5 pre-industrial and LGM simulations (Braconnot et al., 2012). These models reproduce relatively well the modern deep ocean temperature-salinity (T-S) structure in the Atlantic basin, but exhibit considerable spread for the LGM simulations, suggesting that processes responsible for such paleoclimate changes may not be well reproduced in contemporary climate models, and this must be kept in mind when viewing projected changes in deep ocean properties.

[INSERT FIGURE 9.18 HERE]

Figure 9.18: Temperature and salinity for the modern period (open symbols) and the Last Glacial Maximum (LGM, filled symbols) as estimated from data (black symbols) at ODP sites (Adkins et al., 2002) and predicted by the PMIP2 (small triangles) and PMIP3/CMIP5 (big triangles) simulations. Site 981 (triangles) is located in the North Atlantic (Feni Drift, 55'N, 15'W, 2184 m). Site 1093 (upside-down triangles) is located in the South Atlantic (Shona Rise, 50'S, 6'E, 3626 m). In PMIP2 only CCSM included a 1 psu adjustment of ocean salinity at initialization to account for freshwater frozen into LGM ice sheets; the other PMIP2 model-predicted salinities have been adjusted to allow a comparison. In PMIP3 all simulations include the 1 psu adjustment as required in the PMIP2/CMIP5 protocol (Braconnot et al., 2012), T he dotted lines allow a comparison of the values at the northern and the southern sites for a same model. This figure is adapted from Otto-Bliesner et al. (2007b) and shows quantitatively how deep-ocean properties can be evaluated for both modern and palaeo climates. In particular, models have difficulties to reproduce the cold and salty water found at depth at the LGM.

9.4.2.4 Simulation of Surface Fluxes and Meridional Transports

Surface fluxes play a large part in determining the fidelity of ocean simulations. As noted in the AR4, large uncertainties in surface heat and fresh water fluxes observations (usually obtained indirectly) do not allow useful validation of models. This is still the case and so we focus here on an integrated quantity, meridional heat transport, which is less prone to errors. Surface wind stress is better observed and models are evaluated against observed products below.

9.4.2.4.1 Surface wind stress

The main ocean surface currents are wind-driven, and the zonal component of wind stress is particularly important. The annually averaged zonal mean zonal wind stress is reasonably well simulated by the CMIP3 models and also for CMIP5 (Figure 9.19). At most latitudes, the reanalysis estimates lie within the range of model results. At middle to high latitudes, the model-simulated wind stress maximum tends to lie slightly equatorward of that in the reanalysis. This equatorward shift in the southern ocean is slightly reduced in CMIP5. At these latitudes, the largest near surface wind speed biases in CMIP5 are located over the Pacific sector and the smallest are in the Atlantic sector (Bracegirdle et al., 2012). Such wind stress errors may adversely affect oceanic heat and carbon uptake (Swart and Fyfe, 2011). At middle to low latitudes, the CMIP3 and CMIP5 model spreads are relatively small and all the model results lie fairly close to the reanalysis, although near the equator this can occur through compensated zonal errors (Figure 9.20).

9.4.2.4.2 Meridional heat transport

In steady state, the surface heat flux balances the convergence of ocean heat transport, which is therefore a convenient quantity for evaluation. Recent papers by Probst et al. (2012) and Donohoe and Battisti (2012) suggest that the main source of discrepancy in the total meridional transport simulated by the CMIP3 models is the spread in cloud reflection properties, with surface albedo differences playing a minor role. While the atmosphere is the dominant contributor to the total meridional transport (Trenberth and Caron, 2001), biases in the ocean transport have to be compensated for by the modelled atmosphere, and can therefore lead to discrepancies in the coupled system.

The models simulations qualitatively agree with the various observational estimates on the most important features of ocean heat transport (Figure 9.21) and, in a multi-model sense, no major change from CMIP3 can be seen. All CMIP5 models are able to the represent the strong asymmetry with respect to the equator, with the largest values in the Northern Hemisphere, consistent with the observational estimates. At most latitudes the majority of CMIP5 model results fall within the range of observational estimates, although there is some suggestion of modest underestimate between 15°N and 25°N and south of about 60°S. Some models show an equatorward transport at southern hemisphere midlatitudes that is also featured in the observation estimate of Large and Yeager (2009). This highlights the difficulties in the representing large-scale energy processes in the Southern ocean as discussed by Trenberth and Fasullo (2010b).

[INSERT FIGURE 9.19 HERE]

Figure 9.19: Zonal mean zonal wind stress over the oceans in CMIP3 and CMIP5 simulations during 1970–1999. Individual simulations: thin coloured; ensemble mean: thick red. The black solid, dashed, and dash-dotted curves represent QuikSCAT satellite measurements (Risien and Chelton, 2008), NCEP/NCAR reanalysis I (Kalnay et al., 1996), and ERA-Interim (Dee et al., 2011), respectively. The figure shows that CMIP3 & CMIP5 mid-latitude westerly winds are too strong and that circumpolar westerly wind are too weak compared to QuikSCAT, NCEP1, and ERA-Interim. CMIP3 and CMIP5 ensemble means are very similar.

[INSERT FIGURE 9.20 HERE]

Figure 9.20: Equatorial (2ºS–2ºN averaged) zonal wind stress for the Indian, Pacific, and Atlantic Oceans; comparison of simulations by CMIP3 (a) and CMIP5 (b) with QuikSCAT observations. Individual simulations: thin coloured; ensemble mean: thick red. The black solid, dashed, and dash-dotted curves represent QuikSCAT satellite measurements (Risien and Chelton, 2008), NCEP/NCAR reanalysis I (Kalnay et al., 1996), and ERA-Interim (Dee et al., 2011), respectively. The figure shows that in both CMIP3 and CMIP5, the ensemble-mean equatorial zonal wind stress is too weak in the Atlantic and Indian Oceans and too strong in the western Pacific. CMIP3 and CMIP5 ensemble means are very similar.

[INSERT FIGURE 9.21 HERE]

Figure 9.21: Annual mean, zonally averaged oceanic heat transport implied by net heat flux imbalances at the sea surface for CMIP5 simulations, under an assumption of negligible changes in oceanic heat content. Observational estimates include: (blue and green dashed lines) the dataset from Trenberth and Caron (2001) for the period February 1985 to April 1989, derived from reanalysis products from the National Centers for Environmental Prediction (NCEP/NCAR, Kalnay et al., 1996) and European Centre for Medium Range Weather Forecasts 40-year reanalysis (ERA40, Uppala et al., 2005), (purple dashed line) an updated version by Trenberth and Fasullo (2008) with improved TOA radiation data from the Clouds and Earth's Radiant Energy System (CERES) for March 2000 to May 2004, and updated NCEP reanalysis (Kistler et al., 2001) up to 2006, (red dashed line) the Large and Yeager (2009) analysis based on the range of annual mean transport estimated over the years 1984–2006, computed from air-sea surface fluxes

adjusted to agree in the mean with a variety of satellite and in situ measurements, and (black diamonds) direct estimates by Ganachaud and Wunsch (2003) obtained from hydrographic sections during the World Ocean Circulation Experiment combined with inverse models. The model climatologies are derived from the years 1986 to 2005 in the historical simulations in CMIP5. The multi-model mean is shown as a thick black line. The CMIP3 multi-model mean is added as a cyan blue line. Note: climate models should feature a vanishing net energy balance when long time averages are considered but unphysical sources and sinks lead to energy biases (Trenberth and Fasullo, 2009; Trenberth and Fasullo, 2010a; Lucarini and Ragone, 2011) that are also found in reanalysis constrained by observations (Trenberth et al., 2009). When correcting for the imperfect closure of the energy cycle, as done here, comparison between models and observational estimates become possible.

9.4.2.5 Simulation of Tropical Mean State

9.4.2.5.1 Tropical Pacific Ocean

From CMIP1 through CMIP5, models have shown persistent biases in important properties of the mean state of the tropical Pacific (AchutaRao and Sperber, 2002; Guilyardi et al., 2009b; Randall et al., 2007). Among these are the mean thermocline depth and slope along the equator, the structure of the equatorial current system, and the equatorial cold tongue (Brown et al., 2010a; Reichler and Kim, 2008a). Many of the reasons for these biases have, in principle, been identified, such as: too strong trade winds; a too diffusive thermocline; deficient horizontally isotropic mixing coefficients; insufficient penetration of solar radiation; and too weak tropical instability waves (Lin, 2007; Meehl et al., 2001; Wittenberg et al., 2006). Because of strong interactions between the processes involved, it is difficult to identify the ultimate source of these errors, although new approaches using the initial adjustment of seasonal hindcasts suggest that the equatorial wind stress may be at the origin of several errors (Vannière et al., 2012). A heat budget analysis in the CMIP3 models indicates that errors in both net surface heat flux and total upper ocean heat advection significantly contribute to the excessive cold tongue in the equatorial Pacific (Zheng et al., 2012). In that regard it is noteworthy that CMIP5 models exhibit modest improvements in the western equatorial Pacific when compared to CMIP3, with reduced trade wind errors (Figure 9.20). Studies using AOGCMs with eddy permitting ocean resolution report improvements in aspects such as tropical instability waves and coastal upwelling (Jochum and Murtugudde, 2006; Roberts et al., 2009; Sakamoto et al., 2011; Shaffrey et al., 2009).

A particular problem in simulating the seasonal cycle in the tropical Pacific arises from the "double Intertropical Convergence Zone (ITCZ)", defined as the appearance of a spurious ITCZ in the Southern Hemisphere associated with excessive tropical precipitation. Further problems are too strong a seasonal cycle in simulated SST and winds in the eastern Pacific and the appearance of a spurious semi-annual cycle. The latter has been attributed to meridional asymmetry in the background state that is too weak, possibly in conjunction with incorrect regional water vapour feedbacks (De Szoeke and Xie, 2008; Guilyardi, 2006; Li and Philander, 1996; Timmermann et al., 2007; Wu et al., 2008b).

A further persistent problem in AOGCMs is insufficiet marine stratocumulus cloud in the eastern tropical Pacific, caused presumably by weak coastal upwelling off South America leading to a warm SST bias (Lin, 2007). Although the problem persists, improvements are being made (AchutaRao and Sperber, 2006; Reichler and Kim, 2008a).

The equatorial undercurrent (EUC) is a major component of the tropical Pacific Ocean circulation. Even though EUC velocity in most CMIP3 models is sluggish relative to observations, it does not appear to impair other major components of the tropical circulation including upwelling and poleward transport (Karnauskas et al., 2012). These latter transports play a critical role in a theory for how the tropical Pacific may change under increased radiative forcing, i.e., the ocean dynamical thermostat mechanism (Clement et al., 1996; Seager and Murtugudde, 1997; An et al., 2011). These findings suggest that, in the mean, global climate models may not under-represent the role of equatorial ocean circulation, nor bias the balance between competing mechanisms governing the response of tropical Pacific to climate change.

9.4.2.5.2 Tropical Atlantic Ocean

CMIP3 models exhibit severe biases in the tropical Atlantic Ocean, so severe that some of the most fundamental features—the east-west SST gradient and the eastward shoaling thermocline along the equator—cannot be reproduced (e.g., Chang et al., 2008; Chang et al., 2007; Richter and Xie, 2008). In many models, the warm SST bias along the Benguela coast is in excess of 5°C and the Atlantic warm pool in the western basin is grossly underestimated. As in the Pacific, CMIP3 models suffer the double ITCZ syndrome in the Atlantic, with a southern ITCZ that is not observed. Hypotheses for the complex Atlantic bias problem tend to draw on the fact that the Atlantic Ocean has a far smaller basin, and thus encourages a tighter and more complex land-atmosphere-ocean interaction. A recent study using a high-resolution coupled model suggests that the warm eastern equatorial Atlantic SST bias is more sensitive to the local rather than basin-wide trade wind bias and to a wet Congo basin instead of a dry Amazon—a finding that differs from previous studies (Patricola et al., 2011). Recent ocean model studies show that a warm subsurface temperature bias in the eastern equatorial Atlantic is common to virtually all ocean models forced with "best estimated" surface momentum and heat fluxes, owing to problems in parameterization of vertical mixing (Hazeleger and Haarsma, 2005).

9.4.2.5.3 Tropical Indian Ocean

CMIP3 models simulate equatorial Indian Ocean climate reasonably well, though most models produce weak westerly winds and a flat thermocline on the equator. The models show a large spread in the modelled depth of the 20°C isotherm in the eastern equatorial Indian Ocean (Saji et al., 2006). The reasons are unclear but may be related to differences in the various model parameterisations of vertical mixing (Schott et al., 2009).

CMIP3 models generally simulate the Seychelles Chagos thermocline ridge in the Southwest Indian Ocean, a feature important for the Indian monsoon and tropical cyclone activity in this basin (Xie et al., 2002). The models, however, have significant problems in accurately representing its seasonal cycle because of the difficulty in capturing the asymmetric nature of the monsoonal winds over the basin, resulting in too weak a semi-annual harmonic in the local Ekman pumping over the ridge region compared to observations (Yokoi et al., 2009).

9.4.2.6 Summary

It is likely (robust evidence and medium agreement) that the ocean component of CMIP3 and CMIP5 models simulate the physical and dynamical processes at play during transient ocean heat uptake, ocean CO_2 uptake, sea level rise, and coupled modes of variability. There is little evidence that CMIP5 models differ significantly from CMIP3, although there is some evidence of modest improvement. Many improvements are seen in individual CMIP5 ocean components (some now including interactive ocean biogeochemistry) and the number of relatively poor-performing models has been reduced (thereby reducing intermodal spread). New since the AR4, process-based model evaluation is now helping identify the cause of some specific biases.

There is robust evidence and high agreement that SST is well simulated, with limited progress since AR4 and still significant regional biases (e.g., North Atlantic). There robust evidence and high agreement that the AMOC is simulated with mixed skill with limited improvements from CMIP3 to CMIP5, limited evidence and medium agreement that western boundary currents are simulated with mixed skill, medium evidence and high agreement that the meridional heat transport is simulated with high skill and robust evidence and high agreement that zonal mean zonal wind stress is simulated with high skill. The tropical Pacific mean state is simulated with mixed skill, although the cold tongue erroneous extension in the west Pacific is reduced in CMIP5, medium evidence and high agreement that the tropical Atlantic mean state is not correctly simulated, and limited evidence and medium agreement that the tropical Indian Ocean is well simulated. There is high confidence that, with medium quality, the CMIP3 models (which include the effects of volcanic eruptions) simulate the global scale characteristics of observed increases in ocean heat content.

9.4.3 Sea Ice

Evaluation of sea-ice performance requires accurate information on ice concentration, thickness, velocity, salinity, snow cover and other factors. The most reliably measured characteristic of sea ice remains sea ice extent (usually understood as the area covered by ice with a concentration above 15%). Caveats, however, exist related to the uneven reliability of different sources of sea ice extent estimates (e.g., satellite vs. pre-satellite observations) and the accuracy of corresponding retrieval algorithms (see Chapter 4), as well as to limitations of this characteristic as a metric of model performance (Notz, 2012).

Despite substantial differences in performance between individual models, the CMIP3 and CMIP5 multi-model mean annual cycles of sea ice extent in both hemispheres agree reasonably well with observations (Figure 9.22). The CMIP5 multi-model ensemble exhibits modest improvements in simulation of the extent in the both hemispheres, although there has been no substantial increase in the sophistication of sea-ice treatment. In many models the regional distribution of sea ice concentration is poorly simulated, even if the hemispheric extent is approximately correct. In Figure 9.23, however, one can see that the median ice edge position (indicated by the color at which half of the models have ice of 15% concentration) agrees reasonably well with observations in both hemispheres (except austral summer in Antarctica), as was the case for the CMIP3 models shown in the AR4.

[INSERT FIGURE 9.22 HERE]

Figure 9.22: Mean sea ice extent (the ocean area with a sea ice concentration of at least 15%) seasonal cycle in the Northern (upper panel) and Southern (lower panel) Hemispheres as simulated by 40 CMIP5 and 17 CMIP3 models. The observed sea-ice extent cycles (1980–1999) are based on the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST; Rayner et al., 2003a) and the National Snow and Ice Data Center (NSIDC; Fetterer et al., 2002) data sets. The shaded areas show the inter-model standard deviation for each ensemble. Adapted from Pavlova et al. (2011).

[INSERT FIGURE 9.23 HERE]

Figure 9.23: Sea ice distribution in the Northern Hemisphere (upper panels) and the Southern Hemisphere (lower panels) for March (left) and September (right). A) AR5 baseline climate (1986–2005) simulated by 40 of CMIP5 AOGCMs. For each 1° × 1° longitude-latitude grid cell, the figure indicates the number of models that simulate at least 15% of the area covered by sea ice. B) AR4 baseline climate (1980–1999) differences between 40 CMIP5 and 17 CMIP3 (AR4 (Randall et al., 2007) Figure 8.10) AOGCMs. For each 2.5° × 2.5° longitude-latitude grid cell, the figure indicates the difference in the normalized number of CMIP5 minus CMIP3 models that simulate at least 15% of the area covered by sea ice. The observed 15% concentration boundaries (red line) are based on the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) data set (Rayner et al., 2003a). Adapted from Pavlova et al. (2011).

A widely discussed feature of the CMIP3 models is a pronounced underestimation of the trend in the September (annual minimum) sea-ice extent in the Arctic over the past several decades (e.g., Stroeve et al., 2007; Winton, 2011). Possible reasons for the discrepancy include observational uncertainties, vigorous unforced climate variability inherent to high-latitudes, and limitations and shortcomings of the models stemming from gaps in understanding physical process (e.g., Kattsov et al., 2010). Compared to CMIP3, the CMIP5 models better simulate the observed trend of September Arctic ice extent (Figure 9.24). It has been suggested (Stroeve et al., 2012) that in some cases model improvements, such as new sea-ice albedo parameterization schemes that allow for melt ponds (e.g., Holland et al., 2012a; Pedersen et al., 2009), have led to better representation of historical ice conditions. (Holland et al., 2010) show that models with initially thicker ice generally retain more extensive ice throughout the 21st century, and indeed several of the CMIP5 models start the 20th century with rather thin winter ice cover promoting more rapid melt (Stroeve et al., 2012). Notz et al. (2012) caution, however, against direct comparison of modeled trends with observations unless the models' internal variability is carefully taken into account. Their analysis of the MPI-ESM ensemble simulations shows that internal variability in the Arctic can result in individual model realizations to exhibiting a range of trends (negative, or even positive) for the 29-year long period starting in 1979, even if the background climate is warming.

The majority of CMIP5 (and CMIP3) models exhibit a decreasing trend in Southern hemisphere sea ice extent over the satellite era, in contrast to the weak observed increase (see Chapter 4). A large spread in the modelled trends is present, and a comparison of multiple ensemble members from the same model suggests large internal variability during the late 20th century and the first decade of the 21st century (e.g., Landrum et al., 2012; Mahlstein et al., 2012; Zunz et al., 2012).

[INSERT FIGURE 9.24 HERE]

Figure 9.24: Time-series of CMIP5 modelled (coloured lines) and NSIDC (Fetterer et al., 2002) observed (thick red line) Arctic September (upper panel) and Antarctic March (lower panel) sea-ice extent from 1900 to 2012. The CMIP5 multi-model ensemble mean (black) is based on 34 CMIP5 models, with ±1 standard deviation shown as dashed black lines. The dashed thick red line for the Arctic (upper panel) relates to the pre-satellite period of observationally based time series. The upper and lower panel insets are based on the corresponding multi-model ensembles from CMIP5 and CMIP3, ±1 standard deviation. Note that these are monthly means, not yearly minima. Adapted from Stroeve et al. (2012).

Sea ice is a product of atmosphere-ocean interaction. There are a number of ways in which sea ice is influenced by and interacts with the atmosphere and ocean, and some of these feedbacks are still poorly quantified. As noted in the AR4, among the primary causes of biases in simulated sea ice extent, especially its geographical distribution, are problems with simulating high-latitude winds, ocean heat advection and mixing. For example (Koldunov et al., 2010) have shown, for a particular CMIP3 model, that significant ice thickness errors originate from biases in the atmospheric component. Similarly, (Melsom et al., 2009) note sea-ice improvements associated with improved description of heat transport by ocean currents. Biases imparted on modelled sea ice, common to many models, may also be related to representation of high-latitude processes (e.g., polar clouds) or processes not yet commonly included in models (e.g., deposition of carbonaceous aerosols on snow and ice).

Some CMIP5 models show improvements in simulation of sea ice which are likely associated with improvements in simulation of the atmosphere (e.g., Notz et al., 2012). Some recent models are able to simulate rapid changes in the Arctic sea ice due mainly to internal variability (Holland et al., 2008).

There is high confidence that CMIP5 models capture the first-order behaviour of the Arctic sea ice in the climate system, and do so better than CMIP3, particularly the seasonality and the trend of Arctic minimum sea-ice extent. The performance improvements are likely a result of improvements not only in sea-ice components themselves, but also in atmospheric circulation. For the Antarctic, there is less difference between CMIP5 and CMIP3, with the exception of a decreased cold-season bias in sea ice extent. Most CMIP5 models simulate a decrease in Antarctic sea ice extent over the past few decades compared to the slight increase observed.

9.4.4 Land Surface, Fluxes, and Hydrology

9.4.4.1 Snow Cover and Permafrost

The modelling of snow and permafrost processes has received increased attention since the AR4, in part because of the recognition that these processes can provide significant feedbacks on climate change (e.g., Koven et al., 2011; Lawrence et al., 2011). The SnowMIP2 project compared results from thirty-three snowpack models of varying complexity, including some snow models that are used in AOGCMs, using driving data from five Northern Hemisphere locations (Rutter et al., 2009). Consistency between the models and observations was found to be good at open sites, but there was much greater discrepancy at forested sites due to the complex interactions between plant canopy and snow cover. Despite these difficulties, the CMIP5 ensemble is found to simulate large-scale snow-covered area reasonably well (Figure 9.25). Brutel-Vuilmet et al. (2012) show that in the Northern Hemisphere, on average, the models reproduce fairly accurately the seasonal cycle of snow cover over the northern parts of continents, with more disagreement in southerly regions where snow cover is sparse, particularly over China and Mongolia. The latter weaknesses are associated with incorrect timing of the snow onset and melt, and possibly with the choice of thresholds for diagnosing snow cover in the model output. In spite of the good performance of the "mean model", there is a fairly large inter-model scatter of spring snow cover extent in some regions. There is a strong linear correlation between northern hemisphere spring snow cover extent and annual mean surface air temperature in the models, consistent with available observations. The negative trend in spring snow cover (1979–2005) is underestimated by the CMIP5 (and CMIP3) models. The main reason appears to be an underestimate of the boreal land surface warming over that period (Brutel-Vuilmet et al., 2012).

[INSERT FIGURE 9.25 HERE]

Figure 9.25: Terrestrial snow-cover distribution in the Northern Hemisphere simulated by 24 CMIP5 AOGCMs for February, updated for CMIP5 from (Pavlova et al., 2007). For each 1° × 1° longitude-latitude grid cell, the figure indicates the number of models that simulate at least 5 kg m^{-2} of snow water equivalent. The observationally based boundaries (red line) mark the territory with at least 20% of the days per month with snow cover (Robinson and Frei, 2000a) over the period 1986–2005. The annual mean 0°C isotherm at 3.3 m depth averaged across the 24 AOGCMs (yellow line) is a proxy for the permafrost boundary. Observed permafrost extent in the Northern hemisphere (magenta dashed line) is based on Brown et al. (1997, 1998).

Although not typically a direct output, there are a number of approaches to diagnose permafrost extent. One is to use snow depths and skin temperatures generated by climate models to drive a stand-alone multi-layer permafrost model (e.g., Pavlova et al., 2007). A result using this approach is shown in Figure 9.25, and compares the CMIP5-based estimate of the permafrost boundary (as indicated by the 0°C soil temperature isotherm) to that observed, and relatively good agreement is seen for the multi-model mean. However, (Slater and Lawrence, 2012) used CMIP5 model outputs to indirectly estimate permafrost based on climatic indices and to compare them to permafrost extent directly diagnosed via soil temperatures simulated by CMIP5 models. Significant air temperature and snow depth biases in some models degrades both directly and indirectly diagnosed permafrost conditions. . The range of present-day (1986–2005) permafrost area inferred from individual models is therefore extremely large (~4 to 25 × 10^6 km^2) due, firstly, to differences in simulated surface climate and, secondly, to varying abilities of the underlying land surface models. Even though many CMIP5 models include some representation of soil freezing in mineral soils, very few include key processes necessary to accurately model permafrost changes, such as the distinct properties of organic soils, the existence of local water-tables, and the heat released by microbial respiration (Koven et al., 2011; Nicolsky et al., 2007; Wania et al., 2009).

Despite large differences in the absolute permafrost area, the relationship between the decrease in permafrost area and the warming air temperature over the present-day permafrost region is similar, and approximately linear, in many models (Slater and Lawrence, 2012).

9.4.4.2 Soil Moisture and Surface Hydrology

An important land-surface influence is to partition the incoming precipitation into evapotranspiration and runoff, and the net radiation into latent and sensible heat fluxes. In both cases the partitioning is highly-dependent on the moisture status of the land-surface, especially the amount of soil moisture available for evapotranspiration, which in turn depends on properties of the land-cover such as the root-depth of plants.

There has been a long history of off-line evaluation of land-surface schemes, aided more recently by the increasing availability of data from FLUXNET sites (Blyth et al., 2010; Friend et al., 2007). Throughout this time, representations of the land-surface have significantly increased in complexity, allowing the representation of key processes such as links between stomatal conductance and photosynthesis, but at the cost of increasing the number of poorly known internal model parameters. These more sophisticated land-surface models are based on physical principles that should make them more appropriate for projections of future conditions, such as high CO_2. However for specific data-rich sites, current land-surface models still struggle to perform as well as statistical models in predicting year-to-year variations in latent and sensible heat fluxes (Abramowitz et al., 2008) and runoff (Materia et al., 2010).

There are few evaluations of the performance of land-surface schemes in coupled climate models, but those that have been undertaken find major limitations associated with the atmospheric forcing rather than the land-surface schemes themselves. For example, Li et al. (2007) evaluated the soil moisture simulations of CMIP3 models, and found that long-term soil moisture trends could only be reproduced in models that simulated "global dimming". Land-surface schemes are believed to play a key role in determining the ability of climate models to simulate the influence of soil moisture anomalies on rainfall, droughts, and high-temperature extremes (Fischer et al., 2007a). Comparison of climate model simulations to observations suggests that the models correctly represent the soil-moisture impacts on temperature extremes in south-eastern Europe, but overestimate them in central Europe (Hirschi et al., 2011a). The influence of soil moisture on rainfall varies with region, and with the lead-time between a soil moisture anomaly and a rainfall event (Seneviratne et al., 2010). In some regions, such as the Sahel, enhanced precipitation can even be induced by dry anomalies (Taylor et al., 2011). A recent analysis of CMIP5 models reveals considerable spread in the ability of the models to reproduce observed correlations between precipitation and soil moisture in the tropics (Williams et al., 2012).

9.4.4.3 Dynamic Global Vegetation and Nitrogen Cycling

At the time of the AR4 very few climate models included dynamic vegetation, with vegetation being prescribed and fixed in all but a handful of coupled climate-carbon cycle models (Friedlingstein et al., 2006). Dynamic Global Vegetation Models (DGVMs) certainly existed at the time of the AR4 (Cramer et al., 2001) but these were not typically incorporated in climate models. Since the IPCC AR4 there has been continual development of offline DGVMs, and many climate models incorporate dynamic vegetation in at least a subset of the runs submitted to CMIP5.

DGVMs are designed to simulate the large-scale geographical distribution of plant functional types and how these patterns will change in response to climate change, CO_2 increases, and other forcing factors (Cramer et al., 2001). These models typically include rather detailed representations of plant photosynthesis but less sophisticated treatments of soil carbon, with a varying number of soil carbon pools and lifetimes between models. In the absence of nitrogen limitations on CO_2 fertilization, offline DGVMs agree qualitatively that CO_2 increase alone will tend to enhance carbon uptake on the land while the associated climate change will tend to reduce it. There is also good agreement on the degree of CO_2 fertilization in the case of no nutrient limitation (Sitch et al., 2008). However, under more extreme emissions scenarios the responses of the DGVMs diverge markedly. Large uncertainties are associated with the responses of tropical and boreal ecosystems to elevated temperatures and changing soil moisture status. Particular areas of uncertainty are the high-temperature response of photosynthesis (Galbraith et al., 2010), and the extent of CO_2 fertilization (Rammig et al., 2010) in Amazonian rainforest.

Most of the 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 CO_2 (Wang and Houlton, 2009). Recent extensions of two DGVMs to include nitrogen limitations to CO_2-fertilization improve the fit of these models to “Free-Air CO_2 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).

9.4.4.4 Land-Use Change

A major innovation in the land component of ESMs since the AR4 is the inclusion of the effects of land-use change associated with the spread of agriculture, urbanization and deforestation. These affect climate by altering the biophysical properties of the land-surface, such as its albedo, aerodynamic roughness and water-holding capacity (Bondeau et al., 2007; Bonan, 2008; Levis, 2010; Bathiany et al., 2010). Land-use change has also contributed almost 30% of total anthropogenic CO_2 emissions since 1850 (see Table 6.1), and affects emissions of trace gases, and volatile organic compounds such as isoprene. The latest ESMs used in CMIP5 attempt to model the CO_2 emissions implied by prescribed land-use change and many also simulate the associated changes in the biophysical properties of the land-surface. This represents a major advance on the CMIP3 models which typically neglected land-use change, aside from its assumed contribution to anthropogenic CO_2 emissions.

However, the increasing sophistication of the modelling impacts of land-use change has introduced additional spread in climate model projections. The first systematic model intercomparison demonstrating that large-scale land cover change can significantly affect regional climate (Pitman et al., 2009) showed a large spread in the response of different models to the same imposed land-cover change.

9.4.5 Carbon Cycle

9.4.5.1 Terrestrial Carbon Cycle Component Models

When driven with the observed climate, dynamic global vegetation models are able to reproduce the observed land-atmosphere fluxes of CO_2 to within 30% and can replicate the greater carbon uptake observed in the 1990s compared to the 1980s (Sitch et al., 2008). These models also correctly simulate a reduction in the land carbon sink during El Nino and an increase during La Nina, but with varying sensitivities (Sitch et al., 2008). Several coupled biogeochemistry/land-surface models underestimate the seasonal amplitude of CO_2 in the northern hemisphere by factors of 2 to 3 (Randerson et al., 2009), but in most models the phasing of the annual cycle of CO_2 over northern latitudes is accurate, and the timing of observed spring drawdown of CO_2 is reproduced to within 1 month in the tropics with increasing phasing errors between 60°N and 90°N (Cadule et al., 2010).

Accurate simulation of the Amazon is important for representing its buffering of atmospheric CO_2 and for projecting the effects of climate change on the amount of carbon stored in forests (Lewis et al., 2011). While two biogeochemical sub-models reproduced the gross primary productivity (GPP) of the Amazonian forests to within 14% of observational estimates (Lewis et al., 2011), the models overestimated the above-ground live biomass by 130 to 190% and underestimated soil carbon by 33 to 40% (Randerson et al., 2009). The overestimation of live biomass in the Amazon is consistent with excessive allocation of net primary productivity (NPP) to wood and an underestimationof plant respiration (Randerson et al., 2009).

Wildland and human-induced fires have been estimated to contribute 2.3 PgC yr^{-1} to the atmosphere in the period 1997–2004 (Randerson et al., 2009), which is equivalent to about 35% of the CO_2 emissions from fossil fuel burning over the same period. Inadequate parameterisations of fires can lead to underestimation of this flux by a factor of 3 and to errors in its spatial and temporal variability caused by deforestation-linked fires and the effects of drought. Recent advances in parameterisations in off-line simulations yield reasonably good agreement with satellite-based emission retrievals on interannual timescales (Kloster et al., 2010). However, forest fires are typically not simulated in the CMIP5 ESMs.

9.4.5.2 Oceanic Carbon Cycle Component Models

As with land models, it is possible to evaluate ocean models “offline” driven by observed meteorological variables, rather than “online” within a coupled ESM. Recent advances in the observational evaluation of offline ocean ecosystem-biogeochemical (OBGC) models include new diagnostic frameworks (Doney et al., 2009) and protocols to evaluate the impact of ocean circulation on the marine carbon cycle, including export production, dissolved organic matter, and dissolved oxygen (Najjar et al., 2007). Results so far show that the fidelity of biological properties of OBGCs is contingent on corresponding accuracy of the simulated physical ocean (Doney et al., 2009; Najjar et al., 2007), in particular the SSTs, mixed-layer depths (MLDs), upwelling rates, and vertical structure near the surface.

Evaluation of OBGCs has often focused on regional oceanic uptake of CO_2 (Roy et al., 2011b).Based on available results, declining rates of net ocean CO_2 uptake observed in the temperate North Atlantic are broadly reproduced by historical OBGC model simulations (Thomas et al., 2008). These trends represent a superposition of interannual variability associated with the NAO and secular trends in surface temperature. The positive trend in observed sea-air CO_2 partial pressure differences between 1997 and 2004, which is indicative of reduced oceanic uptake or greater outflow of CO_2, is also simulated. However, models that have been evaluated against estimates of surface chlorophyll concentrations cannot reproduce the regime shifts observed in the Northern Atlantic since 1948 (Henson et al., 2009) or the broad-scale shifts from lower to higher biomass-normalized primary productivity between the 1980s and 1990s (Friedrichs et al., 2009). The greater skill in reproducing surface CO_2 fields compared to ecological variables including chlorophyll concentrations is consistent with the relative skills in these fields observed by Doney et al (2009). The errors in reproducing decadal regime shifts are due to challenges in modelling the phytoplankton community structure, the impact of the Gulf Stream on biological variability downstream, and transitions between ecological states (Henson et al., 2009).

9.4.5.3 The Carbon Cycle in Earth System Models

The transition from climate models to ESMs was motivated in part by the results from the first generation coupled climate-carbon cycle models, which suggested that feedbacks between the climate and the carbon cycle were uncertain but potentially very important in the context of 21st century climate change (Cox et al., 2000; Friedlingstein et al., 2001). The first generation models used in the Coupled Climate Carbon Cycle Model Intercomparison Project (C^4MIP) included both extended AOGCMS and EMICs. The C^4MIP experimental design involved running each model under a common emission scenario (SRES A2) and calculating the evolution of the global atmospheric CO_2 concentration interactively within the model. The impacts of climate-carbon cycle feedbacks were diagnosed by carrying-out parallel “uncoupled” simulations in which increases in atmospheric CO_2 did not influence climate. Analysis of the C^4MIP runs highlighted: (a) a greater than 200 ppmv range in the CO_2 concentration by 2100 due to uncertainties in climate-carbon cycle feedbacks, and (b) that the largest uncertainties were associated with the response of land ecosystems to climate and CO_2 (Friedlingstein et al., (2006)).

For CMIP5 a different experimental design was proposed in which the core simulations use prescribed Representative Concentration Pathways (RCPs) of atmospheric CO_2 and other greenhouse gases (Moss et al., 2010). Under such a prescribed CO_2 scenario, ESMs still calculate land and ocean carbon fluxes interactively, but these fluxes do not affect the evolution of atmospheric CO_2. Instead the modelled land and ocean fluxes, along with the prescribed increase in atmospheric CO_2, can be used to diagnose the "compatible" emissions of CO_2 consistent with the simulation (Miyama and Kawamiya, 2009; Arora et al., 2011a). The compatible emissions for each model can then be evaluated against the best estimates of the actual historical CO_2 emissions. Parallel model experiments in which the carbon cycle does not respond to the simulated climate change (which are equivalent to the "uncoupled" simulations in C^4MIP) provide a means to diagnose climate-carbon cycle feedbacks in terms of their impact on the compatible emissions of CO_2 (Hibbard et al., 2007).

Figure 9.26 shows modelled annual mean net land-atmosphere (Net Biome Productivity, "NBP") and ocean-atmosphere CO_2 ("fgCO2") fluxes from the historical RCP simulations in the CMIP5 archive (Anav et al., Submitted). Also shown are the estimates provided by the Global Carbon Project (GCP) which are derived from offline ocean carbon cycle models, measurements of atmospheric CO_2, and best estimates of the CO_2 fluxes from fossil fuels and land-use change (Le Quere et al., (2009)). Uncertainties in these latter annual estimates are approximately ±0.5 PgC yr^{-1}, arising predominantly from the uncertainty in the model-derived ocean CO_2 uptake. The confidence limits for the ensemble mean are derived by assuming that the CMIP5 models form a t-distribution centred on the ensemble mean (Anav et al., Submitted). An evaluation of climate-carbon cycle simulations produced with Earth System Models of Intermediate Complexity (EMICS) has also been carried-out (Eby et al., 2012).

The top panel of Figure 9.26 shows the variability in global land carbon uptake evident in the GCP estimates, with the global land carbon sink being strongest during La Nina years and after volcanoes, and turning into a source during El Nino years. The CMIP5 models cannot be expected to precisely reproduce this year-to-year variability as these models will naturally simulate chaotic ENSO variability that is out of phase with the historical variability. However, the ensemble mean does successfully simulate a strengthening global land carbon sink during the 1990s, especially after the Pinatubo volcano in 1991. The CMIP5 ensemble mean land-atmosphere flux evolves from a small source of –0.35 ± 0.6 PgC yr^{-1} over the period 1901–1930, predominantly due to land-use change, to a sink of 0.6 ± 0.7 PgC yr^{-1} in the period 1960–2005. The GCP estimates give a weaker sink of 0.36 ± 1 PgC yr^{-1} for the 1960–2005 period. The evolution of the global ocean carbon sink is shown in the bottom panel of Figure 9.26. The CMIP5 ensemble mean global ocean uptake increases from 0.52 ± 0.26 PgC yr^{-1} over the period 1901–1930 to 1.47 ± 0.58 PgC yr^{-1} for the period 1960–2005. For comparison, GCP estimates a stronger ocean carbon sink of 1.92 ± 0.3 PgC yr^{-1} for 1960–2005 (Anav et al., Submitted).

Figure 9.27 shows the mean land-atmosphere fluxes and ocean-atmosphere fluxes simulated by each of the CMIP5 models for the period 1986-2005, and compares these to observation-based estimates from GCP, atmospheric inversions ("JMA", Gurney et al., 2003), and ocean pCO_2 measurements (Takahashi et al., 2009). The error-bars represent the interannual variation in the form of the standard deviation of the annual fluxes. Here, as in Figure 9.26, the net land-atmosphere flux is "Net Biome Productivity (NBP)" which includes the net CO_2 emissions from land-use change as well as the changing carbon balance of undisturbed ecosystems. The observation-based estimates of GCP and JMA agree well on the mean global land carbon sink over the period 1986–2005, and most models fit within the uncertainty bounds of these estimates (i.e., 1.17 ± 1.06 PgC yr^{-1} for JMA). The exceptions are INMCM4 which has a larger land carbon sink, and CCSM4, NorESM1-ME and NorESM2-M which model a net land carbon source rather than a sink over this period. Some models (notably GFDL-ESM2M and GFDL-ESM2G) significantly overestimate the interannual variation in the global land-atmosphere CO_2 flux, with a possible consequence being an overestimate of the vulnerability of tropical ecosystems to future climate change (Cox et al., 2012, and see Figure 9.46). There are also systematic differences between the CMIP5 models and the JMA inversion estimates for the large-scale zonal means, with the ESMs tending to produce weaker uptake in the northern hemisphere (perhaps because of the neglect of nitrogen-fertilization) and simulating a net land carbon sink rather than a source in the tropics.

For the period 1986–2005 the observation-based estimates of the global ocean carbon sink are 1.71 PgC yr^{-1} (JMA), 2.19 PgC yr^{-1} (GCP) and 2.33 PgC yr^{-1} (Takahashi et al., 2009). Most of the CMIP5 models simulate ocean sinks within the range of these estimates, with the exception being INMCM4 which simulates a much larger sink of 2.97 ± 0.37 PgC yr^{-1}. At the regional scale most CMIP5 models also simulate the expected pattern of outgassing of CO_2 in the tropics and an uptake of CO_2 in the mid and high latitudes. The models are generally able to reproduce the estimated sinks in the Southern Ocean (0.73 PgC yr^{-1} (JMA) to 1.28 PgC

yr^{-1} (Takahashi et al., 2009)), and the mid-latitude Northern Hemisphere (0.74 PgC yr^{-1} (JMA) to 1.15 PgC yr^{-1} (Takahashi et al., 2009)), as well as the source of CO_2 from the tropical oceans (–0.73 PgC yr^{-1} (JMA) to –1.25 PgC yr^{-1} (Takahashi et al., 2009)). The most obvious exception to this is INMCM4 which simulates a tropical ocean carbon sink rather than a source and overestimates the net sink in the northern mid-latitudes.

[INSERT FIGURE 9.26 HERE]

Figure 9.26: Ensemble mean of annual global land carbon uptake (top) and annual global ocean carbon uptake (bottom panel) in the CMIP5 ESMs for the historical period 1900–2005. For comparison, the observation-based estimates provided by the Global Carbon Project ("GCP", Le Quere et al., 2009) are also shown (black line). The confidence limits on the ensemble mean are derived by assuming that the CMIP5 models come from a t-distribution. The grey areas show the range of annual fluxes simulated across the model ensemble.

[INSERT FIGURE 9.27 HERE]

Figure 9.27: Simulation of net (a) atmosphere-land CO_2 fluxes ("NBP", top) and (b) atmosphere-ocean CO_2 fluxes ("fgCO2", bottom) in the CMIP5 ESMs, for the period 1986–2005. The error bars represent the interannual variability in the fluxes calculated as the standard deviation of the annual fluxes. . For comparison, the observation-based estimates provided by the Global Carbon Project ("GCP", Le Quere et al., 2009), the "JMA" atmospheric inversion (Gurney et al., 2003), and the Takahashi ocean pCO_2 dataset (Takahashi et al., 2009) are also shown as the red symbols, with the horizontal red line representing the estimate from JMA.

In summary, there is robust evidence and medium agreement that CMIP5 ESMs can simulate the mean global land and ocean carbon sinks, and the broad-scale spatial pattern of ocean-atmosphere CO_2 fluxes, within the range of observational estimates. EMICS can also reproduce the recent global ocean sink, but appear to underestimate the contemporary land carbon sink (Eby et al., 2012). With few exceptions, the CMIP5 ESMs also reproduce the large-scale pattern of ocean-atmosphere CO_2 fluxes, with uptake in the Southern Ocean and northern mid-latitudes, and outgassing in the tropics. However, the geographical pattern of simulated land-atmosphere fluxes agrees much less well with inversion estimates, which suggest a larger sink in the northern mid-latitudes, and a net source rather than a sink in the tropics. While there are also inherent uncertainties in atmospheric inversions, discrepancies of these types might be expected from known deficiencies in the CMIP5 generation of ESMs—namely the neglect of nitrogen fertilization in the mid-latitudes, and the fairly rudimentary treatment of the net CO_2 emissions arising from land-use change and forest regrowth.

9.4.6 Aerosol Burdens and Effects on Insolation

9.4.6.1 Recent Trends in Global Aerosol Burdens and Effects on Insolation

The historical emissions data used to drive the CMIP5 simulations of the 20th century reflect two recent trends in regional and global anthropogenic SO_2 emissions. During the last three decades, anthropogenic emissions of SO_2 from North America and Europe have declined due to the imposition of emission controls, while the emissions from Asia have increased. The combination of the European, North American, and Asians trends has yielded a global reduction in SO_2 emissions by 20 $GgSO_2$, or 15% between 1970 and 2000 although emissions subsequently increased by 9 $GgSO_2$ between 2000 and 2005 (Smith et al., 2011b). For the period 2001 to 2005, CMIP5 models understimeate the mean AOT due to all tropospheric aerosol speices relative to satellite-retrieved AOT by at least 20% over virtually all land surfaces (Figure 9.28).

The effects of sulphate and other aerosol species on surface insolation through direct and indirect forcing appear to be one of the principal causes of the "global dimming" between the 1960s and 1980s and subsequent "global brightening" in the last two decades. This inference is supported by trends in aerosol optical depth and trends in surface insolation under cloud-free conditions. Thirteen out of fourteen CMIP3 models examined by (Ruckstuhl and Norris, 2009) produce a transition from "dimming" to "brightening" that is consistent with the timing of the transition from increasing to decreasing global anthropogenic aerosol emissions. The transition from "dimming" to "brightening" in both Europe and Asia is well simulated with the HadGEM2 model (Haywood et al., 2011).

These recent trends are superimposed on a general upward trend in aerosol loading since 1850 reflected by an increase in global-mean ocanic AOTs from the CMIP5 historical and RCP 4.5 simulations from 1850 tp 2010 (Figure 9.29). Despite the use of common anthropogenic aerosol emissions for the historical simulations (Lamarque et al., 2010), the simulated oceanic AOTs for 2010 range from 0.08 to 0.215, a factor of 2.7, with nearly equal numbers of models over and underestimating the satellite retrieved AOT of 0.12 (Figure 9.29). This range in AOTs results from differing estimates of the trends and of the initial global-mean oceanic AOT at 1850 across the CMIP5 ensemble (Figure 9.29).

[INSERT FIGURE 9.28 HERE]

Figure 9.28: The relative error in visible aerosol optical thickness (AOT) from the median of a subset of CMIP5 models' historical simulations, relative to satellite retrievals of AOT. Units are such that +1 is equivalent to satellite AOT exceeding model AOT by 100%. The figure is constructed following (Kinne et al., 2006). The satellite AOT is from the MODIS instrument on the NASA Terra satellite from 2001 through 2005. The data version is MODIS 4; the model outputs are from ACCESS1-0, ACCESS1-3, BNU-ESM, CESM1-CAM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadGEM2-CC, HadGEM2-ES, IPSL-CM5A-LR IPSL-CM5A-MR, IPSL-CM5B-LR, MIROC-ESM, MIROC-ESM-CHEM, MIROC5, MRI-CGCM3, NorESM1-M, and NorESM1-ME.

[INSERT FIGURE 9.29 HERE]

Figure 9.29: Time series of global oceanic-mean AOT from individual CMIP5 models' historical (1850–2005) and RCP4.5 (2006–2010) simulations, corrected MODIS satellite observations by Shi et al. (2011) and Zhang et al. (2008a), and the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) simulations for the 1850s by Shindell et al. (2012b). ACCMIP model output is from CICERO-OsloCTM2, GISS-E2-R, HadGEM2, LMDzORINCA, NCAR-CAM3.5, and NCAR-CAM5.1.

9.4.6.2 Principal Sources of Uncertainty in Projections of Sulphate Burdens

In contrast to CMIP3, the CMIP5 ensemble is driven by a single internally consistent set of SO_4 concentrations and SO_2 emissions. The use of a single set of emissions removes an important, but not dominant, source of uncertainty in the AR5 simulations of the sulphur cycle. In experiments based upon a single chemistry-climate model with perturbations to both emissions and sulphur-cycle processes, uncertainties in emissions accounted for 53.3% of the ensemble variance (Ackerley et al., 2009). The next largest source of uncertainty was associated with the wet scavenging of sulphate, which accounted for 29.5% of the intra-ensemble variance and represents the source/sink term with the largest relative range in the aerosol models evaluated by AEROCOM (Faloona, 2009). Similarly, AEROCOM simulations run with heterogeneous or harmonized emissions data sets yielded approximately the same intermodel standard deviation in sulphate burden of 25 Tg for both sets of experiments. These results show that a dominant source of the spread among the sulphate burdens is differences in the treatment of chemical production, transport, and removal from the atmosphere (Liu et al., 2007; Textor et al., 2007).

Natural sources of sulphate from oxidation of natural dimethylsulphide (DMS) emissions from the ocean surface are not specified under the RCP protocol and therefore represent an additional source of uncertainty in the sulphur cycle simulated by the CMIP5 ensemble. In simulations of present-day conditions, DMS emissions span a 5 to 95% confidence interval of 10.7 to 28.1 TgS yr^{-1} (Faloona, 2009). After chemical processing, DMS contributes between 18 to 42% of the global atmospheric sulphate burden and up to 80% of the sulphate burden over most the southern hemisphere (Carslaw et al., 2010). The effects from differences in DMS emissions and its subsequent oxidation to sulphate on sulphate burdens in the CMIP5 ensemble remain to be quantified. Several of the CMIP5 models include prognostic models of the biogenic DMS source.

9.5 Simulation of Variability and Extremes

9.5.1 Importance of Simulating Climate Variability

The ability of a model to simulate the mean climate, and the slow, externally-forced change in that mean state, is important and was evaluated in the previous Section. However, the ability to simulate climate variability, both unforced internal variability and forced variability (e.g., diurnal and seasonal cycles) is also important. This has implications for the signal-to-noise estimates inherent in climate change detection and attribution studies where low-frequency climate variability must be estimated, at least in part, from long control integrations of climate models. It also has implications for the ability of models to make quantitative projections of changes in climate variability and the statistics of extreme events under a warming climate. In many cases, the impacts of climate change will be experienced more profoundly in terms of the frequency, intensity or duration of extreme events (e.g., heat waves, droughts, extreme rainfall events). The ability to simulate climate variability is also central to the topic of climate prediction, since it is the ability to simulate the specific evolution of the varying climate system, beyond that due to external forcing, that provides useful predictive skill.

Evaluating model simulations of climate variability also provides a means to explore the representation of certain processes, such as the coupled processes underlying the El Niño Southern Oscillation (ENSO) and other important modes of variability. A model's representation of the diurnal or seasonal cycle – both of which represent responses to external (rotational or orbital) forcing – may also provide some insight into a model's 'sensitivity' and by extension, the ability to respond correctly to greenhouse gas, aerosol, volcanic and solar forcing.

In this Section we will also investigate the extent to which biases in the simulation of the mean climate and its long-term evolutionare related to biases in variability, and we will explore to some extent model features, such as resolution, that may affect the simulation of variability, particularly aspects such as atmospheric blocking and convective precipitation events.

9.5.2 Diurnal-to-Seasonal Variability

9.5.2.1 Diurnal Cycles of Physical Climate Variables

The diurnally varying solar radiation received at a given location drives, through complex interactions with the atmosphere, land surface, and upper ocean, easily observable diurnal variations not only in surface and near-surface temperature, but also precipitation, low level stability and winds, and many other geophysical parameters. A good representation of the diurnal cycle requires therefore a correct representation of various interactions and feedbacks. The AR4 noted that climate models simulated the global pattern of the diurnal temperature, zonally and annually averaged over the continent, but tended to underestimate the magnitude in many regions (Randall et al., 2007). The diurnal cycle of precipitation for several CMIP3 models is shown in Figure 9.30. The overall spatial distribution of the diurnal amplitude is realistic, but most models tend to start moist convection prematurely (Wang et al., 2011a) and thus often rain too frequently at reduced intensity, resulting in a rainfall peak too early in the day and the so-called "drizzling bias" (Dai, 2006; Stephens et al., 2010) that can have large adverse impacts on surface evaporation and runoff (Qian et al., 2006). Many models also produce too much convective rain but too little stratiform precipitation compared with satellite data (Dai, 2006). Over the tropical continent models tend to produce earlier development of diurnal precipitation (Dai, 2006). In addition, it was shown that CMIP3 models are able to reproduce the observed surface pressure tides despite the low top in many of the models (Covey et al., 2011), and the associated diurnal variations in tropospheric wind are also broadly reproduced, though with relatively weak amplitudes over oceans based on limited analyses (Dai and Trenberth, 2004)

[INSERT FIGURE 9.30 HERE]

Figure 9.30: Composite diurnal cycle precipitation from observations (black) and a subset of CMIP3 models (coloured lines) averaged over land (left) and ocean (right) areas for three different zones at each local time and seasons (JJA, DJF). Black solid line: surface-observed precipitation frequency; black dashed line: TRMM 3B42 dataset, 1998–2003 mean]. Each colored curve is for one model: CCSM2 (red solid), GFDL-CM2.0 (green solid), GISS-ER (blue), MIROC3.2 (red dashed), and MRI-CGCM2.3.2a (green dashed). Adapted from (Dai, 2006).

These conclusions still hold for most of the CMIP5 models, even though there has so far been limited analysis of the diurnal cycle. It has been suggested that a weak diurnal cycle of surface air temperature is produced over the ocean because of a lack of diurnal variations in sea surface temperature (SST). In recent simulations using an AOGCM that explicitly represents SST diurnal variations, some long-standing model biases, such as cold biases in the tropical Pacific, have been reduced (Bernie et al., 2008). However, most models still have difficulties in simulating the diurnal variations in SST due to coarse vertical resolution and coupling frequency in these models (Dai and Trenberth, 2004; Danabasoglu et al., 2006). In the atmosphere, the focus has been on deficiencies in cumulus convection, but it is likely that model deficiencies in surface-atmosphere interactions and the planetary boundary layer also contribute to some of the diurnal cycle errors.(Lindvall et al., 2012). Model agreement with observations depends on region, vegetation type and season.

Improved representation of the diurnal cycle has been found with increased atmospheric resolution (Ploshay and Lau, 2010; Sato et al., 2009) or with improved representation of cloud physics (Khairoutdinov et al., 2005), but the reasons for these improvements remain poorly understood. Other changes such as the representation of entrainment in deep convection (Stratton and Stirling, 2012), improved coupling between shallow and deep convection, and inclusion of density currents (Peterson et al., 2009) have been shown to greatly improve the diurnal cycle of convection over tropical land and provide a good representation of the timing of convection over land in coupled ocean-atmosphere simulations (Hourdin et al., in press).

9.5.2.2.1 Blocking

In the mid latitudes, climate is often characterised by weather regimes (see Chapter 2). Recent work has underlined the importance of blocking regimes for the occurrence of extreme weather events (Buehler et al., 2011; Sillmann et al., 2011). During blocking, the prevailing midlatitude westerly winds and storm systems are interrupted by a local reversal of the zonal flow. The AR4 was not conclusive on the ability of climate models to represent this phenomenon, though climate and weather prediction models in the past have universally underestimated the occurrence of blocking. Recent work has shown that very high resolution atmospheric models can now simulate the observed level of blocking in both hemispheres (Matsueda, 2009; Matsueda et al., 2009b; Matsueda et al., 2010b) These improvements arise from increased resolution of orography and improved atmospheric dynamics (Berckmans et al., 2012; Jung et al., 2012) while others have shown that reduced ocean surface temperature errors in the extratropics are a key source of the blocking improvement in high resolution coupled models (Scaife et al., 2011a). However, significant underestimation of blocking is still a feature of all the CMIP3 models (Barnes et al., 2012; Scaife et al., 2010) and most of the CMIP5 models as well (Anstey et al., 2012; Masato et al., 2012).

Since the AR4 there has been a renewed focus on the diagnostic methods used to characterize blocking. There are still important differences between methods (Barriopedro et al., 2010a), and the diagnosed blocking frequency can be very sensitive to details such as in the choice of latitude (Barnes et al., 2012). In particular, blocking indices based on the identification of reversed meridional gradients in quantities such as geopotential height can be sensitive to mean state biases in the models (Scaife et al., 2010) as well as problems with modelled variability (Barriopedro et al., 2010b; Vial and Osborn, 2012). When blocking is measured via anomaly fields, rather than reversed absolute fields, model skill can appear better (eg (Sillmann and Croci-Maspoli, 2009)).

Recent work has confirmed the link between blocking events and stratospheric flow anomalies (Martius et al., 2009). This link mostly comprises blocking events perturbing the stratospheric flow through upwards propagating Rossby wave activity, and is well represented in climate models with enhanced stratospheric resolution (Woollings et al., 2010a) . Indeed, models tend to show a small but consistent improvement in blocking when stratospheric representation is improved (Anstey et al., 2012).

In summary, the representation of blocking events is improving in models, even though still underestimated, and there is medium confidence that high-resolution, high-top models better represent blocking.

9.5.2.2.2 Madden Julian Oscillation

During the boreal winter the eastward propagating feature known as the Madden-Julian Oscillation (MJO; (Madden and Julian, 1972, 1994) predominantly affects the deep tropics, while during the boreal summer there is also northward propagation over much of southern Asia (Annamalai and Sperber, 2005). The MJO has received a lot of attention given the prominent role it plays in tropical climate variability in general, and in specific phenomena such as monsoons and ENSO. Cassou (2008) and Pan and Li (2008) also suggest that the quality of the MJO affects model ability to properly reproduce weather regimes over the North Atlantic. Previous assessments reported that most AOGCMs have difficulty in representing intraseasonal MJO variability with most models underestimating the strength and the coherence of convection and wind variability (Lin and Li, 2008; Lin et al., 2006). Simulation of the Madden-Julian Oscillation is still a challenge for climate models (Kim et al., 2009; Lin et al., 2006; Xavier et al., 2010), however, Sperber and Annamalai (2008) have shown that CMIP3 models were able to simulate eastward propagating intraseasonal convection over the Indian Ocean. This represents an improvement over earlier models (Waliser et al., 2003), though it must be noted that only two of seventeen models were able simulate the observed northward propagation during boreal summer.

Several diagnoses have been proposed to evaluate MJO in climate models (Waliser et al., 2009a) and process-oriented diagnostics are being applied (e.g., Xavier, 2012). The simplified metric shown in figure 9.31 provides a synthesis of CMIP3 and CMIP5 model results (Sperber and Kim, 2012). The maximum positive correlation of the two leading principal component time series and the time lag at which it occurs indicate that all models have less coherent eastward propagation than observed. There is a diverse representation of the time scale of the simulated MJO, and some models are incorrectly dominated by westward propagation. However, (Sperber et al., submitted) note that several CMIP5 models are able to capture the northward propagation of convection over India as the equatorial eastward propagation extends into the western Pacific, thus exhibiting an improved capability compared to CMIP3.

An important reason for model errors in representing the MJO is that convection parameterizations do not provide sufficient build-up of moisture in the atmosphrer for the large scale organized convection to occur (Kim et al., 2012a; Mizuta et al., 2012). High frequency coupling with the ocean was also highlighted as an important factor (Bernie et al., 2008). While new parameterizations of convection may improve the MJO, this sometimes occurs at the expense of a good representation of the mean tropical climatology in the current generation of climate models (Kim et al., 2012a). In addition, high resolution models with improved diurnal cycle do not necessarily produce improved MJO (Mizuta et al., 2012).

In summary, there is robust evidence that the MJO is still poorly represented in climate models, even though some CMIP5 models are now able to reproduce the northeastward propagation in summer in the Indian ocean.

[INSERT FIGURE 9.31 HERE]

Figure 9.31: Two leading Empirical Orthogonal Functions (EOF's) of Outgoing Longwave Radiation (OLR) from years of strong MJO variability (Sperber, 2003). The 20–100 day filtered OLR from observations and each of the CMIP5 historical simulations and the CMIP3 simulations of 20th century climate is projected on these two leading EOF's to obtain MJO Principal Component time series. The scatter plot shows the maximum positive correlation between the resulting MJO Principal Components and the time lag at which it occurred for all winters (November-March). The maximum positive correlation is an indication of the coherence with which the MJO convection propagates from the Indian Ocean to the Maritime Continent/western Pacific, and the time lag is approximately 1/4 of the period of the MJO. Most models have weaker coherence in the MJO propagation (smaller maximum positive correlation), and some have periods that are too short compared to observations. One CMIP3 model has its maximum positive correlation at Day –16, indicating that this model is incorrectly dominated by westward propagation. Constructed following Sperber and Kim (2012).

9.5.2.3 Large Scale Monsoon Rainfall and Circulation

The monsoon is the dominant mode of annual variation in the tropics (Trenberth et al., 2000; Wang and Ding, 2008), and so high fidelity simulation of the mean monsoon and its variability is of great importance for simulating future climate impacts (Colman et al., 2011; Sperber et al., 2010; Turner and annamalai, 2012; Wang et al., 2006). The monsoon is characterised by an annual reversal of the low level winds and well defined dry and wet seasons (Wang and Ding, 2008), and its variability is connected to the MJO and ENSO. The AR4 reported that most CMIP3 models poorly represent the characteristics of the monsoon and monsoon teleconnections.

The different monsoon systems are connected through the large-scale tropical circulation, offering the possibility to evaluate a model's representation of monsoon domain and intensity through the global monsoon concept (Wang and Ding, 2008). Relevant diagnostics, shown in Figure 9.32, are based on the annual range of hemispheric precipitation and provide a large-scale view of the 'time-mean' monsoon (Wang et al., 2011a). The CMIP3 multi-model ensemble generally reproduces the observed spatial patterns but somewhat underestimates the extent and intensity, especially over Asia and North America. In terms of the threat score (a categorical metric (Wilks, 1995) which indicates how well a model simulates the monsoon precipitation domain) the best CMIP3 model outperforms the multi-model mean, whereas the poorest models fail to capture the monsoon precipitation domain over the Sahel, Central America, and Australia. (Fan et al., 2010) also show that CMIP3 simulations capture the observed trend of weakening of the South Asian summer circulation over the past half century, but are unable to reproduce the magnitude of the corresponding observed trend in precipitation.

A cold bias in SST over the Arabian Sea during boreal summer, as simulated in the HadGEM3 model for example, leads to reduced evaporation and too weak inland moisture transport (Levine and Turner, 2012). Poor simulation of the monsoon has been attributed to SST patterns and unrealistic development of the Indian Ocean dipole (Achuthavarier et al., 2012; Boschat et al., 2012), and similar biases contiribute to model-data mismatch in the simulation of the mid-Holocene Asian monsoon (Ohgaito and Abe-Ouchi, 2009). This bias has been improved in some models (e.g., Meehl et al., 2012). Another factor contributing to model improvements is intraseasonal variability (Joseph et al., 2012), and in some models the improved representation of monsoon results from a more realistic northward propagation of boreal-summer intraseasonal variability (Sperber et al., submitted). In some models topography-related monsoon precipitation is also better simulated at high horizontal resolution

[INSERT FIGURE 9.32 HERE]

Figure 9.32: Monsoon precipitation intensity (shading, mm/day) and monsoon precipitation domain (lines) are shown for (a) observations from GPCP, (b) the CMIP3 multi-model mean, (c) the best model, and (d) the worst model in terms of the threat score for this diagnostic. The threat scores indicate how well the models represent the monsoon precipitation domain compared to the GPCP data. The threat score in panel (a) is between GPCP and CMAP rainfall to indicate observational uncertainty. A threat score of 1.0 would indicate perfect agreement between the two datasets. See Wang and Ding (2008); Wang et al. (2011a); and Kim et al. (2011) for details of the calculations.

In summary, there is some evidence that the CMIP5 models simulate more realistic monsoon climatology and variability than their CMIP3 predecessors, but the still suffer from systematic biases that limit model credibility at the regional scale both in terms of pattern and magnitude.

9.5.3 Interannual-to-Centennial Variability

In addition to the annual and diurnal cycles described above, a number of other modes of variability arise from interactions between the various components on a number of time and space scales. Here we limit the scope to modes of variability whose timescale ranges from a few years to the multi-decadal features that can modulate the trend arising from changes in GHGs. Most of these modes have a particular regional manifestation whose amplitude is ususally much larger than that of human-induced climate change (see Chapter 14). The observational record is sometimes too short to fully evaluate the representation of variability in models and this motivates the use of re-analysis or proxies, even though these have their own limitations. In the following, we also emphasize recent research on the interactions between modes of variability via teleconnections, the processes involved, and model improvements since the AR4.

9.5.3.1 Global Surface Temperature Variability

Accurately simulating climate variability on various time scales is important both for detection and attribution and for predictions and projections of future climate. The AR4 concluded that modelled global variance at decadal to inter-decadal time scales was consistent with 20th century observations. In addition results from the last millennium suggest that simulated coupled ocean-atmosphere internal variability is consistent with indirect estimates (Hegerl et al., 2007).

Figure 9.33 illustrates that the model spread in the simulated internal variability from CMIP5 pre-industrial control simulations is largest in the tropics and mid to high latitudes (Jones et al., 2012), where variability is also large; however, compared to CMIP3, the spread is smaller in the tropics due to improved representation of ENSO variability (Jones et al., 2012). The power spectral density of simulations forced by both natural and anthropogenic forcings encompasses the variability of observational datasets. (Sen Gupta et al., 2012) found that some of the CMIP3 models have local drift magnitudes that are typically between 15 and 35% of the 20th century simulation trend magnitudes for 1950–2000, which could explain part of the model spread at low frequency.

Similar conclusions are drawn from the comparison of spectra estimated from last millennium simulations, performed with a subset of the CMIP5 models, and different northern hemisphere temperature proxy records (Figure 9.33) (see Chapter 5 for details). The simulations include natural and anthropogenic forcings (solar, volcanic, greenhouse gases, land use) (Schmidt et al., 2012). Significant differences between unforced and forced simulations are seen at low frequencies, indicating the role of forced variability (Fernandez-Donado et al., 2012). Lower variability is found in the two MPI-ESM spectra at low frequencies owing to an underestimation of CO_2 variability (Jungclaus et al., 2010). At regional scales, the comparison of a subset of Holocene simulations with instrumental and proxy SST data suggest that models tend to underestimate low-frequency SST variability (Laepple and Huybers, 2012). These lines of evidence suggest with high confidence that models have an adequate representation of variability at the global scale, with larger errors at regional scales and a tendency for underestimation at long time scales.

[INSERT FIGURE 9.33 HERE]

Figure 9.33: Global climate variability represented as (top): Standard deviation of zonal-mean surface temperature of the CMIP5 pre-industrial control simulations (after Jones et al., 2012). (Middle): Power spectral density for 1901–2010 global-mean surface temperature for both historical CMIP5 simulations and the observations (after Jones et al., 2012). (Bottom): Power spectral density for Northern-Hemisphere surface temperature from the CMIP5-PMIP3 last-millennium simulations using common external forcing configurations (colour, (Schmidt et al., 2012), together with the corresponding pre-industrial simulations (colour, dashed), previous last-millennium AOGCM simulations (black, (Fernandez-Donado et al., 2012), and temperature reconstructions from different proxy records (see Section 5.3.5). The small panel included in the bottom panel shows for the different models and reconstructions the percentage of spectral density cumulated for periods above 50 year, so as to better highlight the differences between unforced (pre-industrial control) and forced (PMIP3 and pre-PMIP3) simulations, compared to temperature reconstruction for the longer time scales.

9.5.3.2 Extra-Tropical Circulation, North Atlantic Oscillation and Other Related Dipolar and Annular Modes

Based on CMIP3 models, Gerber et al. (2008) confirmed the AR4 assessment that climate models are able to capture the broad spatial and temporal features of these modes as well as their main inter-hemispheric differences. While models successfully simulate the broad features of the NAO[1], there are substantial differences in the spatial patterns amongst individual models (Miller et al., 2006; Stephenson et al., 2006) especially in non-winter seasons (Stoner et al., 2009; Zhu and Wang, 2010). Climate models have a tendency to overestimate the teleconnection between the Atlantic and Pacific basins, so that patterns of variability tend to be more annular in character than observed (Xin et al., 2008). Models substantially over-estimate persistence on subseasonal and seasonal time scales, particularly during austral spring and summer, and have difficulty simulating the annual cycle of annular mode timescales found in re-analyses for either hemisphere, but with strongest biases in the NAO (Gerber et al., 2008). The unrealistically long timescale of jet variability is worse in models with particularly strong equatorward biases in the mean jet location, a result which has been found to hold in the North Atlantic and in the Southern Hemisphere (Barnes and Hartmann, 2010; Kidston and Gerber, 2010).

As described in the AR4, several climate models have been unable to simulate the observed level of multi-decadal variability in the NAO, in particular variations as strong as the positive trend over the latter half of the 20th century (Gillett, 2005; Stephenson et al., 2006; Stoner et al., 2009). Underestimation of NAO trends can contribute substantially to underprediction of future warming in certain regions (Knutson et al., 2006). Scaife et al. (2009) showed that atmospheric models forced with observed sea surface temperatures, sea-ice and radiative forcings are also unable to simulate the strong NAO trend over the period 1965–1995. However, several coupled climate models do exhibit multi-decadal variability in unforced control simulations which is sometimes as large as the observed 50 year (Raible et al., 2005) and even 30 year trends (Selten et al., 2004; Semenov et al., 2008). Sampling variability may therefore be an explanation for the mismatch, but other explanations have also been suggested, so it is unclear to what extent the underestimation of late 20th century trends reflects real model shortcomings. In addition, the observed NAO trend has weakened in recent years and is no longer significant compared to simulated internal variability (see Figure 10.11). While some studies suggest that greenhouse gas forcing could have played a role in the positive NAO trend (Paeth et al., 2008), model projections have a vertical structure of circulation change which is quite different from the NAO (Woollings, 2008). Scaife et al. (2005) showed that the trend can be reproduced reliably and repeatedly when the upper atmospheric winds are relaxed to the observed trend.

---

[1] The term NAO is used in the widest sense to denote wintertime NAO, AO and NAM unless further distinction is required.

Further evidence has emerged of the coupling of NAO variability between the troposphere and the stratosphere, and even models with improved stratospheric resolution appear to underestimate the vertical coupling (Morgenstern et al., 2010a). Furthermore, the representation of the stratosphere seems to significantly influence the NAO response to anthropogenic forcing (Scaife et al., 2011b), with some studies finding a negative NAO response in stratosphere-resolving Chemistry Climate Models (Morgenstern et al., 2010a). Improved representation of storms in higher-horizontal resolution climate models has also been shown to improve model ability to simulate the NAO (Marti et al., 2010).

The Pacific basin analogue of the NAO, the North Pacific Oscillation (NPO) is a prominent pattern of wintertime atmospheric circulation variability characterized by a north-south dipole in sea level pressure (Linkin and Nigam, 2008; Rogers, 1981; Walker, 1924). Although climate models simulate a realistic NPO spatial pattern under present-day GHG concentrations, many of them are unable to capture the observed linkage with the North Pacific Ocean gyre circulation and most fail to show the observed connection between tropical central Pacific warming phases and the NPO (Furtado et al., 2011a).

There are also considerable biases in the Southern Hemisphere eddy-driven jet stream in the CMIP3 models and these appear to have a direct bearing on the magnitude of the SAM response to forcing (Barnes and Hartmann, 2010; Kidston and Gerber, 2010). In terms of spatial patterns, Raphael and Holland (2006) showed that coupled models produce a clear SAM but that there are relatively large differences between models in terms of the exact shape and orientation of this pattern. Karpechko et al. (2009) found that the CMIP3 models have problems in accurately representing the impacts of the SAM on SST, surface air temperature, precipitation and particularly sea-ice in the Antarctic region.

9.5.3.3 Atlantic Modes

9.5.3.3.1 AMOC variability

Previous comparisons of the observed and simulated AMOC were restricted to its mean strength, as it had only been sporadically observed (see Chapter 3). Continuous AMOC time-series now exist for latitudes 41°N (reconstructions since 1993) and 26°N (estimate based on direct observations since 2004) (Willis, 2010; Cunningham et al., 2010). At 26°N, CMIP3 model simulations show realistic variability for the total AMOC over the available observational record (Baehr et al., 2009; Marsh et al., 2009; Balan Sarojini et al., 2011).

Most AMOC observations (continuous or sporadic) estimate the total AMOC as the sum of a wind-driven component and a mid-ocean geostrophic component. The wind-driven variability appears well represented (though a bit strong in some models); the mid-ocean geostrophic variability is also generally well represented (though a bit weak in some models) (Baehr et al., 2009; Balan Sarojini et al., 2011). The under-representation of the variability of the mid-ocean geostrophic contribution might point to deficiencies in the simulation of hydrographic characteristics (Baehr et al., 2009), which might improve at higher resolution (Marsh et al., 2009).

9.5.3.3.2 Atlantic multi-decadal variability / AMO

The Atlantic Multidecadal Oscillation (AMO) is a mode of climate variability found in the instrumental climate record, with an apparent period of about 70 years, a pattern centred in the North Atlantic Ocean, and near-global reach (see Section 14.2.5). In the AR4, it was shown that a number of unforced climate models produced AMO-like multidecadal variability in the North Atlantic linked to variability in the strength of the AMOC. Subsequent analyses have not changed this picture, with more models showing Atlantic multidecadal variability. Despite this, detailed agreement between models is lacking, with, for example, simulated timescales ranging from 40–60 years (Park and Latif, 2010; Frankcombe et al., 2010), to a century or more (Msadek and Frankignoul, 2009; Menary et al., 2011). In addition, the spatial patterns of variability related to the AMOC differ in many respects from one model to another as shown in Figure 9.34. Models also tend to lack ‘convergence’ in that a model with good AMO characteristics often does not retain this when upgraded to the next version (Hurrell et al., 2010). Recent modelling confirms the link between the AMO and AMOC, but models tend to differ on the mechanisms involved. These include: coupled atmosphere-ocean interactions in the far North Atlantic (Msadek and Frankignoul, 2009), water mass exchange with the Arctic (Frankcombe et al., 2010), and advected tropical salinity feedbacks.

The presence of AMO-like variability in unforced simulations, and the fact that forced 20th century simulations in the CMIP3 multi-model ensemble produce AMO variability that is not in phase with that observed, implies the AMO is not a result of the forcings imposed on the models (Kravtsov and Spannagle, 2008; Knight, 2009; Ting et al., 2009). However, a better reproduction of historical AMO fluctuations has been achieved in a model with a more sophisticated aerosol treatment than was typically used in CMIP3 (Booth et al., 2012). This would suggest that at least part of the AMO may in fact be forced, and that aerosols play a role. Further evidence for this comes from the fact that changes in atmospheric loading of African dust may also be a strong driver of multidecadal temperature variability in the tropical Atlantic (Evan et al., (2009), and could act as a positive feedback on the AMO (Foltz and McPhaden, 2008). In addition to tropospheric aerosols, Otterå et al. (2010) showed the potential for simulated volcanic forcing to have influenced AMO fluctuations over the last 600 years.

[INSERT FIGURE 9.34 HERE]

Figure 9.34: From top to bottom: SST composites using AMOC time series; precipitation composites using cross-equatorial SST difference time series; equatorial salinity composites using ITCZ-strength time series; subpolar-gyre depth-averaged salinity (top 800–1,000 m) using equatorial salinity time series; subpolar gyre depth averaged density using subpolar gyre depth averaged salinity time series. From left to right: HadCM3, MPI-ESM, and KCM. Black outlining signifies areas statistically significant at the 5% level for a two-tailed t test using the moving-blocks bootstrapping technique (Wilks, 1995). Figure 3 from Menary et al. (2011).

9.5.3.3.3 Tropical zonal and meridional modes

Atlantic Meridional Mode (AMM)

The AMM is the dominant mode of internannual variability in the tropical Atlantic in all seasons except for boreal summer, when the Atlantic Niño becomes slightly more prominent. The AMM is characterized by an anomalous meridional shift in the inter-tropical convergence zone (ITCZ) caused by variations in SST and easterly trade winds in the tropical Atlantic (Chiang and Vimont, 2004). Variations in the AMM have been shown to be related to variations in hurricane tracks over the North Atlantic (Xie et al., 2005; Smirnov and Vimont, 2011). Virtually all CMIP models simulate AMM-like SST variability in their 20th century climate simulations. However, most models underestimate the SST variance associated with the AMM, and position the north tropical Atlantic SST anomaly too far equatorward. More problematic is the fact that the development of the AMM in many models is led by a zonal mode during boreal winter—a feature that is not observed in nature (Breugem et al., 2006). This spurious AMM behavior in the models is likely to be associated with the severe model biases in simulating the ITCZ.

Atlantic Niño

CMIP3 models have considerable difficulty simulating Atlantic Niño in their 20th century climate simulations. For many models the so-called ‘Atl-3’ SST index (20°W–0°W, 3°S–3°N) displays the wrong seasonality, with the maximum value in either DJF or SON instead of JJA as in observations (Breugem et al., 2006). Of the two models that capture the observed seasonality, one severely over-estimates the Atl-3 SST variance, while the other severely underestimates it. The models’ inability to capture the observed Atlantic Niño activity is likely caused by mean biases in the region. Almost all of the CMIP3 models fail to simulate a fundamental feature of the equatorial Atlantic Ocean—the east-west equatorial SST gradient and the eastward shoaling thermocline (e.g., Richter and Xie, 2008).

9.5.3.4 Indo-Pacific Modes

9.5.3.4.1 El Niño-Southern Oscillation

The El Niño-Southern Oscillation (ENSO) phenomenon is the dominant mode of climate variability in the tropical Pacific on seasonal to interannual time scales (see Wang and Picaut, 2004; and Chapter 14). The representation of ENSO in AOGCMs has steadily improved and now bears considerable similarity to observed ENSO properties (AchutaRao and Sperber, 2002; Guilyardi et al., 2009b; Randall et al., 2007). However, as was the case in the AR4, simulations of both background climate (time mean and seasonal cycle) and internal variability, exhibit serious systematic errors (Capotondi et al., 2006; Guilyardi, 2006; van Oldenborgh et al., 2005; Wittenberg et al., 2006; Stevenson et al., 2012; Dufresne and co-authors, 2011; Watanabe et al., 2011), many of which can be traced the representation of deep convection, trade wind strength and cloud feedbacks (Braconnot et al., 2007a; Guilyardi et al., 2009a; L'Ecuyer and Stephens, 2007; Lloyd et al., 2010; Lloyd et al., 2009; Sun et al., 2009). Some models have been identified that perform particularly well (e.g., GFDL2.1 in CMIP3 and CNRM-CM5 in CMIP5, (Kakitha et al., 2011)).

While a number of CMIP3 models do not exhibit an ENSO variability maximum at the observed 2–7 year time scale, most CMIP5 models do have a maximum near the observed range and fewer models have the tendency for biannual oscillations (Figure 9.35, see also Stevenson, 2012). In CMIP3 the amplitude of El Niño ranged from less than half to more than double the observed amplitude (AchutaRao and Sperber, 2006; Guilyardi, 2006; Guilyardi et al., 2009b; van Oldenborgh et al., 2005). By contrast, the CMIP5 models show less inter-model spread (Figure 9.36, Kim and Yu, 2012). The observed seasonal amplitude phase locking—El Niño and La Niña anomalies tend to peak in boreal winter and are weakest in boreal spring—is often not captured by models, although some do show a tendency to have the ENSO peak in boreal winter (Kakitha et al., 2011). The CMIP5 models still display the narrow bias in the ENSO-related SST pattern width around the equator relative to the observations, although the bias is slightly reduced when compared to CMIP3. The improvement is partly due to relatively stronger trade winds, and partly due to a relatively longer ENSO periods (Zhang and Jin, 2012). The biases that persist combine to generate errors in ENSO amplitude, period, irregularity, skewness or spatial patterns (Guilyardi et al., 2009b; Leloup et al., 2008). Ohba et al. (2010) separately investigate the simulated transition process of a warm-phase and a cold-phase ENSO in the CMIP3 models. Some of the models reproduce the features of the observed transition process of El Niño/La Niña, whereas most models fail to concurrently reproduce the process during both phases.

Since AR4, new analysis methods have emerged and are now being applied. For example, Jin et al. (2006) and Kim and Jin (2010a) identified five different feedbacks affecting the Bjerknes (or BJ) index, which in turn characterizes ENSO stability. Kim and Jin (2010b) applied this process-based analysis to the CMIP3 multi-model ensemble and demonstrated a significant positive correlation between ENSO amplitude and the BJ index. When respective components of the BJ index obtained from the coupled models were compared with those from observations, it was shown that most coupled models underestimated the negative thermal damping feedback and the positive zonal advective and thermocline feedbacks.

[INSERT FIGURE 9.35 HERE]

Figure 9.35: Maximum entropy power spectra of surface air temperature averaged over the NINO3 region (i.e., 5°N to 5°S, 150°W to 90°W) for (a) the CMIP3 models and (b) the CMIP5 models. Note that ECMWF reanalysis in (a) refers to the European Centre for Medium Range Weather Forecasts (ECMWF) 15-year reanalysis (ERA15). The vertical lines correspond to periods of two and seven years. The power spectra from the reanalyses and for SST from the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) version 1.1, HadCRU 4, ERA40 and NCEP/NCAR data set are given by the series of solid, dashed and/or dotted black curves. Adapted from (AchutaRao and Sperber, 2006).

[INSERT FIGURE 9.36 HERE]

Figure 9.36: ENSO and mean tropical Pacific metrics for pre-industrial control simulations in CMIP3 (blue) and CMIP5 (red). (a) and (b): SST anomaly standard deviation in Niño 3 and Niño 4 (^{o}C), (c) SST annual cycle amplitude in Niño3, (^{o}C), (d) precipitation response (standard deviation) in Niño4 (mm/day), (g) ENSO power spectrum (Niño3) RMS error, (^{o}C). Reference datasets, shown as black solid circles and dashed lines: HadISST1.1 for (a), (b), (c), and (g); CMAP for (d). The CMIP3 and CMIP5 multi-model mean are shown as squares on the left of each panel with the whiskers representing the model standard deviation.

CMIP3 models display a wide range of skill in simulating the interdecadal variability of ENSO (Lin, 2007). The models can be categorized into three groups: those that show an oscillation with a constant period shorter than the observed ENSO period, and sometimes with a constant amplitude; those that do not produce statistically significant peaks in the ENSO frequency band, but usually produces one or two prominent peaks (episodes) at period longer than 6 years; and those that display significant interdecadal variability of ENSO in both amplitude and period. (Yu and Kim, 2011) have further shown that the spatial asymmetries of El Niño and La Niña enable an ENSO–tropical Pacific mean state interaction mechanism that gives rise to a decadal modulation of ENSO intensity in at least three CMIP3 models.

Detailed quantitative evaluation of ENSO performance is hampered by the short observational record (Li et al., 2011b; Wittenberg, 2009; Deser et al., 2011b) and the complexity and diversity of the paradigms and processes involved (Wang and Picaut, 2004). While shortcomings remain (van Oldenborgh et al., 2005; Leloup et al., 2008; Guilyardi et al., 2009b), the CMIP5 model ensemble shows some improvement compared to CMIP3, but there has been no major breakthrough and the improvement is mostly due to a reduced number of poor-performing models.

9.5.3.4.2 Indian Ocean basin and dipole modes

Indian Ocean SST displays a basin-wide warming following El Niño (Klein et al., 1999). This Indian Ocean basin (IOB) mode peaks in boreal spring and persists through the following summer. Recent observational analysis suggests that it is not simply a thermodynamic response to ENSO but involves ocean dynamics and active ocean-atmosphere coupling within the Indian Ocean basin (Du et al., 2009). Only about half of CMIP3 models capture this IOB mode, and the same models tend to simulate ENSO-forced ocean Rossby waves in the tropical south Indian Ocean. In-depth analysis of one model (GFDL CM2.1) confirmed that slow propagation of ocean Rossby waves south of the equator underpins ocean-atmospheric patterns that maintain the IOB through the following summer (Zheng et al., 2011).

The Indian Ocean zonal dipole mode (IOD) (Saji et al., 1999; Webster et al., 1999) appears to be part of a hemispheric response to tropical atmospheric forcing (Fauchereau et al., 2003; Hermes and Reason, 2005). Most CMIP3 models are able to reproduce the general features of the IOD, including its phase lock onto the July-November season (Saji et al., 2006). The modelled SST anomalies, however, tend to show too strong a westward extension along the equator in the eastern Indian Ocean. CMIP3 models exhibit considerable spread in IOD amplitude, some of which can be explained by differences in the strength of the simulated Bjerknes feedback (Liu et al., 2011).

Many models simulate the observed correlation between IOD and ENSO. The magnitude of this correlation varies substantially between models, but is apparently not tied to the amplitude of ENSO (Saji et al., 2006). Models that simulate a deeper thermocline off Sumatra tend to show a larger correlation between IOD and ENSO indices than do models with a shallower thermocline. A subset of CMIP3 models show a spurious correlation with ENSO following the decay of ENSO events, instead of during the ENSO developing phase, possibly due to erroneous representation of oceanic pathways connecting the equatorial Pacific and Indian Oceans (Cai et al., 2011).

9.5.3.4.3 Pacific Decadal Oscillation (PDO) and Interdecadal Pacific Oscillation (IPO)

The PDO refers to a mode of variability involving sea surface temperature (SST) anomalies over the North Pacific (north of 20°N) (Mantua et al., 1997). It exhibits anomalies of one sign along the west coast of North America, and of opposite sign over the western and central North Pacific. Although the PDO time series exhibits considerable decadal variability, it is difficult to ascertain whether there are any robust spectral peaks given the relatively short observational record (Minobe, 1997, 1999; Deser et al., 2004; Pierce, 2001). The ability of climate models to represent the PDO has been assessed by Stoner et al. (2009) and Furtado et al. (2011b). Their results indicate that approximately half of the CMIP3 models simulate a realistic spatial pattern and temporal behaviour (e.g., enhanced variance at low frequencies); however, spectral peaks are consistently higher in frequency than those suggested by the short observational record. The modelled PDO correlations with SST anomalies in the tropical Indo-Pacific are strongly underestimated by the CMIP3 models (Lienert et al., 2011; Deser et al., 2011a; Furtado et al., 2011b; Wang et al., 2010). On the other hand, climate models have been shown to simulate the closely related Interdecadal Pacific Oscillation (IPO, based on SSTs over the entire Pacific basin see Chapter 14, Section 14.2.5) reasonably well (Meehl et al., 2009; Power and Colman, 2006; Power et al., 2006).

9.5.3.4.4 Tropical ocean decadal variability

Pacific Subtropical Cells (STCs) are the shallow meridional cells in which water flows out of the tropics in the surface layer, subducts in the subtropics, flows equatorward in the thermocline and upwells in the equatorial ocean (Blanke and Raynaud, 1997; McCreary and Lu, 1994). The STCs provide a pathway by which extra-tropical atmospheric variability can force tropical variability. Observational studies have shown that these wind driven cells are major drivers of SST change in the tropical Pacific (McPhaden and Zhang, 2002), where a decrease in tropical Pacific SST is significantly correlated with a spin-up of the STCs and an increase in SST with a spin-down. Several studies have shown that this relationship is absent from the CMIP3 historical climate simulations (Zhang and McPhaden, 2006). Hence the impact of a weakening of the Walker Circulation with climate change (Vecchi et al.; Vecchi et al., 2006b) may not be fully accounted for. Solomon and Zhang (2006) suggest that the CMIP3 models may be reproducing the observed *local* ocean response to changes in forcing but inadequately reproduce the *remote* STC-forcing of the tropical Pacific due to the underestimate of extratropical winds that force these ocean circulations.

9.5.3.5 Teleconnections

In general terms, teleconnections characterize the response of the climate system in one location to forcings in another. SST variability provides a significant forcing of atmospheric teleconnection response and drives a significant portion of the climate variability over land (Goddard and Mason, 2002; Shin et al., 2010). Although local forcings and feedbacks can play an important role (Pitman et al., 2010), the simulation of land surface temperatures and precipitation requires accurate predictions of SST patterns (Compo and Sardeshmukh, 2009; Shin et al., 2010).

9.5.3.5.1 Teleconnections affecting North America

The Pacific North American (PNA) pattern is a wavetrain-like pattern in mid-level geopotential heights. The majority of CMIP3 models simulate a realistic spatial structure of the PNA pattern in wintertime (Stoner et al., 2009). The PNA pattern has contributions from both internal atmospheric variability (Johnson and Feldstein, 2010) and ENSO and PDO teleconnections (Deser et al., 2004). The power spectrum of this temporal behavior is generally captured by the CMIP3 models, although the level of year-to-year autocorrelation varies according to the strength of the simulated ENSO and PDO (Stoner et al., 2009). Surface air temperature and precipitation patterns are more challenging, and the quality of CMIP5 teleconnections for these is mixed. Wintertime surface air temperature anomaly patterns associated with El Niño events are qualitatively captured in the CMIP5 multi-model mean (Sheffield and al., 2012) although the amplitude is weak. For precipitation, ENSO teleconnection pattern correlations over the southern United States exhibit no significant improvement in the CMIP5 models relative to CMIP3 (Langenbrunner and Neelin, 2012), but models do correctly reproduce the sign of the observed pattern. The CMIP5 models do well at capturing the surface air temperature pattern over North America associated with the PDO (Sheffield and al., 2012).

9.5.3.5.2 Tropical ENSO teleconnections

These moist teleconnection pathways, in which a baroclinic tropospheric warming signal propagates within convection zones, involve mechanisms related to those at play in the precipitation response to global warming (Chiang and Sobel, 2002; Neelin et al., 2003) and provide challenging test statistics for model precipitation response. Compared to earlier generation climate models, CMIP3 and CMIP5 models tend to do somewhat better (Cai et al., 2009; Langenbrunner and Neelin, 2012) (Coelho and Goddard, 2009; Neelin, 2007) at precipitation reductions associated with El Niño over equatorial South America and the Western Pacific, although CMIP5 offers little further improvement over CMIP3 (see for instance the standard deviation of preciptation in the western Pacific in Figure 9.36). CMIP5 models match well to observations in simulating the sign of the precipitation change over broad regions, and do well at predicting the amplitude of the change (for a given SST forcing) provided this is assessed as the mean of the amplitudes from the individual models and not the amplitude of the multi-model mean (Langenbrunner and Neelin, 2012; Langenbrunner and Neelin 2012). Teleconnection patterns from both ENSO and the Indian Ocean Dipole to precipitation over Australia are reasonably well simulated in the key September-November season (Cai et al., 2009) in the CMIP3 multi-model model.

A regression of the West African monsoon precipitation index with global SSTs reveals two major teleconnections (Fontaine and Janicot, 1996). The first highlights the strong influence of ENSO, while the second reveals a relationship between the SST in the Gulf of Guinea and the northward migration of the monsoon rainbelt over West Africa. Most CMIP3 models show a single dominant Pacific teleconnection, which is, however, of the wrong sign for half of the models (Joly et al., 2007). Only one model shows a significant second mode, emphasizing the difficulty in simulating the response of the African rainbelt to Atlantic SST anomalies that are not synchronous with Pacific anomalies.

Both CMIP3 and CMIP5 models have been evaluated and found to vary in their abilities to represent both the seasonal cycle of correlations between the Niño 3.4 and North Australian SSTs, and the evolution of SSTs during composite El Niño and La Niña situations (Catto et al., 2012a, 2012b) with little change in quality from CMIP3 to CMIP5. Generally the models do not capture the strength of the negative correlations during the second half of the year. The models also still struggle to capture the SST evolution in the North Australian region during El Niño and La Niña.

9.5.3.5.3 The Quasi-Biennial Oscillation (QBO)

Significant progress has been made in recent years to model and understand the impacts of the QBO (Baldwin et al., 2001). Many climate models have now increased their vertical domain, and some of these reproduce a QBO (e.g., HadGEM2, MPI-ESM-LR, MIROC). To help simulate the small scale waves which drive the QBO some models have employed reduced diffusion or high vertical or horizontal resolution (Takahashi, 1999; Kawatani et al., 2011 respectively), while others use parameterised wave spectra to circumvent the need for such high resolution (Scaife et al., 2000; Giorgetta et al., 2002; Giorgetta, 2006; McLandress, 2002). These model results are consistent with recent observations which confirm that small scale gravity waves carry a large proportion of the momentum flux which drives the QBO (Sato and Dunkerton, 1997; Ern and Preusse, 2009). Many features of the QBO such as its width and phase asymmetry also appear spontaneously in these simulations due to internal dynamics (Dunkerton, 1991; Scaife et al., 2002; Haynes, 2006). Some of the QBO effects on the extratropical climate (Holton and Tan, 1980; Hamilton, 1998; Naoe and Shibata, 2010) as well as ozone (Butchart et al., 2003; Shibata and Deushi, 2005) are also reproduced in models. Subsequent influences on the Arctic/North Atlantic Oscillation have also been suggested from observational and modelling studies (Thompson et al., 2002; Boer and Hamilton, 2008; Marshall and Scaife, 2009).

9.5.3.6 Carbon Cycle Variability

The two coupled biogeochemistry/land component models evaluated by Randerson et al. (2009) reproduce the interannual variability in land fluxes during 1988–2004 when compared against Atmospheric Tracer Transport Model Intercomparison Project (TRANSCOM). The models are significantly and positively correlated with the time series of annual-mean fluxes and explain between 43% and 53% of the fluctuations in TRANSCOM. The models produce year-to-year variability that agrees to within 30% with the interannual standard deviation from TRANSCOM of 1.0 PgC yr^{-1}. Over the longer time period spanning 1860 to 2002, the inclusion of nitrogen cycling and deposition on global carbon sequestration accounts for less than 20% of recent changes in annual NPP due to atmospheric composition and climate (Zaehle et al., 2010b).

When these components are linked to fully coupled Earth system models, these models tend to overestimate the long-term trend in global-mean atmospheric CO$_2$ concentrations (Cadule et al., 2010). The quality of the simulation of various types of interannual variability, including the oscillations in CO$_2$ associated with volcanic eruptions and the positive and negative phases of ENSO (Cox et al., submitted), and long-term trends in the seasonal amplitude (Cadule et al., 2010), vary significantly amongst models.

9.5.3.7 Summary

In summary, the assessment of interannual to interdecadal variability in climate models presents a varied picture. CMIP5 models show a modest improvement over CMIP3. New since the AR4, process-based model evaluation is now helping identify sources of specific biases, although the observational record is sometime too short or innacurate to offer strong enough constraints. The assessment in reproducing the modes and patterns discussed in this section is summarised in Table 9.3.

Table 9.3: Summary of assessment of interannual to interdecadal variability in climate models. See also Figure 9.45.

EG	Short Name	Level of Confidence	Level of Evidence for Evaluation	Level of Agreement	Model Quality	Difference with AR4 (including CMIP5 vs. CMIP3)	Panel	Section
Global SST variability	SST-var	High	Robust	Medium	Medium	Slight improvement in the tropics	B	9.5.3.1
North Atlantic Oscillation and Northern annular mode	NAO	Medium	Medium	Medium	High	No assessment	B	9.5.3.2
Southern Annual Mode	SAM	Low	Limited	Medium	Medium	No assessment	B	9.5.3.2
Atlantic Meridional	AMOC	Low	Limited	Medium	Medium	No assessment	B	9.5.3.3

Overtuning Circulation								
Atlantic Multi-decadal Variability	AMV	Low	Limited	Medium	Medium	No assessment	B	9.5.3.3
Atlantic Meridional Mode	AMM	High	Medium	High	Low	No assessment	B	9.5.3.3
Atlantic Niño	AN	Low	Limited	Medium	Low	No assessment	B	9.5.3.3
El Niño Southern Oscillation	ENSO	High	Medium	High	Medium	Inter-model amplitude diversity reduced and peak frequency improved	B	9.5.3.4
Indian ocean dipole	IOD	Medium	Medium	Medium	Medium	No assessment	B	9.5.3.4
Quasi-Biennal Oscillation	QBO	Medium	Medium	Medium	High	No assessment	B	9.5.3.5
Pacific Decadal Oscillation	PDO	Low	Limited	Medium	Medium	No assessment	B	9.5.3.4
Interdecadal Pacific Oscillation	IPO	Low	Limited	Medium	High	No assessment	B	9.5.3.4
Pacific North American	PNA	High	Medium	High	Medium	No assessment	B	9.5.3.5
Tropical ENSO teleconnections	ENSOtele	High	Robust	Medium	Medium	No change	B	9.5.3.5

9.5.4 *Extreme Events*

Extreme events are realizations of the tail of the probability distribution of weather and climate variability. They are higher-order statistics and thus generally expected to be more difficult to realistically represent in climate models. Shorter time scale extreme events are often associated with smaller scale spatial structure, which cannot be captured by coarse resolution models but may be better represented as model resolution increases. In the AR4, it was concluded that models could simulate the statistics of extreme events better than expected from generally coarse resolutions of the models at that time, especially for temperature extremes (Randall et al., 2007).

The IPCC has conducted an assessment of extreme events in the context of climate change -- the Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX) (IPCC, 2012). Although there is no comprehensive climate model evaluation with respect to extreme events in SREX, there is some consideration of model performance taken into account in assessing uncertainties in projections.

9.5.4.1 Extreme Temperature

Since the AR4, evaluation of CMIP3 and CMIP5 models has been undertaken with respect to extreme events. Both model ensembles simulate present-day warm extremes in terms of 20-year return values reasonably well on the global scale, as compared to reanalyses (Kharin et al., 2007; Kharin et al., 2012). The CMIP5 models perform comparably to the CMIP3 for various temperature extreme indices, such as annual maximum daily maximum surface air temperature, but with smaller inter-model spread in CMIP5 (Sillmann et al., 2012). Figure 9.37 shows relative error estimates of available CMIP5 models for various extreme indices over global land based on Sillmann et al. (2012). While the relative performance of an individual model may depend on the choice of the reference dataset (four different reanalyses), the mean and median models tend to outperform individual models with respect to all the reanalyses used. According to the standardized multi-model median errors ($\mathrm{RMSE}_{\mathrm{std}}$) for CMIP3 and CMIP5 ensembles shown on the right side of the figure, the CMIP5 ensemble performs slightly better (indicated by lighter gray shading) than CMIP3 for some indices and reference datasets.

In terms of historical trends, CMIP3 and CMIP5 models generally capture observed trends in temperature extremes in the second half of the 20th century (Sillmann et al., 2012), as illustrated in Figure 9.37. The modelled trends are consistent with both reanalyes and station-based estimates. It is also clear in the figure that model-based indices respond coherently to major volcanic eruptions. Interestingly, detection and attribution studies based on CMIP3 models suggest that models tend to overestimate the observed warming of warm temperature extremes and underestimate the warming of cold extremes in the second half of 20th century (Christidis et al., 2011; Zwiers et al., 2011) as noted in SREX (Seneviratne et al., 2012) See also Chapter 10 -- a result at odds with the global mean results in Figure 9.37.

[INSERT FIGURE 9.37 HERE]

Figure 9.37: (a) Portrait plot display of relative error metrics for the CMIP5 temperature and precipitation extreme indices based on Sillmann et al. (2012). Reddish and bluish colours indicate, respectively, larger and smaller RMS errors for an individual model relative to the median model. The gray-shaded columns on the right side indicate the RMS error for the multi-model median standardized by the spatial standard deviation of the index climatology in the reanalysis, representing absolute errors for CMIP3 and CMIP5 ensembles. Results for four different reference datasets, ERA-interim (top), ERA40 (left), NCEP/NCAR (right) and NCEP-DOE (bottom) reanalyses, are shown in each box. The analysis period is 1981–2000 and only land areas are considered. The indices shown are simple daily precipitation intensity index (SDII), consecutive dry days (CDD), annual maximum 5-day/1-day precipitation (RX5day/RX1day), tropical nights (TR), frost days (FD), annual minimum/maximum daily maximum surface air temperature (TXn/TXx) and annual minimum/maximum daily minimum surface air temperature (TNn/TNx). Note that only a small selection of the indices analysed in Sillmann et al. (2012) is shown, preferentially those that appear in other Chapters (2, 10, 11, 12, 14). (b)-(e) Time series of global mean temperature extreme indices over land from 1948 to 2010 for CMIP3 (green) and CMIP5 (black) models, ERA40 (blue) and NCEP/NCAR (cyan) reanalyses and HadEX station-based observational dataset (red) based on Sillmann et al. (2012). Shading for model results indicates the 25th to 75th quantile range of inter-model spread. Grey shading along the horizontal axis indicates the evolution of globally averaged volcanic forcing according to Sato et al. (1993). The indices shown are the frequency of daily minimum/maximum surface air temperature below 10th percentile (b: Cold nights/c: Cold days) and that above 90th percentile (d: Warm nights/e: Warm days) of 1961–1990 base period. Note that, as these indices essentially represent changes relative to the base period, they are particularly suitable for being shown in time series and not straightforward for being shown in (a).

9.5.4.2 Extreme Precipitation

For extreme precipitation, uncertainty in observational data is much larger than for temperature, which makes model evaluation more challenging. Discrepancies between different reanalyses for extreme precipitation are substantial, while station-based observations have limited spatial coverage (Kharin et al., 2007; Kharin et al., 2012; Sillmann et al., 2012). Moreover, a station-based observational dataset, which is an interpolated field from station measurements, has a potential mismatch of spatial scale when compared to model results or reanalyses (Chen and Knutson, 2008). The uncertainties are especially large in the tropics. In the extratropics, precipitation extremes in terms of 20-year return values simulated by CMIP3 and CMIP5 models compared relatively well with the observational datasets (Kharin et al., 2007; Kharin et al., 2012). Figure 9.37, shows relative errors of CMIP5 models for four precipitation-related indices. Darker gray shadings in the RMSE columns for precipitation indicate larger discrepancies between models and reanalyses for precipitation extremes in general. Sillmann et al. (2012) found that the CMIP5 models tend to simulate more intense precipitation and fewer consecutive dry days than the CMIP3, and thus are closer to the observations as represented by the HadEX indices. This improvement could in part be due to generally higher spatial resolution of CMIP5 models compared to CMIP3 (Sillmann et al., 2012).

It is known from sensitivity studies that simulated extreme precipitation is strongly resolution dependent. Growing evidence has shown that high-resolution models (50 km or finer atmospheric horizontal resolution) can capture the intensity of extreme precipitation fairly realistically (Kawazoe and Gutowski, 2012; Sakamoto et al., 2012; Wehner et al., 2010), though some of these results are based on models with observationally-constrained surface or lateral boundary conditions (i.e., AGCMs or RCMs).

In terms of historical trends, a D & A study by Min et al. (2011) found consistency in sign between the observed increase in heavy precipitation over Northern Hemisphere land areas in the second half of the 20th century and that simulated by CMIP3 models, but they found that the models tend to underestimate the observed trend (see also Chapter 10). Related to this, it has been pointed out from comparisons between CMIP3 models and satellite-based datasets that models underestimate the temperature dependence of precipitation intensity over the tropical ocean (Allan and Soden, 2008) and globally (Liu et al., 2009). It is currently unclear whether CMIP5 models also have the same tendency.

9.5.4.3 Other Extremes

With regard to tropical cyclones, it was concluded in the AR4 that high-resolution AGCMs produced generally good simulation of their frequency and distribution, but underestimated their intensity (Randall et al., 2007). Since then, Mizuta et al. (2011) have shown that a newer version of the MRI-AGCM with improved parameterizations but with the same resolution as the previous version (20 km) simulates tropical cyclones as intense as those observed with improved distribution as well. This implies that 20 km atmospheric resolution might be enough to simulate realistic intensity of tropical cyclones if adequate parameterizations are used. Another remarkable finding since AR4 is that the observed year-to-year counts of Atlantic hurricanes can be well simulated by AGCMs driven only by observed sea surface temperature (Larow et al., 2008; Zhao et al., 2009).

One of the important extreme events at longer timescale (months or longer) is drought, which is caused by variability of both precipitation and evaporation. Sheffield and Wood (2008) found that models in the CMIP3 ensemble simulated large-scale droughts in 20th century reasonably well. However, it should be noted that there are various definitions of drought (see Chapter 2) and the performance of simulated drought might depend on the definition. Moreover, different models can simulate drought with different mechanisms (McCrary and Randall, 2010). A comprehensive evaluation of CMIP5 models for drought is currently not available except that Sillmann et al. (2012) found that consecutive dry days simulated by CMIP5 models agree fairly well with HadEX.

9.5.4.4 Summary

There is medium evidence (i.e., a few multi-model studies) and high agreement that global distribution of temperature extremes are represented well by CMIP3 and CMIP5 models. The observed warming trend of temperature extremes in the second half of the 20th century is well captured in models, but there is medium evidence (a few CMIP3 studies) and medium agreement (not evident in a preliminary look at CMIP5) that models tend to overestimate the warming of warm temperature extremes and underestimate the warming of cold temperature extremes. There is medium evidence (single multi-model study) and medium agreement (as inter-model difference is large) that CMIP5 models tend to simulate more intense and thus more realistic precipitation extremes than CMIP3, which could be partly due to generally higher horizontal resolution in CMIP5 than in CMIP3. There are medium but different lines of evidence and high agreement that CMIP3 models tend to underestimate the observed increase in heavy precipitation over the northern hemisphere land areas in the second half of 20th century. There is medium evidence and high agreement that higher resolution models tend to simulate more realistic intensity of extreme precipitation. There is medium evidence, though a limited number of models, and medium agreement, that high-resolution (~20km) AGCMs can simulate realistic distribution and intensity of tropical cyclones and year-to-year count of Atlantic hurricanes. Finally, there is medium evidence (a few multi-model studies) and medium agreement (as it might depend on definitions of drought) that models can simulate drought reasonably well. It should be noted that analysis of CMIP5 models is still limited and the performance of CMIP5 models to simulate observed trends of extreme events is particularly unclear at present.

9.6 Downscaling and Simulation of Regional-Scale Climate

In the above sections, climate model evaluation is discussed in terms of the simulation of different components of the climate system as well as climate averages, variability and extremes. This section addresses climate model evaluation for geographical regions, which both complements the previous sections and provides an assessment of model quality related to regional-scale climate projections in climate change impact research.

Regional-scale climate information can be extracted from the AOGCMs, but the horizontal resolution of the AOGCMs used for centennial simulations may be too low to resolve features that are important at regional scales. To overcome this, high resolution AGCMs, variable-resolution AOGCMs, statistical or dynamical downscaling (i.e., regional climate modelling) are used to generate region-specific climate information.

9.6.1 Global Models

9.6.1.1 Regional-Scale Simulation by AOGCMs

The general performance of the CMIP5 models in terms of annual mean temperature and precipitation was illustrated in section 9.4. Evaluation of how AOGCMs capture observed regional trends" is contained in Box 11.2 in Chapter 11. Here we focus on specific regions and show, in Figure 9.38, a comparison of the seasonal cycles of temperature and precipitation for different regions. These results show, first, that temperature is generally better simulated than precipitation in terms of the amplitude and phase of the seasonal cycle, although precipitation is also well simulated in many models and regions. Second, as was the case for larger-scale quantities, the multi-model mean is closer to observations than most of the individual models. Third, the systematic difference between the CMIP5 and CMIP3 ensembles is small in most cases compared to the inter-model variation. In the monsoon regions of South Asia and Amazon, CMIP3 models often severely underestimated precipitation in the rainy seasons. These errors are often reduced in CMIP5 models. The phase of the annual cycles of temperature and precipitation are also, on average, improved in CMIP5 over CMIP3 models.

[INSERT FIGURE 9.38 HERE]

Figure 9.38: Mean annual cycle of (a) temperature (ºC) and (b) precipitation (mm day^{-1}). The average is taken over land areas within the indicated rectangles, and over the period 1980–1999. The red thick line is the average over 37 CMIP5 models (red thin lines), the green thick line is the average over 22 CMIP3 models (green thin lines) and the black thick line is the CRU TS3.10 observational data for temperature and CRU TS3.10.1 for precipitation. Note that some of the sub-plots have different axis-ranges.

Regional biases in seasonal and annual temperature and precipitation are shown for the SREX land regions (cf. Seneviratne et al., 2012) in Figure 9.39, and for additional regions shown in the Atlas (Annex I) in Figure 9.40. CMIP5 median biases in temperature range from about –3°C to 1.5°C. Substantial cold biases over Northern Hemisphere regions are more prevalent in winter (December-February) than summer (June-August). However, the median biases are in most cases slightly less negative for CMIP5 than CMIP3. The spread amongst models, as characterized by the 25–75% and 5–95% ranges (and by the standard deviation; not shown) has been slightly reduced from CMIP3 to CMIP5 in a majority of the regions. Still, the inter-model spread remains large, particularly in high-latitude regions in winter, and in regions with steep orography (such as CAS, SAS, TIB and WSA—see definitions in Figure caption). The inter-model temperature spread has decreased from CMIP3 to CMIP5 over most of the oceans and over the Arctic and Antarctic land regions, and the cold winter bias over the Arctic has been reduced (Figure 9.40). There is little systematic inter-ensemble difference in temperature over lower latitude oceans, except for a larger cold bias in CMIP5 than CMIP3 over the equatorial Tropical Pacific in June-August.

Biases in precipitation are shown in in the right column of Figures 9.39 and 9.40 for the Northern Hemisphere winter (October to March) and summer (April to September) half years as well as the annual mean. The largest systematic biases over land regions occur in ALA, WSA and TIB, where the annual precipitation exceeds the CRU TS 3.10.01 analysis in all CMIP5 models, with a median bias on the order of 100%. All these regions are characterized by high orography and / or a large fraction of solid precipitation, both of which are likely to introduce a negative bias in gauge-based precipitation (Adam et al., 2006; Yang and Ohata, 2001), and may artificially amplify the model-observation discrepancy. A large negative relative bias in SAH occurs in October-March, but is of negligible magnitude in absolute terms. In nearly all other seasonal and regional cases over land, the observational estimate falls within the range of the CMIP5 simulations. Compared with CMIP3, the CMIP5 median precipitation is slightly higher in most regions; however, there is no systematic change in agreement with observations between the two ensembles. The inter-model spread is similar between CMIP5 and CMIP3, being typically largest in arid areas when expressed in relative terms.

Over the oceans and polar regions, observational uncertainty complicates the evaluation of simulated precipitation. Of two commonly used datasets, CMAP indicates systematically more precipitation than GPCP over low-latitude oceans, but less precipitation in high latitudes (Yin et al., 2004; Shin et al., 2011; Figure 9.40). Over most low-latitude ocean regions, annual precipitation in most CMIP3 and CMIP5 models exceeds GPCP, while the difference from CMAP is smaller although mostly of the same sign. In Arctic and

Antarctic sea areas, simulated precipitation is much above CMAP and somewhat more similar to GPCP. Over Antarctic land, precipitation in most CMIP3 and CMIP5 models is slightly below CMAP and far below GPCP. In contrast, Maris et al. (2012) found a general overestimation of GCM-simulated Antarctic land precipitation, especially inland. Underestimation was seen closer to the coasts and the western side of the Antarctic Peninsula. They used, however, a different observational reference; a regional climate simulation driven by the ERA-Interm reanalysis.

[INSERT FIGURE 9.39 HERE]

Figure 9.39: Seasonal and annual mean biases of (left) temperature (in °C) and (right) precipitation (in %) in the SREX land regions (cf. Seneviratne et al., 2012, page 12. The region's coordinates can be found from their online Appendix 3.A). The 5th, 25th, 50th, 75th and 95th percentiles of the biases in 34 CMIP5 models are given in box-whisker format, and the corresponding values for 23 CMIP3 models with crosses, as indicated in the legend in the top-right panel. The CMIP3 models' 20C3M simulations are complemented with the corresponding A1B runs for the 2001–2005 period. The biases are calculated over the period 1986–2005, using CRU T3.10 as the reference for temperature and CRU TS 3.10.01 for precipitation. The regions are: Alaska/NW Canada (ALA), Eastern Canada/Greenland/Iceland (CGI), Western North America (WNA), Central North America (CNA), Eastern North America (ENA), Central America/Mexico (CAM), Amazon (AMZ), NE Brazil (NEB), West Coast South America (WSA), South-Eastern South America (SSA), Northern Europe (NEU), Central Europe (CEU), Southern Europe/the Mediterranean (MED), Sahara (SAH), Western Africa (WAF), Eastern Africa (EAF), Southern Africa (SAF), Northern Asia (NAS), Western Asia (WAS), Central Asia (CAS), Tibetan Plateau (TIB), Eastern Asia (EAS), Southern Asia (SAS), South-Eastern Asia (SEA), Northern Australia (NAS) and Southern Australia/New Zealand (SAU). Note that the region WSA is poorly resolved in the models.

[INSERT FIGURE 9.40 HERE]

Figure 9.40: As Figure 9.39, but for various polar and ocean regions, and using ERA Interim reanalysis as the reference for temperature and GPCPCMAP for precipitation. Global land, ocean and overall means are also shown. The regions shown are defined as; Arctic: 67.5-90°N, Caribbean: 10°N–25°N, 85°W–60°W, West Indian Ocean: 25°S–5°N, 52°E–75°E; North Indian Ocean: 5°N–30°N, 60°E–95°E; Northern Tropical Pacific: 5°N–25°N, 155°E–150°W; Equatorial Tropical Pacific: 5°S–5°N, 155°E–130°W; Southern Tropical Pacific: 5°S–25°S, 155°E–130°W; Antarctic: 50°S–90°S. As an indicator of observational uncertainty, the normalised difference between CMAP and GPCP precipitation is shown with dotted lines.

Continental to sub-continental mean values frequently hide smaller-scale biases; therefore, biases generally increase in magnitude with decreasing spatial averaging (Masson and Knutti, 2011a; Raisanen and Ylhaisi, 2011). A typical order of magnitude for grid-box-scale annual mean biases in individual CMIP3 models was 2°C for temperature and 1 mm day^{-1} for precipitation (Masson and Knutti, 2011a; Raisanen, 2007), although this was geographically variable and the annual mean biases occasionally represented a balance between larger but compensating biases in individual seasons. The importance of local present-day biases for projection of climate change remains difficult to assess, both because the spread in modelled 21st century climate change is less than the spread in present-day biases (Raisanen, 2007) and because the simulated climate change typically shows only a weak relationship to the simulated present-day climate (Raisanen et al., 2010; Whetton et al., 2007b). Exceptions are areas near the winter sea ice edge, where inter-model variations in present-day ice conditions and temperature biases are linked via local sea ice feedbacks (Bracegirdle and Stephenson, 2012a).

Another issue related to the use of AOGCM data in climate projection concerns the need of spatial averaging or smoothing. Räisänen and Ylhäisi (2011) suggest that explicit smoothing is unnecessary for multi-model mean temperature and precipitation projections, essentially because the most unreliable small-scale features are filtered out in averaging over multiple models. Working from another line of argument, 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. Both studies focus on the direct use of AOGCM data and do not assess the potential for obtaining additional information from downscaling methods. The results of a particular AOGCM for a particular region are also affected both by how the model describes processes specific to the region in question and by the large-scale performance of the model. For example, (van Haren et al., 2012) showed that the failure of CMIP3 models to capture recent precipitation trends in Europe was linked to their failure to simulate the trends in atmospheric circulation in winter and SST in summer. Similar results have been reported earlier (e.g., van Oldenborgh et al., 2009).

On the whole, the CMIP5 models simulate regional-scale temperature somewhat better than the CMIP3 models did. For precipitation, there is no clear evidence of either improvement or deterioration. This is based on an analysis of both ensemble means and intermodal spread. There are as yet rather few published studies in which regional behaviour of the CMIP5 models is evaluated in greater detail, although Cattiaux et al. (Cattiaux et al., 2012) recently noted seasonal temperature biases for Europe similar to those discussed above. Overall, the finding is assessed to be one of high agreement and medium evidence.

9.6.1.1.1 Net precipitation over Antarctica

Net precipitation (P-E) over Antarctica is of particular importance for future sea-level rise as any net water storage on the ice sheet effectively withdraws water from the ocean, and is therefore discussed separately here. Detailed analyses of 15 CMIP3 models (Uotila et al., 2007) suggested that most AOGCMs reproduce late 20th century Antarctic P-E quite well, with a multi-model area mean of 184 mm yr^{-1} over the period 1979–2000, as compared to observationally-based range of 150–190 mm yr^{-1}. However, the range between models was large, from 123 to 269 mm yr^{-1}. A 5-model subset of these models chosen on the basis of their ability to reproduce the observed near surface circulation and P-E over Antarctica, yielded a mean of 171 mm/yr. Within the whole ensemble, as well as the 5-model subset, some models indicated a positive and some a negative trend in the Antarctic P-E over the period 1979-2000. Bromwich et al. (2011) reported on discrepancies between reanalysis data sets for the period 1989-2009), and so a trend and its magnitude remain uncertain in observationally-based data estimates as well. Connolley and Bracegirdle (2007) looked at Antarctic surface mass balance from 19 CMIP3 models' 20th century simulations and concluded the ensemble mean exhibits an overestimate of 30 mm yr^{-1}. A 9-model subset had a bias of 15 mm yr^{-1}, but individual models had substantial positive or negative biases.

The studies above suggest that there is medium confidence that CMIP3 models simulate realistic mean P-E over Antarctica, albeit with values around the high end of those observed and substantial intermodal spread. Further analysis of CMIP5 model results will be required to determine if this is a robust result. There is low agreement, in both models and observations, regarding a trend in Antarctic P-E.

9.6.1.2 Regional-Scale Simulation by AGCMs

Stand-alone global atmospheric models (AGCMs) run at higher resolution than AOGCMs can provide complementary regional-scale climate information. This is sometimes referred to as 'global downscaling'. One example is the simulation of tropical cyclones, which is generally difficult for both AOGCMs and limited-domain RCMs. High-resolution AGCMs have been successfully utilized for this purpose (e.g., Murakami and Sugi, 2010; Murakami et al., 2012; Zhao et al., 2012; Zhao et al., 2009). A number of advantages of high-resolution AGCMs have been identified, including improved regional precipitation (Kusunoki et al., 2011; Zhao et al., 2009) and blocking (Matsueda et al., 2009a; Matsueda et al., 2010a). As in lower-resolution models, performance is affected not only by resolution but also by the quality of physical parameterizations (Lin et al., 2012b; Mizuta et al., 2012; Zhao et al., 2012). Also, the fact that AGCMs do not simulate coupled interactions may limit their ability to represent some high-resolution phenomena (e.g., Hasegawa and Emori, 2007; Zhou et al., 2009).

9.6.1.3 Regional-Scale Simulation by Variable-Resolution GCMs

An alternative to global high-resolution is a global model with variable resolution (or 'stretched grid') with higher resolution over some region of interest. Abiodun et al. (2011) showed that global simulations with such a model improve the simulation of West African monsoon systems and African easterly jets. Fox-Rabinovitz et al. (2008) showed that regional biases in the high-resolution portion of a stretched grid model were similar to that of a global model with the same high resolution everywhere. Markovic et al. (2010) and Déqué (2010) reported similar results, suggesting that, although not widely used, such methods are feasible.

9.6.2 Regional Climate Downscaling

An alternative to global models is a regional climate model (RCM) applied over a limited-area domain and provided with boundary conditions either from observationally-based reanalyses or AOGCM output. For even more localized information, statistical methods can be used, taking advantage of relationships between larger-scale meteorology and local conditions. Such methods are discussed in this section, however results are difficult to generalise as the studies often concern different regions, observational data and periods.

9.6.2.1 Recent Developments of Statistical Methods

Satistical downscaling (SD) involves deriving empirical relationships linking large-scale atmospheric variables (predictors) and local/regional climate variables (predictands). These relationships may then be applied to equivalent predictors from AOGCM projections in order to downscale future climate scenarios. SD methods have also been applied to RCM output (e.g., Boe et al., 2007; Deque, 2007; Paeth, 2011; Segui et al., 2010; van Vliet et al., 2011). The stationarity hypothesis underlies the use of SD, i.e. that the established present-day relationships also apply for the future. Most SD applications use so-called Perfect Prognosis methods in which relationships between predictors and predictands are established using observations. The availability of sufficiently long observational data sets can be a limiting factor in regions with sparse observations.

The development of SD since the AR4 has been quite vigorous (e.g., Fowler et al., 2007; Maraun et al., 2010b), and many state-of-the-art SD approaches now combine different methods , 2011; Vrac and Naveau, 2008). There is also an increasing number of studies that focus on extremes, making use of extreme value theory (e.g., Vrac and Naveau, 2008; Wang and Zhang, 2008), and on features not explicitly represented in global models such as hurricanes (Emanuel et al., 2008), river flow and discharge, sediment, soil erosion and crop yields (e.g., Lewis and Lamoureux, 2010; Prudhomme and Davies, 2009; Zhang, 2007). Techniques have also been developed to consider multiple climatic variables simultaneously in order to preserve some physical consistency (e.g., Zhang and Georgakakos, 2011).

The methods used to evaluate SD approaches vary with the downscaled variable and include metrics related to intensities (e.g., Ning et al., 2011; Tryhorn and DeGaetano, 2011), temporal behaviour (e.g.,Brands et al., 2011; Maraun et al., 2010a; May, 2007; Timbal and Jones, 2008), and physical processes (Lenderink and Van Meijgaard, 2008; Maraun et al., 2010a). SD capabilities are also examined through secondary variables like river discharge and stream flow, which pertains to coherency between variables and/or their spatial autocorrelation (e.g., Boe et al., 2007; Teutschbein et al., 2011).

In general, SD approaches provide a valuable addition to the suite of tools available to provide local and regional climate information, but owing to the variety of schemes, the very specific applications, and the lack of targeted intercomparisons, it is not possible to provide an overall evaluation of SD at this point.

9.6.2.2 Recent Developments of Dynamical Methods

As for AOGCMs, the resolution of RCMs has increased since the AR4, and they have seen further development regarding process-descriptions, new components, and coordinated experimentation based on multi-model ensembles (Laprise, 2008; Rummukainen, 2010).

Much of the motivation for employing RCMs is that they allow more detailed representation of local processes, topography, and other features that shape regional climate. Higher resolution may also better capture other details such as extremes (Seneviratne et al., 2012). For example, Kawazoe and Gutowski (2012) compared six RCMs over the US, two GCMs providing boundary conditions, and high resolution gridded observations and concluded that the RCMs provided better precipitation extremes than the GCMs. Vautard et al. (2012) found that warm extremes were generally better simulated in a set of RCMs run for Europe at 12km resolution compared to 50km runs.

Since the AR4, the typical regional climate model resolution has increased from around 50 km to around 25 km (e.g., Christensen et al., 2010). This has led to some improvements, but increased resolution does not by itself guarantee a better model. For example, Woollings et al. (2010b) investigated the effect of different resolution of Atlantic SST as boundary conditions in an RCM. While higher spatial resolution improved the simulation of storm tracks, higher temporal resolution led to some degradation. Rojas (2006) found a non-linear relationship between improved simulation quality and increasing resolution with more improvement when increasing RCM resolution from 135 km to 45 km that when going from 45 km to 15 km. This was despite the highly variable regional orography in the domain considered. Walther et al. (2011) found that the peak timing and amplitude of the diurnal precipitation cycle, as well as the frequency of light precipitation, improved more when going from 12 km to the 6 km resolution than when going from 50 to 25 km or from 25 to 12 km. Güttler et al. (2012) considered the same runs for two regions with very variable orography, for which the 12 km run had the best overall performance. However, the 6 km run gave more improvement for Norway than Switzerland. Figure 9.41 shows an example of their simulated geographical patterns of precipitation. 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. Longer RCM runs at resolution higher than about 10 km are still rather few (e.g., Yasutaka et al., 2008), and even fewer at convection-permitting resolutions. Kendon et al. (2012) and Chan et al. (2012) found mixed results in daily precipitation simulated at 12 km and 1.5 km resolution, but noted that the latter improved sub-daily features. Some improvements were attributed to the fact that convection in the 1.5 km version could be accounted for on the model's grid scale rather than using a convection parameterisation.

[INSERT FIGURE 9.41 HERE]

Figure 9.41: Summer seasonal mean (JJA, 1987–2008) in southern-Norway gridded observational precipitation with 1 km resolution from Met.no (Mohr, 2008), and RCM-simulated precipitation with boundary conditions from the ERA40 reanalysis and ECMWF operational analysis (top row). The RCM was run at four different resolutions ranging from 50 to 6 km. Differences between the simulated precipitation and the gridded observations aggregated from 1 km to respectively 50, 25, 12 and 6 km grids are shown in the bottom row. After Güttler et al. (2012).

There are now more coupled regional climate models with interactive ocean and when appropriate also sea ice (Artale et al., 2010; Dorn et al., 2009; Doscher et al., 2010; Somot et al., 2008). Results indicate that this can improve the quality of the RCM results in certain cases. Döscher et al. (2010) found that a coupled RCM reproduced empirical relationships between Arctic sea ice extent and sea ice thickness and NAO. Samuelsson et al. (2010) showed that coupling a lake model with an RCM captured the effect of lakes on the air temperature over adjacent land in Europe. Lenaerts et al. (2012) added drifting snow in an RCM run for the Antarctica and found this increased the total area of ablation, improving the fit to observations. Smith et al. (2011a) added vegetation dynamics-ecosystem biogeochemistry into an RCM, and found some evidence of local feedback to near-surface temperature.

At the time of the AR4, RCMs were exclusively used for time-slice experiments. Since then, multi-decadal simulations and transient centennial projections with RCMs have emerged in larger numbers (e.g., de Elia et al., 2012; Diffenbaugh et al., 2011; Kjellstrom et al., 2011). Furthermore, coordinated RCM experiments and ensembles have become much more common. Overall, studies since the AR4 have covered Europe (e.g., Christensen et al., 2010), North America (e.g., Gutowski et al., 2010; Mearns et al., 2012), South America (e.g., Chou et al., 2012; Krüger et al., 2012; Menendez et al., 2010), Africa (e.g., Druyan et al., 2010; Paeth et al., 2011; Ruti et al., 2011), the Arctic region (e.g., Inoue et al., 2006) and Asian regions (e.g., Feng and Fu, 2006; Feng et al., 2011). A further development has been the globally-coordinated CORDEX downscaling experiment (Giorgi et al., 2009) with regional foci on Africa (Hernández-Díaz et al.; Kim et al., 2012b; Nikulin et al., 2012), Europe (Vautard et al., 2012), the Americas (Costa et al., 2012; Lucas-Picher et al., 2012a) and Asian regions (Ozturk et al., 2012; Shkolnik et al., 2007; Suh et al., 2012). In addition to providing ensembles of regional climate projections for impact research, coordinated intercomparisons enable better characterization of uncertainty as well as exploration of performance-based metrics (Christensen et al., 2010). For example, Suh et al. (2012) noted that the 10 RCMs run for Africa overall did well for average and maximum temperature, but systematically overestimated the daily minimum temperature. Precipitation was generally simulated betted for wet regions than for dry regions. Nikulin et al. (2012) found significant improvements in simulated precipitation compared to the ERA-Interim reanalysis used as boundary conditions. The quality of the diurnal cycle was, however, clearly related to the choice of convection scheme.

9.6.3 Skill of Downscaling Methods

Studies that have compared different SD techniques and/or SD and dynamical downscaling approaches (e.g., Maurer and Hidalgo, 2008; Schmidli et al., 2007) reiterate that downscaling skill varyies with location, season, parameter, and the AOGCM used as boundary conditions. A reasonable necessary condition for a skilful downscaling is that the driving global model, reanalysis or other such data provide a realistic large-scale setting (e.g., Diaconescu et al., 2012; van Oldenborgh et al., 2009). Limitations in the available boundary conditions remain an important consideration in assessing the skill of dowscaling methods.

Consequently, applying an RCM that has been developed for a specific region to other regions exposes it to a wider range of conditions and thus provides opportunities for more objective model evaluation. This is a feature of the coordinated experiments mentioned above, as the same RCMs are used for many regions. Transferability experiments target this explicitly. They involve running sets of RCMs for specific regions (cf. Jacob et al., 2012), holding process-descriptions constant for all domains. Not surprisingly, it is found that RCMs exhibit different skill for different regions (Gbobaniyi et al., 2011; Takle et al., 2007). Often no single model is found to systematically outperform others, which supports the usefulness of the multi-model approach.

The skill of downscaling methods can, in principle, be improved by correcting known biases, though this assumes that the biases remain constant and can be applied to future projections. Some studies highlight that this assumption needs to be carefully examined (e.g., de Elia and Cote, 2010) and (Boberg and Christensen, 2012; Christensen and Boberg, 2012). A range of bias correction methods have been proposed since the AR4 and some positive benefits for daily temperature and precipitation have resulted. For example, Dosio and Paruolo (2011) showed how biases in temperature and precipitation, including extremes, could be ameliorated with statistical bias correction. Yang et al. (2010) found the same for several statistics of RCM-simulated precipitation. Chen et al. (2012) found, on the other hand, that sub-daily precipitation errors were difficult to address with bias correction methods.

A complement to bias correction that has been explored is performance-based ranking or weighting of RCM ensembles. Christensen et al. (2010) examined metrics based on the ability of RCMs to simulate extremes, mesoscale features, trends, aspects of variability and consistency with the driving boundary conditions. One of these metrics led to striking differentiation among RCMs (Lenderink, 2010), whereas others did not. The latter may imply general skilfulness of models, but may simply indicate that the other metrics were not very informative. Coppola et al. (2010), Kjellström et al. (2010) and Sobolowski and Pavelsky (2012) demonstrated, for two different regions, that weighted sets of RCMs outperformed sets without weighting for both temperature and precipitation. (Wehner, 2012) highlighted that in the presence of outliers among the RCMs, skill-based weighting improved the ensemble result for extreme precipitation.

9.6.4 Value Added

RCMs are regularly tested to evaluate whether improvements over global models materialize on smaller scales (Laprise et al., 2008), i.e., whether they do indeed ‘add value’. In addition to better capturing topography-influenced phenomena and extremes with relatively small spatial and/or short temporal character, RCMs may also modify information at larger scales via non-linear interactions involving their detailed representation of small scales and processes. Effects that travel up-scale have been found improve specific aspects of the larger-scale simulated climate (Inatsu and Kimoto, 2009; Inatsu et al., 2012; Lorenz and Jacob, 2005), but these can also degrade others (Laprise et al., 2008; Sanchez-Gomez et al., 2009).

Downscaled results appear very rich in small-scale features; however, when averaged in time, the degree of added detail is considerably reduced, irrespective of the variable analysed. As a result, differences between time-averaged RCM and GCM fields are not always obvious, especially in fairly homogeneous regions. Added value is much more obvious in studies that go beyond time averages (e.g., Feser and Barcikowska, 2012; Feser et al., 2011; Shkol’nik et al., 2012), although quantification of added value is a difficult task.

In essence, added value is a measure of the extent to which downscaled climate variables are closer to observations than the model from which the boundary conditions were obtained. There are several cases of reported added value in this sense, including improved simulation of convective precipitation (Rauscher et al., 2010), near-surface temperature (Feser, 2006), near-surface temperature and wind (Kanamaru and Kanamitsu, 2007), temperature and precipitation (Lucas-Picher et al., 2012b), extreme precipitation (Kanada et al., 2008), coastal features (Kawazoe and Gutowski, 2012; Vautard et al., 2012; Winterfeldt and Weisse, 2009; Winterfeldt et al., 2011), European storm damage estimates (Donat et al., 2010), strong mesoscale cyclones in the Mediterranean (Cavicchia and Storch, 2011) and cutoff lows in Australia (Grose et al., 2012). Distinctive polar lows can be simulated in RCMs, even when they are not conspicuous in the driving data (Zahn and von Storch, 2008). Studies consistently indicate that added value does arise, particularly for regions with distinct mesoscale phenomena, variable orography or other variable surface characteristics, for example in coastal areas (Feser et al., 2011). Added value for higher statistical moments of precipitation or the water budget has also been noted on (Bresson and Laprise, 2011), although care is needed in order to avoid attributing unwarranted added value when RCM-downscaled improvements can be a result of compensating errors (Grose et al., 2012; Kanamitsu and DeHaan, 2011; Tjernstrom et al., 2008).

9.6.5 Sources of Model Errors and Uncertainties

In the case of SD, errors and uncertainties arise from estimating empirical relationships between predictors and predictands from limited data sets. For RCMs, in addition to resolution effects and process parameterisations, errors and uncertainties arise from the choice and application of boundary conditions (driving data) and domain.

An RCM is strongly constrained by the driving data close to the lateral boundaries, and so in small domains, there is less freedom for the RCM to generate additional information. On the other hand, larger domains allow the RCM solution to become increasingly 'decoupled' from the driving data and so systematic errors can grow. This can be constrained by techniques such as spectral nudging (Misra, 2007; Separovic et al., 2012), which precludes the larger scales of the RCM solution from deviating from the large-scale solution that underlies the boundary conditions. However, the desirability of such nudging is debated (Veljovic et al., 2010). 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. Improvements with increasing domain size in mean temperature, precipitation and wind, as well as of daily variability and extremes were found by Køltzow et al. (2008). The quality of the RCM results may also vary according to the synoptic situation, season, and the lateral boundary location (Alexandru et al., 2007; Leduc and Laprise, 2009; Nikiema and Laprise, 2010; Rapaić et al., 2010; Xue et al., 2007), but these effects are generally found to be small in climate-length applications (Laprise et al., 2008; Separovic et al., 2008).

Process-oriented evaluation is a powerful complement to evaluation in terms of state variables. For example, Sasaki and Kurihara (2008) examined the ability of an RCM to reproduce observed relationships between precipitation and elevation. Driouech et al. (2010) showed that a specific variable-resolution AGCM reproduced rather well the observed link between north Atlantic weather regimes and local precipitation. Hirschi et al. (2011b) found that a number of RCMs reproduce observed relationships between soil moisture and extreme temperature. When it comes to uncertainties related to the representation of physical processes, the same issues arise as in global models: cloud-related processes and land-surface/atmosphere interactions, cumulus convection schemes (Lynn et al., 2009), the combination of schemes for convection and the planetary boundary layer, horizontal diffusion and/or microphysics (Axelsson et al., 2011; Crétat et al., 2012; Pfeiffer and Zängl, 2010; Solman and Pessacg, 2012; Tjernstrom et al., 2008; Wyser et al., 2008). Roy et al. (2011a) demonstrated that a new land surface scheme with a more sophisticated representation of soil moisture scheme improved the simulation of daily temperature and precipitation extremes. Yhang and Hong (2008) and Cha et al. (2008) tested improvements of land surface, boundary layer and cumulus parameterisation schemes in an RCM, and improved its simulation of the East Asian summer monsoon. Land-surface and atmosphere coupling is in general found to be particularly important for simulating regions with monsoons (Boone et al., 2010; Druyan et al., 2010; van den Hurk and van Meijgaard, 2010).

9.6.6 Relating Downscaling Performance to Credibility of Regional Climate Information

A fundamental issue is how past performance of a downscaling method relates to its ability to provide local and regional information regarding future climate.

When it comes to SD, the statistical stationarity hypothesis (that statistical relationships inferred from historical data remain valid under a changing climate) cannot be directly tested. Some recent studies have proposed ways to examine its validity using RCM outputs (e.g., Driouech et al., 2010; Vrac and Naveau, 2008) or using long series of observations (e.g., Schmith, 2008). However, while the stationarity of biases can be explored in part, it cannot be fully tested (Maraun, 2012), and remains unsettled.

For RCMs, their credibility is conditional on both the quality of the RCMs themselves, and of the boundary conditions (e.g., Dawson et al., 2012; Deque et al., 2012; Eum et al., 2012). Giorgi and Coppola (2010) argued that regional-scale climate change signals in the CMIP3 models were not sensitive to their temperature biases, over land. For precipitation, the same was found for about two thirds of the global land area. However, there is recent evidence that regional-scale model biases may be non-linear for temperature extremes (Boberg and Christensen, 2012; Christensen et al., 2008; Christensen and Boberg, 2012) for both global and regional models, and this could be important for the application of bias corrections. One mechanism at play may be that RCMs tend to dry out the soil too effectively at high temperatures in some regions, which can lead to systematic biases in projected changes in warm summertime conditions, including heat waves (Christensen et al., 2008; Kostopoulou et al., 2009). This is illustrated in Figure 9.42 for the Mediterranean region based on the RCMs participating in ENSEMBLES and CMIP3 GCMs. There is a tendency towards an enhanced warm bias in the warmer months of the year. In climate change projections, consequently, part of the typically large warming signal in these regions could be due to model bias (Boberg and Christensen, 2012). These issues can also be seen in North America (Mearns et al., 2012).

[INSERT FIGURE 9.42 HERE]

Figure 9.42: Ranked modelled versus observed monthly mean temperature for a Mediterranean region for the 1961–2000 period. The RCM data (panel a) are from Christensen et al. (2008) and are adjusted to get a zero mean in model temperature with respect to the diagonal. The smaller insert shows uncentred data. The GCM data in panel b are from CMIP3 and adjusted to get a zero mean in model temperature with respect to the diagonal. Figure after Boberg and Christensen (2012).

Di Luca et al. (2012) analysed the climate change signal from six RCMs run at 50km resolution on a North American domain and found that in general, the climate change signals were quite similar to the driving data, after downscaling, for precipitation and surface temperature. They attributed this to the stationary character of the smaller-scale structure associated with orographic features, i.e., the spatial detail gained in surface temperature by downscaling was comparable in both present and future climate and thus not so conspicuous in the climate change signal.

Systematic or individual model biases can be addressed by means of bias correction and possibly performance-based weighting as was discussed in Section 9.6.3, which is more straightforward if the biases do not significantly depend on the (changing) climate state. Results from Buser et al. (2009) suggest this may not be the case. They found that the projected summertime warming in the European Alpine region, obtained from a combination of several RCMs, ranged from 3.4°C to 5.4°C depending on which of two assumptions was made regarding bias behaviour under changing climate conditions. Projected changes for the wintertime warming were not as sensitive.

In summary, although there are caveats and uncertainties, regional downscaling methods do offer a means of providing credible, physically-based climate information at the small scales needed for many impact studies. While errors and biases from the global model projections propagate directly to regional downscaling, the overall finding is that downscaling adds value, compared to driving data from AOGCMs or global reanalyses both in regions with highly variable topography (e.g., distinct orography, coast lines) and for various small-scale phenomena. These results arise from a variety of distinct studies with different RCMs, rather than coordinated experiments, so there is high agreement, but medium evidence. Added value to time-averaged temperature and precipitation is much less obvious, and there is so far only limited evidence of added value in terms of higher-order statistics and extremes.

9.7 Understanding Model Performance and Climate Sensitivity

The previous sections have dealt with the ability of climate models to simulate recent and longer-term records, variability and extremes, and regional-scale climate. We have assessed this capability mainly by comparing model results to observations and by evaluating inter-model spread, the latter being a minimum estimate of model uncertainty. In this section we provide some assessment of *why* models show errors and spread. This identification is crucial not only for understanding why models fail to reproduce observations, but also for diagnosing whether models obtain the right answer for the right reason.

Error in model results can be conceptually subdivided into “modelling error”, caused by the difference between model formulation and physical process, and “approximation error”, caused by the difference between true model solution and numerical approximation (Oden and Prudhomme, 2002). No general framework exists for diagnosing modelling error. In contrast, for approximation error in geophysical fluid dynamics a general framework has just been formulated (Rauser et al., 2011), but application has so far been restricted to a shallow-water model. When we assess the causes of errors in *current* climate models, we thus cannot build on a general conceptual framework and must instead rely on more ad-hoc approaches, governed by practicality.

9.7.1 Understanding Model Performance

9.7.1.1 Uncertainty in Process Representation

Some model errors can be traced to uncertainty in representation of processes in models, i.e., parameterizations or their interactions. Some of them have been long-standing issues in climate modelling, reflecting our limited, though gradually increasing, understanding of the processes and inherent challenges in mathematically representating the highly complex processes.

For the atmosphere, cloud processes, including convection and its interaction with boundary layer and larger-scale circulation, remain major sources of uncertainty. This is evident for the vertical distribution of water vapour and the distribution of clouds, particularly over subtropical and arctic regions. These in turn cause errors or uncertainties in radiation which propagate into the coupled atmosphere-ocean system. The omission or underestimation of absorbing aerosols, the omission of weak-line absorption by water vapour and overestimation of surface albedo cause additional common errors in radiative fluxes. Distribution of aerosols is also a source of uncertainty arising from modelled microphysical processes and transport. It has also been pointed out that simulation of storm tracks and extratropical cyclones, diurnal cycle of precipitation over land in the tropics and extratropics in summer, the Madden-Julian Oscillation and tropical cyclones are strongly dependent on parameterizations, particularly those related to convection. For example, a bias common to most models is that moist convection starts prematurely and it rains too frequently at reduced intensity. This so called "drizzle bias" is improved by changes in the representation of entrainment in deep convection, improved coupling between shallow and deep convection, and inclusion of density currents.

Ocean models, on the other hand, are subject to uncertainty in parameterizations of vertical and horizontal mixing and convection: common warm and saline biases in the intermediate and deep layers in the Labrador and Irminger Seas are caused by too shallow convection; inter-model spread in the volume transport of the Antarctic circumpolar current appears to be mainly due to uncertainty in the eddy-induced thickness diffusivity; a common warm bias in the subsurface water in the eastern equatorial Atlantic is traced to problems in vertical mixing; and inter-model spread in subsurface temperature in the eastern equatorial Indian Ocean may also be related to uncertainty in vertical mixing. Some errors in the ocean evidently affect the atmosphere. For example, weak coastal upwelling off South America and a resulting warm SST bias is thought to be responsible for insufficient marine stratocumulus cloud in the eastern tropical Pacific. This in turn affects surface radiative fluxes, generating a feedback loop in the coupled atmosphere-ocean system. To understand common biases in the equatorial Pacific mean state, causes in both the atmosphere and ocean must be considered, including too strong trade winds, a too diffusive thermocline, deficient horizontal isotropic mixing coefficients, insufficient penetration of solar radiation and too weak tropical instability waves. Similarly, simulation of ENSO is known to be affected by deep convection, trade wind strength and cloud feedbacks. Explicit representation of SST diurnal variations has been shown to reduce a long-standing cold bias in the tropical Pacific.

Simulation of sea ice is also affected by errors in both the atmosphere and ocean, including high-latitude winds, polar clouds, oceanic heat advection and mixing. The parameterization of sea ice itself is also important. For example, new sea-ice albedo parameterization schemes that allow for melt ponds have been implicated in the improved representation of summertime Arctic sea-ice decline.

With respect to biogeochemical components in ESMs, parameterizations of nitrogen limitation and forest fires are thought to be important for simulating the carbon cycle, but very few ESMs incorporate this so far. Deposition of carbonaceous aerosols on snow and ice is also recognized as important, but not commonly included in models. There is an indication that more sophisticated treatment of aerosols improves the simulation of historical variations in the Atlantic multi-dicadal oscillation. Regarding permafrost, very few models include the key processes necessary to accurately model permafrost changes, such as the distinct properties of organic soils, the existence of local water-tables, and the heat released by microbial respiration.

9.7.1.2 Error Propagation

Cause of one model bias can sometimes be associated with another. Although the root cause of those biases is often unclear, knowledge on the causal chain of biases or a set of interrelated biases can provide a key to further understanding and improvement of model performance. For example, biases in storm track position are partly due to a SST biases in the Gulf Stream and Kuroshio Current; unrealistic behaviour of the simulated Atlantic meridional mode is likely to be associated with biases in the ITCZ; uncertainty in permafrost simulation is partly caused by biases in air temperature and snow depth; simulation of ocean biogeochemistry is affected by underlying physical ocean simulation (e.g., upwelling and mixed-layer depth); and underestimated polar amplification in paleoclimate simulations can partly be attributed to uncertainties in sea-ice and vegetation feedbacks. Some biases in variability or trend can be partly traced back to biases in mean states. For example, a decreasing trend in September Arctic ice extent tends to be underestimated when sea ice thickness is overestimated. In such a case, improvement of the mean state may improve simulated variability or trend.

9.7.1.3 Sensitivity to Resolution

Some phenomena or aspects of climate are found to be better simulated with models run at higher horizontal and/or vertical resolution. In particular, increased resolution in the atmosphere has improved, at least in some models, storm track and extratropical cyclones, diurnal variation of precipitation over land, extreme precipitation, and tropical cyclones. Similarly, increased horizontal resolution in the ocean is shown to improve sea surface height variability, western boundary currents, tropical instability waves and coastal upwelling, and variability of Atlantic meridional overturning circulation. High vertical resolution and a high model top, as well as high horizontal resolution, are shown to be important for simulating lower stratospheric climate variability, the Quasi-Biennial Oscillation, blocking, and the North Atlantic Oscillation. Higher resolution regional climate models show improvements in aspects of simulated daily and sub-daily precipitation.

9.7.1.4 Uncertainty in Observational Data

In some cases, insufficient length or quality of observational data makes model evaluation challenging, and is a frequent problem in the evaluation of simulated variability or trends. This is evident for evaluation of upper tropical tropospheric temperature, tropical atmospheric circulation, the Atlantic meridional overturning circulation, the North Atlantic Oscillation, and the Pacific Decadal Oscillation. Data quality has been pointed out as an issue for arctic cloud properties, ocean heat content, heat and fresh water fluxes over the ocean, and extreme precipitation. There are also issues when it comes to precipitation over ocean and polar regions. Paleoclimate reconstructions also have large inherent uncertainties.

It is clear therefore that updated or newly available data affect model evaluation conclusions. For example, changes in the observational estimate of cloud radiative effect have reduced the model-data discrepancy, and newly available vertically resolved cloud information has highlighted difficulties in simulating clouds and water vapour fields in models.

9.7.1.5 Other Factors

Model evaluation can be affected by how models are forced. For example, models with prescribed or simulated stratospheric ozone depletion perform better than fixed ozone. Similarly, uncertainties in specified greenhouse gases, aerosols, land-use change, etc. will all affect model results and hence evaluation of model quality. Different statistical methods used in model evaluation can lead to subtle or substantive differences in the assessment of model quality. The quality of boundary conditions used in driving regional models is imprinted on the regional model simulation and so affects conclusions regarding model quality.

9.7.2 Climate Sensitivity and Climate Feedbacks

9.7.2.1 Equilibrium Climate Sensitivity, Idealised Radiative Forcing, and Transient Climate Response in the CMIP5 Ensemble

Equilibrium climate sensitivity (ECS) is the equilibrium change in global-mean surface temperature after doubling the atmospheric concentration of CO_2 relative to pre-industrial levels. It represents the single most important measure of climate response because the response of many other climate variables to an increase in CO_2 scales with the increase in global-mean surface temperature (Meehl et al., 2007b). In the AR4, the range in equilibrium climate sensitivity of the CMIP3 models was 2.1°C to 4.4°C, and the single largest contributor to this spread was differences among modelled cloud feedbacks. These assessments carry over to the CMIP5 ensemble without any substantial change.

The method of diagnosing climate sensitivity in CMIP5 differs fundamentally from the method employed in CMIP3 and assessed in AR4 (Randall et al., 2007). In CMIP3, an AGCM was coupled to a non-dynamic mixed-layer (slab) ocean model with prescribed ocean heat transport convergence. CO_2 concentration was then instantaneously doubled, and the model was integrated to a new equilibrium with unchanged implied ocean heat transport. While computationally efficient, this method had the disadvantage of employing a different model from the AOGCM used for the historical simulations and climate projections. However, in the few comparisons that were made, the resulting disagreement in ECS was less than about 10% (Boer and Yu, 2003; Danabasoglu and Gent, 2009; Li et al., 2012a).

In CMIP5 it was decided to diagnose climate sensitivity directly from the AOGCMs following the approach of Gregory et al (2004). In this case the CO_2 concentration is instantaneously quadrupled and kept constant for 150 years of simulation, and both equilibrium climate sensitivity and radiative forcing (see below) are diagnosed from a linear fit of the model output to the global energy balance relationship

$$N = F - \alpha \Delta T \, , \tag{9.1}$$

where N is the instantaneous radiative imbalance at the top of the atmosphere, F the adjusted radiative forcing (see Chapter 7), α the climate feedback parameter (which is the inverse of the climate sensitivity parameter, see Glossary), and ΔT the perturbation in global-mean surface temperature. If F and α are constant, a linear fit of Equation (9.1) to the annual-mean AOGCM output of N against ΔT yields the ECS as the value of ΔT when N=0 and the forcing F as the value of N when ΔT=0, respectively. Because in CMIP5 the CO_2 concentration is quadrupled and not doubled, the estimates of F and ECS are obtained from the linear fit in the same manner followed by a division by two on the assumption that the forcing and equilibrium climate sensitivity depend logarithmically on CO_2 concentration (Manabe and Bryan, 1985).

While the method employed in CMIP5 obviates the need to maintain a separate slab version of the AOGCM, at least three new sources of uncertainty could be introduced if one or more of the assumptions underlying the method were violated. First, the assumption that F and α are strictly constant is not always valid (Block and Mauritsen, 2012; Boer and Yu, 2003; Giorgetta et al., 2012; Williams et al., 2008). Second, climate feedbacks can be different during the initial transient and subsequent equilibrated phases of the AOGCM integration with increased CO_2 concentration (Yokohata et al., 2008). Third, the method assumes a strictly logarithmic dependence of the climate response to elevated CO_2 concentrations. While there is evidence for this logarithmic dependence (Manabe and Bryan, 1985), in a low-resolution version of the CMIP3 model ECHAM5/MPI-OM the global-mean surface warming was larger than indicated by a logarithmic dependence on CO_2 concentrations (Li et al., 2012a). Despite these potential sources of uncertainty, a rigorous comparison against the equilibrium climate response of an AOGCM to quadrupled CO_2 concentrations showed that the technique is accurate to within 10% (Li et al., 2012a).

As an alternative to estimating radiative forcing using regression against Equation (9.1), Hansen et al (2005) diagnose the forcing from the radiative imbalance at the top of the atmosphere using simulations with fixed SST but quadrupled CO_2 concentrations. Both methods are used in CMIP5 and the results are listed in Table 9.4. The differences between the estimates from the two methods gives a measure of the methodological uncertainty in adjusted radiative forcing (Colman and McAvaney, 2011). To ensure comparability with the equilibrium climate sensitivity, both estimates of radiative forcing are likewise divided by two to obtain estimates equivalent to CO_2 doubling.

The third important quantity in this context is the transient climate response (TCR), defined as the global-mean surface temperature change averaged over the 20-year period centred on the time of CO_2 doubling from an experiment in which the CO_2 concentration is increased by 1% yr^{-1}. The transient climate response is smaller than the equilibrium climate sensitivity because ocean heat uptake delays surface warming. TCR is linearly correlated with ECS in the CMIP5 ensemble (Figure 9.43), implying that although differences in ocean heat uptake across the CMIP5 multi-model ensemble contribute to a spread in transient climate response, the dominant cause of spread in TCR is spread in ECS.

Based on the methods outlined above and explained in Section 9.7.2.2 below, Table 9.4 shows the diagnosed radiative forcings, equilibrium climate sensitivities, transient climate responses, and feedback parameters from doubling or quadrupling CO_2 concentrations obtained from the CMIP5 ensemble. The two estimates of radiative forcing agree with each other to within 5% for six models (CanESM2, INM-CM4, IPSL-CM5A-LR, MIROC5, MPI-ESM-LR, and MPI-ESM-P), although the deviation exceeds 10% for four models (CCSM4, CSIRO-Mk3-6-0, HadGEM2-ES, and MRI-CGCM3) and is indicative of deviations from the basic assumptions underlying one or both forcing estimation methods. However, the mean difference of 0.3 W m^{-2} between the two methods for diagnosing radiative forcing is only about half of the ensemble standard deviation of 0.5 W m^{-2}, or 15% of the mean value for radiative forcing by CO_2 using fixed SSTs. Equilibrium climate sensitivity and transient climate response vary across the ensemble by a factor of approximately two. The multi-model ensemble mean in equilibrium climate sensitivity is 3.4°C, a value identical to that for CMIP3, while the CMIP5 ensemble range is 2.1°C to 4.7°C, a spread which is also nearly indistinguishable from that for CMIP3. While every CMIP5 model whose heritage can be traced to CMIP3 shows some change in equilibrium climate sensitivity, there is no discernible systematic tendency. In particular, including ESMs in CMIP5 has caused no systematic difference in the equilibrium climate sensitivities of the ensemble. This broad similarity between CMIP3 and CMIP5 and the good agreement between different methods where they were applied to the same atmospheric GCM indicate that the uncertainty in methodology is minor compared to the overall spread in equilibrium climate sensitivity. The change in transient climate response from CMIP3 to CMIP5 is generally of the same sign but of smaller magnitude compared to the change in equilibrium climate sensitivity.

Finally, while it would be conceivable that equilibrium climate sensitivity were a function of global mean surface temperature, owing to the dependence of the water vapour feedback and the surface albedo feedback on temperature (see Section 9.7.2.2), Figure 9.43 shows no discernible correlation, a fact that enhances confidence in model simulations of a warming climate even in the presence of errors in the time-mean.

[INSERT TABLE 9.4 HERE]

Table 9.4: Climate sensitivity estimates (Andrews et al., 2012b) and feedback parameters (Soden et al., 2008) for the CMIP5 AOGCMs (see Table 9.1 for model details). The entries were calculated according to Hansen et al (2005) for radiative forcing using fixed SSTs; Gregory et al (2004) for radiative forcing and equilibrium climate sensitivity using regression; Soden et al (2008) for the feedback parameters using radiative kernel methods; and Taylor et al.'s (2012) reference CMIP5 experiment with 1% CO_2 increase per year for the transient climate response using the 20-year mean centred on the year of CO_2 doubling. Notice that the entries for radiative forcing and equilibrium climate sensitivity were obtained by dividing by two the original results, which were obtained for CO_2 quadrupling.

[INSERT FIGURE 9.43 HERE]

Figure 9.43: A) Equilibrium climate sensitivity (ECS) against the global mean surface air temperature of CMIP5 models for the period of 1961 to 1990. B) Equilibrium climate sensitivity against transient climate response (TCR). The ECS and TCR information is taken from Andrews et al. (2012b).

9.7.2.2 Understanding the Range in Model Climate Sensitivity: Climate Feedbacks

The feedback parameters for the CMIP3 and CMIP5 models are compared in Figure 9.44 (panel A). Changes from CMIP3 to CMIP5 in the water-vapour and cloud feedbacks are statistically insignificant. While increases in the albedo, lapse-rate, and combined water-vapor and lapse rate feedbacks of 36%, 26%, and 11% (respectively) are statistically significant at the 98% level, the difference between the sums of all feedbacks are statistically indistinguishable. Advances in determining each of the feedback parameters enumerated in Table 9.4 are described in detail below.

[INSERT FIGURE 9.44 HERE]

Figure 9.44: a) Feedback parameters for CMIP3 and CMIP5 models (left and right columns of symbols) for water vapour (WV), clouds (C), albedo (A), lapse rate (LR), combination of water vapour and lapse rate (WV+LR), and sum of all feedbacks (ALL) updated from Soden and Held (2006). CMIP5 feedbacks are derived from CMIP5 simulations for abrupt four-fold increases in CO_2 concentrations (4 × CO_2). b) ECS obtained using fixed-SST and regression techniques by Andrews et al. (2012b) against ECS estimated from the ratio of CO_2 radiative forcing to the sum of all feedbacks. The CO_2 radiative forcing is one-half the 4 × CO_2 forcings from Andrews et al. (2012b), and the sum of feedbacks (ALL + Planck) is updated from Soden and Held (2006).

9.7.2.2.1 Role of humidity and lapse rate feedbacks in climate sensitivity

Tests of the water-vapour feedback in AOGCMs have advanced since AR4 using combinations of satellite data and meteorological analyses. Correlations between coincident variations in SST and clear-sky outgoing longwave radiation (OLR) provide estimates on the rate of radiative damping of SST fluctuations. Modelled values for clear-sky damping are internally consistent across the CMIP3 multi-model ensemble and are a good approximation of the observationally-derived damping rate, confirming the robustness of a positive water-vapour feedback (Chung et al., 2010b). The modelled and observationally-derived damping rates are consistent with a strong positive correlation between SST and water vapour on regional to global scales (Allan, 2009) and reproduce the tropical correlations observed in reanalyses of ENSO episodes (Dessler and Wong, 2009). Finally, the relationship of fluctuations in SST and upper-tropospheric humidity can be derived from the Atmospheric Infrared Sounder (AIRS), and the results show that AOGCMs can reproduce the positive rate of increase in specific humidity with increased SST, a key term in the water-vapour feedback, to within 10–25% $°C^{-1}$ (Gettelman and Fu, 2008).

The compensation between the water-vapour and lapse-rate feedbacks noted in the CMIP3 AOGCMs is still present in the CMIP5 models, and possible explanations of the compensation have been developed (Ingram, 2010). New formulations of the feedbacks, replacing specific with relative humidity, eliminate most of the cancellation between the water-vapour and lapse-rate feedbacks and reduce the inter-model scatter in the individual feedback terms (Held and Shell, 2012)

9.7.2.2.2 Role of surface albedo in climate sensitivity

The two primary factors that govern the magnitude of the surface albedo feedback are the responses of snow and sea-ice areal coverage, and hence surface albedo, to increased surface temperatures. The spread in CMIP3 snow-albedo feedbacks is highly correlated with intermodel difference in middle to high latitude present-day springtime albedo, and these differences are due to large uncertainties in simulating the relationship between microphysical and optical properties of snow (Levis et al., 2007). These optical properties are strongly affected by snow melt and other temperature-related effects, and recent satellite retrievals suggest that temperature-modulated optical properties contribute nearly 30% of the snow-albedo feedback while snow cover contributes roughly 70% (Fletcher et al., 2012). The magnitude of the temperature dependence of snow albedo in CMIP3 models is 50 to 100% less than observed in the Arctic basin (Fletcher et al., 2012). However, over Antarctica the positive temperature-induced feedbacks from snow optics may be largely offset by projected changes in precipitation since small increases in precipitation introduce small snow grains and hence increase snow albedo (Picard et al., 2012).

Analysis of observed declines in sea-ice and snow coverage from 1979 to 2008 suggests that the Northern Hemispheric albedo feedback is between 0.3 and 1.1 W m^{-2} $°C^{-1}$ (Flanner et al., 2011). This range is substantially above the global feedback of 0.4 ± 0.1 W m^{-2} $°C^{-1}$ of the CMIP5 models analysed in Table 9.4. The difference suggests that the CMIP5 AOGCMs underestimate the strength of the feedback as did the CMIP3 models based upon the systematic errors in simulated sea-ice coverage decline relative to observed rates (Boe et al., 2009c).

9.7.2.2.3 Role of cloud feedbacks in climate sensitivity

Cloud feedbacks represent one of the main causes for the range in modelled climate sensitivity (Chapter 7). The spread due to inter-model differences in cloud feedbacks is approximately 3 times larger than the spread contributed by feedbacks due to variations in water vapour and lapse-rate combined (Dufresne and Bony, 2008), and is a primary factor governing the range of climate sensitivity across the CMIP3 ensemble (Volodin, 2008c). Differences in ECS and TCR are due primarily to differences in the shortwave cloud feedback (Yokohata et al., 2008). In perturbed ensembles of three different models, the primary factor contributing to the spread in equilibrium climate sensitivity is the low-level shortwave cloud feedback (Klocke et al., 2011c; Yokohata et al., 2010a).

An approach particularly well suited for estimation of climate sensitivity in multi-model ensembles is based upon radiative kernel methods (Soden et al., 2008). In the kernel approach, each feedback related to a climate property is represented as the product of the change in the radiative energy budget to incremental changes in that property multiplied by the change in that property to incremental changes in surface temperature. The first term is generally derivable from basic physical principles and can be calculated once and applied to multiple models (Soden et al., 2008) while the second term is model-specific and can be readily derived for each member of the multi-model ensemble. Application of radiative kernel techniques to multiple models forced by doubled CO_2 show that while changes in cloud forcing can be either positive or negative, the cloud feedbacks are generally positive or near neutral (Shell et al., 2008; Soden et al., 2008). All of the models examined in a multi-thousand member ensemble of AOGCMs constructed by parameter perturbations also have net positive or neutral cloud feedbacks (Sanderson et al., 2010). This finding is consistent with the modelled and measured relationships between SSTs and top-of-atmosphere radiative fluxes, which suggest that interannual cloud variations act as a positive feedback in the current climate (Chung et al., 2010a; Dessler, 2010; Zelinka et al., 2012).

Over the north-east Pacific, decadal-scale fluctuations in surface and satellite-based measurements of low-level cloud cover are significantly negatively correlated with variations in SST (Clement et al., 2009). This negative correlation is consistent with a positive low-cloud feedback in this region operating on decadal time scales. Models that reproduce this negative correlation and other relationships between cloud cover and regional meteorological conditions simulate a positive low-cloud feedback over much of the Pacific basin (Clement et al., 2009).

Changes in the high-altitude clouds also induce climate feedbacks due to the large areal extent and significant longwave cloud radiative effects of tropical convective cloud systems. Analysis of the budget terms in the cloud condensate tendency equation shows that inter-model differences in the cloud response are attributable to processes such as condensation-evaporation and ice sedimentation (Ogura et al., 2008a; Ogura et al., 2008b). In experiments with perturbed physics ensembles of AOGCMs, the parameterisation of ice fall speed also emerges as one of the most important determinants of climate sensitivity (Sanderson et al., 2010; Sanderson et al., 2008b; Sexton et al., 2012).

9.7.2.2.4 Relationship of feedbacks to modelled climate sensitivity

The ECS can be estimated from the ratio of the forcing to the total climate feedback parameter using Equation 9.1. This approach is applicable to simulations in which the net radiative balance is much smaller than the forcing and hence the modelled climate system is essentially in equilibrium. This approach can also serve to check the internal consistency of estimates of the ECS, forcing, and feedback parameters obtained using independent methods. The relationship between ECS from Andrews et al (2012a) and estimates of ECS obtained from the ratio of forcings to feedbacks is shown in Figure 9.44 (Panel B). The forcings are estimated using both regression and fixed sea-surface temperature techniques (Gregory et al., 2004; Hansen et al., 2005) by Andrews et al (2012a) and the feedbacks are calculated using radiative kernels (Soden et al., 2008). On average, the ECS from forcing to feedback ratios underestimate the ECS from Andrews et al (2012a) by 5% and 15% using fixed-SST and regression forcings respectively.

9.7.2.2.4 Relationship of feedbacks to uncertainty in modelled climate sensitivity

Objective methods for perturbing uncertain model parameters to optimize performance relative to a set of observational metrics have shown a tendency toward an increase in the mean and a narrowing of the spread of estimated climate sensitivity (Jackson et al., 2008a). This tendency is opposed by the effects of gaining better knowledge regarding structural biases related to incomplete process representations that are shared across a multi-model ensemble. If structural biases are shared among models in a MME, the most likely sensitivity for the MME tends to shift towards lower sensitivities while the possibility of larger sensitivities increases at the same time (Lemoine, 2010). Following Schlesinger and Mitchell (1987), Roe and Baker (2007) suggest that symmetrically distributed (e.g., Gaussian) uncertainties in feedbacks lead to inherently asymmetrically distributed uncertainties in climate sensitivity with increased probability in extreme positive values of the sensitivity. Roe and Baker (2007) conclude that this relationship makes it extremely difficult to reduce uncertainties in climate sensitivity through incremental improvements in the specification of feedback parameters. While subsequent analysis has suggested that this finding could be an artifact of the statistical formulation (Hannart et al., 2009) and linearization (Zaliapin and Ghil, 2010) of the relationship between feedback and sensitivity uncertainty adopted by (Roe and Baker, 2007), these issues remain unsettled (Roe and Armour, 2011; Roe and Baker, 2011). Sexton et al (2012) have also demonstrated methods for reducing climate sensitivity uncertainty using PPEs.

Using a Bayesian framework to analyse perturbed physics experiments using a slab-ocean GCM, Sanderson et al (2008b), Rougier et al (2009b), and Sexton et al (2012) find that the rate of cloud entrainment is the parameter with the largest direct impact on uncertainty in AOGCM climate sensitivity. An additional source of uncertainty in equilibrium sensitivity is that some of the adjustments of clouds to forcing are sufficiently fast that the adjustments should be considered part of the forcing rather than the response (Chapter 8). These experiments involve instantaneously increasing (usually doubling) the concentrations of CO_2 and then monitoring the rate at which radiative equilibrium is restored or estimating the asymptotic equilibrated surface temperature increase. The instantaneous increase induces very rapid atmospheric and terrestrial adjustments analogous to the semi-direct effects of aerosols including adjustments to the cloud field, tropospheric lapse rate and humidity and snow cover (Andrews and Forster, 2008; Gregory and Webb, 2008). These findings have highlighted the importance of separating the fast responses that depend on (instantaneous) changes in forcing and the feedbacks that follow the much slower adjustments in ocean temperature (Chapter 7).

9.7.2.3 Climate Sensitivity and Model Performance

Despite the range in equilibrium sensitivity of 2.1°C to 4.4°C for CMIP3 AOGCMs, these models reproduce the global surface air temperature anomaly of 0.76°C over 1850–2005 to within 25% relative error. The relatively small range of historical climate response suggests that there is another mechanism, for example a compensating non-GHG forcing, present in the historical simulations that counteracts the relatively large range in sensitivity obtained from idealized experiments forced only by increasing CO_2. One possible mechanism is a systematic negative correlation across the multi-model ensemble between climate sensitivity and anthropogenic aerosol forcing (Anderson et al., 2010; Kiehl, 2007; Knutti, 2008). A second possible mechanism is a systematic overestimate of the mixing between the oceanic mixed layer and the full depth ocean underneath (Hansen et al., 2011).

However, despite a comparable range of sensitivities in the CMIP5 models, there is no significant relationship across the multi-model ensemble between the equilibrium climate sensitivities and the adjusted 20th century forcings applied to each individual model (Forster et al., 2012). Differences in ocean heat uptake also do not appreciably affect the spread in projected changes in global mean temperature by 2095 (Forster et al., 2012). The evidence for a mechanism compensating for the large range in ECS to produce a narrow range in 20th century temperature change is therefore weaker in CMIP5 than in CMIP3.

9.7.2.3.1 Model evaluation of EMICs and constraints on climate sensitivity

An EMIC intercomparison project (Eby et al., 2012; Zickfeld et al., submitted) allows an assessment of carbon-cycle feedbacks and model response characteristics, including ECS, TCR, and heat uptake efficiency (Table 9.5). (Eby et al., 2012) confirm results of Plattner et al. (2008) and conclude that the EMICs compare favourably with AOGCM results over the 2000 to 2100 period. Relevant results are shown in Figure 9.8. Overall, these studies imply that EMICs are well suited for simulations beyond the CMIP5 ensemble 2300 period, extending over the future millennium.

Attributes of current EMICs are provided in Table 9.6 (Eby et al., 2012). Significant advances in EMIC capabilities are inclusion of ice-sheets (UVic 2.9 (Weaver et al., 2001), CLIMBER-2.4 (Petoukhov et al., 2000), LOVECLIM (Goosse et al., 2010) and ocean sediment models (DCESS (Shaffer et al., 2008), UVic 2.9 (Weaver et al., 2001), Bern3D-LPJ (Ritz et al., 2011)). These additional interactive components provide critical feedbacks for the ice-sheet response for sea-level rise estimates and for the oceanic carbon-cycle to estimate the appropriate carbon cycle response on the millennial time-scales (Zickfeld et al., submitted).

Table 9.5: Model response metrics for EMICs in Table 9.3. TCR_{2X}, TCR_{4X}, and ECS_{4X} are the changes in global average model surface air temperature from the decades centered at years 70, 140, and 995 respectively, from the

idealized 1% increase to 4 × CO_2 experiment. The ocean heat uptake efficiency, κ_{4X}, is calculated from the global average heat flux divided by TCR_{4X} for the decade centered at year 140, from the same idealized experiment. ECS_{2x} was calculated from the decade centered about year 995 from a 2 × CO_2 pulse experiment. Data from Eby et al. (2012).

Model	TCR_{2X} (°C)	ECS_{2x}(°C)	TCR_{4X} (°C)	ECS_{4X} (°C)	κ_{4X} (W m^{-2} $°C^{-1}$)
B3	2.0	3.3	4.6	6.8	0.58
C2	2.1	3.0	4.7	5.8	0.84
C3	1.9	3.2	4.5	5.9	0.93
DC	2.1	2.8	3.9	4.8	0.72
FA	2.3	3.5	5.2	8.0	0.55
GE	2.5	4.0	5.4	7.0	0.51
IA	1.6	--	3.7	4.3	--
I2	1.5	1.9	3.7	4.5	--
LO	1.2	2.0	2.1	3.5	1.17
ME	2.4	3.7	5.3	6.9	0.55
MI	1.6	2.4	3.6	4.6	0.66
ML	1.6	2.8	3.7	5.5	1.00
SP	0.8	3.6	2.9	5.2	0.84
UM	1.6	2.2	3.2	4.3	--
UV	1.9	3.5	4.3	6.6	0.92
EMIC mean	1.8	3.0	4.0	5.6	0.8
EMIC range	0.8 to 2.5	1.9 to 4.0	2.1 to 5.4	3.5 to 8.0	0.5 to 1.2

[INSERT TABLE 9.6 HERE]

Table 9.6: Features of Earth System Models of Intermediate Complexity (EMICs)

9.7.2.3.2 Climate sensitivity for the Last Glacial Maximum

Climate sensitivity can also be explored in another climatic context. The AR4 reported on attempts to relate the simulated LGM changes in tropical SST to global climate sensitivity (Hegerl et al., 2007; Knutti and Hegerl, 2008). LGM temperature changes in the tropics and in Antarctica have been shown to scale well with climate sensitivity (Hargreaves et al., 2007), because the signal is mostly dominated by CO_2 forcing in these regions (Braconnot et al., 2007b; Jansen et al., 2007). The analogy between the LGM climate sensitivity and future climate sensitivity is, however, not perfect (Crucifix, 2006). In an ensemble of MIROC simulations, the magnitudes of the LGM cooling and the warming induced by a doubling of CO_2 are nonlinear in the forcings applied to generate these altered climates (Hargreaves et al., 2007). Differences in the cloud radiative feedback are at the origin of this asymmetric response to equivalent positive and negative forcings (Yoshimori et al., 2009). There is thus still low confidence that the regional LGM model-data comparisons can be used to evaluate model climate sensitivity. However, even if the results do not scale perfectly with equilibrium or transient climate sensitivity, the LGM simulations allow the identification of the different feedback factors that contributed to the LGM global cooling (Yoshimori et al., 2011) and model spread in these feedbacks. The largest spread in LGM model feedbacks is found for the shortwave cloud feedback, just as for the modern climate. This correspondence between LGM and modern climates adds to the high confidence that the shortwave cloud feedback as the dominant source of model spread in climate sensitivity.

9.7.3.3 Constraints on equilibrium climate sensitivity from climate-model ensembles and observations

Several modelling groups have performed perturbed physics ensembles (PPE) to sample the parametric, and to some degree structural uncertainty (e.g., by switching between alternative parameterizations). The idea is to explore the range of possible model responses and to find relationships between model parameters and the simulated climate. Such relationships, if they exist, can be used to constrain model parameters, climate sensitivity, the transient response, or the regional response of any variable based on observations.

The perturbation of atmospheric and surface albedo feedbacks in the Hadley Centre model leads to ranges of feedbacks and sensitivities much larger than the CMIP range (Collins et al., 2011a; Knight et al., 2007; Piani et al., 2005; Sanderson et al., 2008a; Sanderson et al., 2008c; Stainforth et al., 2005a). On the other hand, the range covered in the ECHAM and NCAR CCSM model are only about 2–5°C and 2.2–3.2°C, respectively (Klocke et al., 2011a; Sanderson, 2011a, 2011b). A PPE based on perturbed land parameters yields an even narrower range of 0.5°C for climate sensitivity (Fischer et al., 2011), indicating that those parameters do not strongly control climate sensitivity. The spread in other variables and on small scales is of course larger. Relationships between climatological quantities and climate sensitivity are often found within a specific PPE, but in many cases the relationship is not robust across PPEs from different models or in CMIP3 (Klocke et al., 2011a; Knutti et al., 2006; Masson and Knutti, 2012; Rougier et al., 2009b; Sanderson, 2011b; Yokohata et al., 2010b). This implies that the model structure underlying a PPE is important, and that a single PPE is probably insufficient to constrain the climate sensitivity in the real world. Feedbacks related to clouds and the water cycle are found to be particularly important for the spread of climate sensitivities (Rougier et al., 2009b; Sanderson et al., 2008c; Webb et al., 2006; Yokohata et al., 2010b). Relationships between observables and climate sensitivities in CMIP3 based on interannual variability (Wu and North, 2003), the seasonal cycle (Knutti et al., 2006; Wu et al., 2008a) and the regional radiation budget (Huber et al., 2011) are generally weaker because models' structural uncertainty is also sampled. A few studies point to higher sensitivities being more realistic (Fasullo and Trenberth, 2012; Shukla et al., 2006), but most studies are unable to narrow the range of climate sensitivities significantly if the observational uncertainty and the imperfect functional relationships are properly accounted for. Several studies in fact point to a most likely value near 3°C that is similar to the CMIP3 mean (Huber et al., 2011; Volodin, 2008a; Wu et al., 2008a) and most studies indicate a likely or very likely range of about 2°C–5°C similar to the CMIP range (Huber et al., 2011; Wu et al., 2008a).

The large scale climatological information available has so far been insufficient to constrain model behaviour to a range tighter than CMIP3, at least on a global scale. Sanderson and Knutti (2012b) suggest that much of the available and commonly used large scale observations have already been used to develop and evaluate models and are therefore of limited value to further constrain climate sensitivity or TCR. The assessed literature suggests that the range of climate sensitivities and transient responses covered by CMIP3/5 cannot be narrowed significantly by constraining the models with observations of the mean climate and variability, consistent with the difficulty of constraining the cloud feedbacks from observations (see Chapter 7). Studies based on PPE and CMIP3 support the conclusion that a credible representation of the mean climate and variability is very difficult to achieve with equilibrium climate sensitivities below 2°C (Fasullo and Trenberth, 2012; Huber et al., 2011; Klocke et al., 2011a; Piani et al., 2005; Sanderson et al., 2008a; Sanderson et al., 2008c; Stainforth et al., 2005a). High climate sensitivity values above 5°C (in some cases above 10°C) are found in the PPE based on HadAM/HadCM3. Several recent studies find that such high values cannot be excluded based on climatological constraints, but are much less likely than values in the consensus range of 2–4.5°C (Knutti et al., 2006; Piani et al., 2005; Rodwell and Palmer, 2007; Sanderson et al., 2008a; Sanderson, 2011a, 2011b; Sanderson et al., 2010; Sanderson et al., 2008c). An overall assessment of climate sensitivity and transient response is given in Box 12.2. Observational constraints on the transient climate response from the observed warming over the last century are discussed in Section 10.9.1 and shown in Box 12.2, Figure 2.

9.8 Relating Model Performance to Credibility of Model Applications

9.8.1 Overall Assessment of Model Performance

This chapter has quantitatively assessed the performance of individual climate models as well as the multi-model mean. In addition, changes between models available now and those that were assessed in AR4 have been documented. The models display a range of abilities to simulate essential climate variables, underlying key processes, and climate phenomena such as monsoon and ENSO. The range in model quality is apparent both when different performance metrics are applied to a single model and when the same performance metric is applied to different models. No model scores high or low in all performance metrics, but some models perform substantially better than others for specific climate variables or phenomena. For a few climate characteristics, the assessment has shown that some classes of models, e.g., those with higher horizontal resolution, higher model top or a more complete representation of the carbon cycle, aerosols or chemistry, agree better with observations, although this is not universally true.

Figure 9.45 provides a synthesis of key model evaluation results from this chapter. The figure makes use of the calibrated language as defined in Mastrandrea et al. 2011). The y-axis refers to the level of confidence which increases towards the right as suggested by the increasing strength of shading. The level of confidence is a combination of the level of evidence and the level of agreement. The level of evidence includes the number of studies and quality of observational data. Generally, evidence is most robust when there are multiple, independent studies that evaluate multiple models using high-quality observations. The level of agreement measures whether different studies come to the same conclusions or not. The figure shows that several important aspects of the climate are simulated well by contemporary models, with varying levels of confidence. The color coding provides an indication of how model quality has changed from CMIP3 to CMIP5. For example, there is very high confidence that global mean surface air temperature (SAT) is well simulated, since SAT has been evaluated in multiple studies using high-quality observations. It is shown in orange because there is limited evidence of improvements since CMIP3. By contrast, the diurnal cycle of SAT is also simulated well but the confidence is only medium, owing to as yet limited analyses. A description that explains the expert judgment for each of the results presented in Figure 9.45 can be found in the body of this chapter, with a link to the specific sections given in the figure caption.

[INSERT FIGURE 9.45 HERE]

Figure 9.45: Summary of the findings of Chapter 9 with respect to how well the CMIP5 models simulate important features of the climate of the 20th century. Confidence in the assessment increases towards the right as suggested by the increasing strength of shading. Model quality increases from bottom to top. The color coding indicates improvements from CMIP3 (or models of that generation) to CMIP5. The assessment is mostly based on the multi-model mean, not excluding that deviations for individual models could exist. Note that assessed model quality is simplified for representation in the figure and it is referred to the text for details of each assessment. The figure highlights the following key features, with the sections that back up the assessment added in brackets:

PANEL a:

AMOC	Atlantic Meridional Overtuning Circulation mean (Section 9.4.2.3)
AntSIE	Annual cycle Antarctic sea ice extent (Section 9.4.3)
ArctSIE	Annual cycle Arctic sea ice extent (Section 9.4.3)
Blocking	Blocking events (Section 9.5.2.2)
CLD	Clouds (Section 9.4.1.1.2)
fgCO2	Global ocean carbon sink (Section 9.4.5.3)
fgCO2-sp	Spatial pattern of ocean-atmosphere CO_2 fluxes (Section 9.4.5.3)
MHT	Meridional heat transport (Section 9.4.2.4)
Monsoon	Global monsoon (Section 9.5.2.3)
NBP	Global land carbon sink (Section 9.4.5.3)
NBP-sp	Spatial pattern of land-atmosphere CO_2 fluxes
PR	Large scale precipitation (Section 9.4.1)
PR-diur	Diurnal cycle precipitation (Section 9.5.2.2)
PR-RS	Regional scale precipitation (Section 9.6.1.1)
SAF	Snow albedo feedbacks (Section 9.8.3)
SSS:	Sea surface salinity (Section 9.4.2.1)
SST:	Sea surface temperature (Section 9.4.2.1
TAS:	Large scale surface air temperature (Section 9.4.1)
TAS-diur	Diurnal cycle surface air temperature (Section 9.5.2.1)
TAS-RS	Regional scale surface air temperature (Section 9.6.1.1)
TropO3	Tropospheric column ozone climatology (Section 9.4.1.3.5)
TrAtlantic	Tropical Atlantic / Pacific mean state (Section 9.4.2.5)
TrInOcean	Tropical Indian Ocean mean state (Section 9.4.2.2.5)
TrPacific	Tropical Pacific mean state (Section 9.4.2.5)
WBC	Western boundary current (Section 9.4.2.3)
ZTaux	Zonal mean zonal wind stress (Section 9.4.2.4)

PANEL b (Trends)

AntSIE-t:	Trend in Antarctic sea ice extent (Section 9.4.3)
ArctSIE-t:	Trend in Arctic sea ice extent (Section 9.4.3)
LST-t	Lower stratospheric temperature trends (Section 9.4.1.3.5)
OHC-t	Global ocean heat content trends (Section 9.4.2.2)
StratO3-t	Total column ozone trends (Section 9.4.1.3.5)
TAS-t	Surface air temperature trends (Section 9.4.1)

UTT-t	Upper tropospheric temperature trends (Section 9.4.1.3.2)
PANEL c (Variability)	
AMM	Atlantic Meridional Mode (Section 9.5.3.3)
AMOC-var	Atlantic Meridional Overtuning Circulation (Section 9.5.3.3)
AMV	Atlantic Multi-decadal Variability (Section 9.5.3.3)
AN	Atlantic Niño (Section 9.5.3.3)
CO2-iav	Interannual variability of atmospheric CO2 (Section 9.4.5)
ENSO	El Niño Southern Oscillation (Section 9.5.3.4)
ENSOtele	Tropical ENSO teleconnections (Section 9.5.3.5)
IOD	Indian ocean dipole (Section 9.5.3.4)
IPO	Indian ocean dipole (Section 9.5.3.4)
MJO	Madden Julian Oscillation (Section 9.5.2.2)
NAO	North Atlantic Oscillation (Section 9.5.3.2)
OHC-var	Global ocean heat content variability (Section 9.4.2.2)
PDO	Pacific Decadal Oscillation (Section 9.5.3.4)
PNA	Pacific North American (Section 9.5.3.5)
QBO	Quasi-Biennal Oscillation (Section 9.5.3.5)
SAM	Southern Annual Mode (Section 9.5.3.2)
SST-var	Global sea surface temperature variability (Section 9.5.3.1)
PANEL d (Extremes):	
Hurric-hr	Year-to-year counts of Atlantic hurricanes in high-resolution AGCMs (Section 9.5.4.3)
PR_ext	Global distributions of precipitation extremes (Section 9.5.4.2)
PR_ext-hr	Global distribution of precipitation extremes in high-resolution AGCMs (Section 9.5.4.2)
PR_ext-t	Global trends in precipitation extremes (Section 9.5.4.2)
TAS_ext	Global distributions of surface air temperature extremes (Section 9.5.4.1)
TAS_ext-t	Global trends in surface air temperature extremes (Section 9.5.4.1)
TC-hr	Tropical cyclone tracks and intensity in high-resolution AGCMs (Section 9.5.4.3)
Droughts	Droughts (Section 9.5.4.3)

9.8.2 Implications of Model Evaluation for Climate Change Detection and Attribution

Climate models are developed, tuned and evaluated based on historical climate conditions. Their evaluation is therefore of direct relevance to detection and attribution studies using a combination of observations and model results to identify the anthropogenic signature in the climate record. The key for detection/attribution studies is that models accurately reproduce climate variability and patterns of response to forcing, while certain biases in magnitude or basic state are less important. For instance, the detection and attribution study by Santer et al. (2009a) showed that the anthropogenic water vapour fingerprint is rather insensitive to model uncertainties, and is governed by basic physical processes that are well-represented in CMIP3 models. Other biases like the warm bias of lower stratosphere temperature trends during the satellite period can be linked to uncertainties in stratospheric ozone forcing (Santer et al., 2012; Solomon et al., 2012b).

Several statements in the AR4 Chapter 9 (Hegerl et al., 2007) indicated that anthropogenic changes were detected on temperature rather accurately at large scale and continental scale, but that there was less confidence in the understanding of forced changes in other variables such as surface pressure and precipitation. This is directly linked to the ability of climate models to reproduce the large scale pattern and variability in these variables as discussed in (Hegerl and Zwiers, 2011). Figure 9.45 shows that this is still valid with high confidence. However some improvement can be noted. In particular 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. The latter was estimated to account for at least half of the inter-model spread in projected climate trends during 2005–2060 in the CMIP3 multi-model ensemble (Deser et al., 2011a). This together with the larger ensembles of simulations in CMIP5 provides evidence that more robust estimates of detection and attribution of human activity on climate can be expected. Methods based on ‘fingerprint ‘are sensitive to uncertainties in the pattern of the response to forcing. The better representation of some of the climate phenomena and modes discussed is Sections 9.5.2 and 9.5.3 suggest that improved detections and attribution results can also be expected at regional scale, although models still suffer from major biases in the representation of the intraseasonal to multidecadal variability that systematically affects regional features. Recent studies of climate extremes also provide evidence that models have reasonable skill in these important attributes of a changing climate, however, there is an indication that models have difficulties in reproducing the right balance between historical changes in cold and warm extremes..They also confirm that resolution affects the confidence that can be placed in the analyses of extreme in precipitation.

9.8.3 Implications of Model Evaluation for Model Projections of Future Climate

The ability of a climate model to make future climate projections cannot be directly evaluated, and so our confidence in model projections must be built upon the knowledge that these models are based on physical principles, and a demonstration of how well models represent a wide range of processes on various spatial and temporal scales (Eyring et al., 2005; Knutti et al., 2010b), especially those related to important feedbacks in the Earth's climate system. A climate model's credibility in representing key processes is increased if the model is able to simulate past variations in climate (e.g., trends over the 20th century, paleoclimatic changes) and the associated covariations of different climate quantities. However, unlike shorter lead forecasts, climate change projections push models into unknown territory since the anticipated future greenhouse gas forcing is outside the range observed in the historical period used for evaluation. The following two subsections assess new approaches on the relation between model performance and projections. The relation between downscaling performance to the credibility of downscaling applications was assessed earlier in Section 9.6.6.

9.8.3.1 The Multi-Model Ensemble Approach

Projections of climate are constructed from Multi-Model Ensembles (MME) of climate model simulations, for example from CMIP5. The models differ in complexity as discussed in Section 9.1, and so the ensemble is not homogeneous. The most common approach is to estimate the climate change response by calculating the arithmetic mean of the individual mean climate model responses, referred to as unweighted Multi-Model Mean (MMM). This approach of "one vote per model" gives equal weight to each climate model regardless of how many simulations each model has contributed, how interdependent the models are or how well each model has fared in objective evaluation. Some climate models share a common lineage and so share common biases (Annan and Hargreaves, 2011; Frame et al., 2005; Jun et al., 2008b; Knutti, 2010; Knutti et al., 2012; Knutti et al., 2010a; Pennell and Reichler, 2011b; Tebaldi and Knutti, 2007b). As a result, collections such as the CMIP5 MME cannot really be considered a random sample of independent models. This complexity creates challenges for how best to make quantitative inferences of future climate climate as discussed further in Chapter 12 (Collins et al., 2012b; Knutti et al., 2010a; Sanderson and Knutti, 2012a; Stephenson et al., 2012) as discussed further in Chapter 12.

Despite the lack of independence, Annan and Hargreaves (2010) have proposed a 'rank histogram' approach to evaluate model ensembles as a whole, rather than individual models, by diagnosing whether observations can be considered statistically indistinguishable from a model ensemble. Studies based on this approach have suggested that MMEs (CMIP3/5) are 'reliable' in that they are not too narrow or too dispersive as a sample of possible models, but existing single-model-based ensembles tend to be too narrow (Yokohata et al., 2012a; Yokohata et al., 2012b). This "reliability test" highlights an advantage of the use of multiple model ensembles for exploring uncertainty and introduces a metric for evaluating the respective ensembles. Although the analysis was done on the current mean climate state, this has implications for ensembles of future projections, but further work is required to study the exact relationships between simulation errors and uncertainties in projections (Collins et al., 2012a).

An open question is whether climate projections can be improved by weighting of models according to their ability to reproduce past observed climate. Several studies have explored the use of unequally weighted means with the weights based on the models' performance at simulating past variations in climate typically using some performance metric or collection of metrics to construct the weights (Abe et al., 2011; Christensen et al., 2010; Connolley and Bracegirdle, 2007; Knutti et al., 2010b; Murphy et al., 2007; Pierce et al., 2009; Raisanen et al., 2010; Shiogama et al., 2011; Watterson and Whetton, 2011; Waugh and Eyring, 2008). Generally, these studies have found only small differences between the weighted and unweighted multi-model means, and so demonstrating the advantage of a weighting scheme remains difficult. However, when applied to projections of Arctic sea ice, averages in which extra weight is given to models with the most realistic historical sea ice do give different results than the unweighted mean (Massonnet et al., 2012; Scherrer, 2011; Stroeve et al., 2007; Stroeve et al., 2012; Wang and Overland, 2012), see further discussion in Section 12.4.6.1.

9.8.3.2 Emergent Constraints

There are several encouraging examples of "emergent constraints", where variations between model projections are found to be related to inter-model variations in mean climate, past trends, or seasonal variability (Boe et al., 2009a; Boe et al., 2009b; Bracegirdle and Stephenson, 2012a; Eyring et al., 2007; Hall and Qu, 2006; Huber et al., 2011; Mahlstein and Knutti, 2010; Qu and Hall, 2012; Schaller et al., 2011). For example, analyzing the CMIP3 ensemble, Hall and Qu 2006) showed that inter-model variations of snow albedo feedback in the contemporary seasonal cycle are strongly correlated with comparably large inter-model variations in this feedback under future climate change. Qu and Hall (2012) performed the same analysis for the CMIP5 ensemble, and found nearly identical results (Figure 9.46, left panel). The relationship likely arises from the fact that surface albedo values in areas covered by snow vary widely across the models, particularly in the heavily-vegetated boreal forest zone. Models with higher surface albedos in these areas have a larger contrast between snow-covered and snow-free areas, and hence a stronger snow albedo feedback whether the context is the seasonal variation in sunshine or anthropogenic forcing. It is possible to estimate the actual climate's snow albedo feedback in the context of the seasonal cycle, and a comparison with models reveals most of them to be biased. If these seasonal cycle biases were reduced, there would be a nearly proportionate reduction in the spread of the manifestation of snow albedo feedback in future climate projections. This would be an improvement in overall ensemble quality because the spread reduction would occur for reasons consistent with well-documented climate physics. Of course, this process, once complete, would not guarantee the models are completely free of errors in snow albedo feedback, especially if the models share some common structural bias. It would only ensure errors made obvious by analyses such as that shown in the left panel of Figure 9.46 have been eliminated.

The right panel of Figure 9.46 shows another example of an emergent constraint, where the sensitivity of tropical land carbon to warming (i.e., without CO_2 fertilization effects) is related to the sensitivity of interannual variability of the growth-rate of atmospheric CO_2 to annual mean tropical temperature (30°N–30°S). The horizontal axis is the regression of the atmospheric CO_2 growth-rate on the tropical temperature anomaly, for each model. The strong statistical relationship between these two variables is consistent with the fact that interannual variability in the CO_2 growth-rate is known to be dominated by the response of tropical land to climatic anomalies, associated particularly with ENSO. Thus the relationship has a physical as well as a statistical basis. The interannual sensitivity of the CO_2 growth-rate to tropical temperature can be estimated from observational data. Like the snow albedo feedback example, this intermodel relationship therefore provides a credible means to reduce model spread in the sensitivity of tropical land carbon to tropical climate change.

[INSERT FIGURE 9.46 HERE]

Figure 9.46: *Left:* Scatter plot of simulated springtime snow albedo feedback ($\Delta\alpha_s/\Delta T_s$) values in climate change (y-axis) versus simulated springtime $\Delta\alpha_s/\Delta T_s$ values in the seasonal cycle (x-axis) in transient climate change experiments from 17 CMIP3 (blue) and 25 CMIP5 models (α_s and T_s are surface albedo and surface air temperature, respectively). Adapted from Hall and Qu (2006) and Qu and Hall (2012). *Right:* Constraint on the climate sensitivity of land carbon in the tropics (30°N–30°S) from interannual variability in the growth-rate of global atmospheric CO_2 (Cox et al., 2012). This is based on results from ESMs with free-running CO_2; C^4MIP GCMs (black labels, (Friedlingstein et al., 2006)), and three land carbon "physics ensembles" with HadCM3 (red labels, (Booth et al., 2012)). The y-axis is calculated over the period 1960-2099 inclusive, and the y-axis is calculated over the period 1960-2010 inclusive. In both cases the temperature used is the mean (land+ocean) temperature over 30°N–30°S. The vertical grey band shows the estimated sensitivity of the observed global CO_2 growth-rate to the observed tropical mean temperature.

On the other hand, many studies have failed to find strong relationships between observables and projections. Whetton et al. (2007a) and Knutti et al. (2010a) found that correlations between local to regional climatological values and projected changes are small except for a few regions. 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. Räisänen et al. (2010) report only small (10–20%) reductions in cross-validation error of simulated 21st century temperature changes when weighting the CMIP3 models based on their simulation of the present-day climatology. They note that the effects of the weighting on real-world temperature projections are sensitive to the predictor variable. The main difficulty in constraining AOGCMs with climatological data to a range much narrower than that covered by the CMIP ensemble is measurement uncertainties, sparse coverage in many observed variables, short time series for observed trends, lack of correlation between observed quantities and projected past or future trends (Jun et al., 2008a; Knutti, 2010; Knutti et al., 2010a; Tebaldi and Knutti, 2007a), the ambiguity of possible performance metrics and the difficulty of associating them with predictive skill (Gleckler et al., 2008; Knutti et al., 2010a; Parker et al., 2007; Pierce et al., 2009; Pincus et al., 2008a; Reichler and Kim, 2008b)

Emergent constraints can be difficult to identify if climate models are structurally similar and share common biases, thereby reducing the effective ensemble size. Comparison of emergent constraints in MMEs from different modeling experiments can help reveal which constraints are robust (Bracegirdle and Stephenson, 2012b; Massonnet et al., 2012). Another issue is that multiple testing of large numbers of predictors will find statistically significant correlations that do not remain significant in a different ensemble. This is particularly important if many predictors are tested using only small ensembles like CMIP3 (DelSole and Shukla, 2009; Huber et al., 2011; Masson and Knutti, 2012; Raisanen et al., 2010). All of these potential pitfalls underscore the need for analysis of the mechanism underpinning the statistical relationship between current and future climate parameters for any proposed emergent constraint.

[START FAQ 9.1 HERE]

FAQ 9.1: Are Climate Models Getting Better, and How Would We Know?

Climate models are extremely sophisticated pieces of software that simulate, with as much fidelity as possible, the complex interactions between the atmosphere, ocean, land surface, snow and ice, the global ecosystem, and a variety of chemical and biological processes.

Complexity in such models has certainly increased since the first IPCC Assessment in 1990, so in that sense, current Earth system models are vastly 'better' than the models in use then. More powerful computers allow current models to resolve detail to much finer scale. Today's models have also benefitted from the past two decades of research into various climate processes, more comprehensive observations, and improved scientific understanding.

Climate models of today are, in principle, better than their predecessors. However, every bit of added complexity also introduces new sources of error (e.g., via uncertain parameters) and new interactions between model components that may, if only temporarily, degrade a model's simulation of other aspects of the climate system.

An important consideration is that model performance can only be evaluated relative to past observations. To have confidence in the future projections of such models, historical climate—and its variability and change—must be well-simulated. But this alone may not sufficient. Whereas weather and seasonal climate predictions can be regularly verified, climate projections spanning a century or more cannot, particularly as anthropogenic forcing is driving the climate system toward conditions not previously observed in the instrumental record. This will always be a limitation.

Quantifying model performance is a topic that has featured in all previous IPCC Working Group I Reports. Reading back over these earlier assessments provides a general sense of the improvements that have been made. Past reports have typically provided a rather broad survey of model performance, showing differences between model-calculated versions of various climate quantities and corresponding observational estimates.

Inevitably, some models perform better than others for certain climate variables, but no individual model clearly emerges as 'the best' overall. Recently, there has been progress in computing various performance metrics, which synthesise model performance relative to a range of different observations according to a simple numerical score. Of course, the definition of such a score, how it is computed, the observations used (which are themselves uncertain to some extent), and the manner in which various scores are combined are all important, and will affect the end result.

[INSERT FIGURE FAQ 9.1, FIGURE 1 HERE]

FAQ 9.1, Figure 1: Model error in simulating annual mean temperature and precipitation as produceed in the three recent phases of the Coupled Model Intercomparison Project (CMIP2, CMIP3 and CMIP5). The upper portion of the figure shows the root mean squared error (a measure of local discrepancies between model and observation) that has been normalized by the observational standard deviation to allow comparison across variables. Each symbol represents the result from a particular model. Larger values indicate larger errors; smaller errors going from left to right indicate model improvement. Across the bottom of the figure, the sketches indicate evolution in model complexity and model resolution, both of which have improved from CMIP2 to CMIP5. [PLACEHOLDER FOR FINAL DRAFT: to be updated as more results from CMIP5 simulations are analyzed and a consistent observational reference is used in the error calculation.] Redrafted from Gleckler et al. (2008) and updated with CMIP5 results.

Nevertheless, if the metric is computed consistently, one can compare different generations of models. Evaluation against a 'performance index' has shown a steady improvement in models participating in the series of Coupled Model Intercomparison Projects: CMIP1 included models from the mid 1990s; CMIP2 included models from around 2000; and CMIP3 from about 2005.

These results showed that, although each generation exhibited a range in model performance, the average model performance index improved steadily between each generation. An example of changes model performance over time is shown in FAQ 9.1, Figure 1, and illustrates the ongoing, albeit modest, improvement. Reasons for this improvement include increased understanding of various climate processes, and better representation of these processes, and their interactions, in climate models. More comprehensive Earth observations are also driving improvemtns.

So, yes, climate models are getting better, and we can demonstrate this with quantitative performance metrics based on historical observations. Although future climate projections cannot be directly evaluated, climate models are based on verifiable physical principles and are able to reproduce many important aspects of past response to external forcing. In this way, they provide a scientifically sound preview of the climate to come.

[END FAQ 9.1 HERE]

References

Abe, M., H. Shiogama, T. Nozawa, and S. Emori, 2011: Estimation of future surface temperature changes constrained using the future-present correlated modes in inter-model variability of CMIP3 multimodel simulations. *Journal of Geophysical Research-Atmospheres*, **116**.

Abiodun, B., W. Gutowski, A. Abatan, and J. Prusa, 2011: CAM-EULAG: A Non-Hydrostatic Atmospheric Climate Model with Grid Stretching. *Acta Geophysica*, **59**, 1158-1167.

Abramowitz, G., R. Leuning, M. Clark, and A. Pitman, 2008: Evaluating the Performance of Land Surface Models. *Journal of Climate*, **21**, 5468-5481.

AchutaRao, K., and K. Sperber, 2002: Simulation of the El Niño Southern Oscillation: results from the coupled model intercomparison project. *Clim. Dyn.*, **19**, 191-209.

——, 2006: ENSO simulations in coupled ocean-atmosphere models: are the current models better? *Clim. Dyn.*, **27**, 1-16.

AchutaRao, K. M., et al., 2007: Simulated and observed variability in ocean temperature and heat content. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 10768-10773.

Achuthavarier, D., V. Krishnamurthy, B. P. Kirtman, and B. H. Huang, 2012: Role of the Indian Ocean in the ENSO-Indian Summer Monsoon Teleconnection in the NCEP Climate Forecast System. *Journal of Climate*, **25**, 2490-2508.

Ackerley, D., E. J. Highwood, and D. J. Frame, 2009: Quantifying the effects of perturbing the physics of an interactive sulfur scheme using an ensemble of GCMs on the climateprediction.net platform. *Journal of Geophysical Research-Atmospheres*, **114**.

Adachi, Y., et al., 2012: Simulations of climate and atmospheric chemistry and carbon cycle for the mid-19th century through the end of the 21st century by a new earth system model: MRI-ESM1. *Papers in Meteorology and Geophysics*, **submitted**.

Adam, J., E. Clark, D. Lettenmaier, and E. Wood, 2006: Correction of global precipitation products for orographic effects. *Journal of Climate*, **19**, 15-38.

Adkins, J. F., K. McIntyre, and D. P. Schrag, 2002: The salinity, temperature, and delta O-18 of the glacial deep ocean. *Science*, **298**, 1769-1773.

Adler, R. F., et al., 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, **4**.

Alekseev, V. A., E. M. Volodin, V. Y. Galin, V. P. Dymnikov, and V. N. Lykossov, 1998: Modeling of the present-day climate by the atmospheric model of INM RAS DNM GCM. Description of the model version A5421 and results of AMIP2 simulations.

Alessandri, A., P. G. Fogli, M. Vichi, and N. Zeng, 2012: Strengthening of the hydrological cycle in future scenarios: atmospheric energy and water balance perspective. *Earth Syst. Dynam. Discuss.*, **3**, 523-560.

Alexander, M. J., and K. H. Rosenlof, 1996: Nonstationary gravity wave forcing of the stratospheric zonal mean wind. *Journal of Geophysical Research-Atmospheres*, **101**, 23465-23474.

Alexander, M. J., et al., 2010: Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. *Quarterly Journal of the Royal Meteorological Society*, **136**, 1103-1124.

Alexandru, A., R. de Elia, and R. Laprise, 2007: Internal variability in regional climate downscaling at the seasonal scale. *Monthly Weather Review*. doi:DOI 10.1175/MWR3456.1, 3221-3238.

Alexandru, A., R. de Elia, R. Laprise, L. Separovic, and S. Biner, 2009: Sensitivity Study of Regional Climate Model Simulations to Large-Scale Nudging Parameters. *Monthly Weather Review*. doi:DOI 10.1175/2008MWR2620.1, 1666-1686.

Allan, R. P., 2009: Examination of Relationships between Clear-Sky Longwave Radiation and Aspects of the Atmospheric Hydrological Cycle in Climate Models, Reanalyses, and Observations. *Journal of Climate*, **22**, 3127-3145.

Allan, R. P., and B. J. Soden, 2008: Atmospheric warming and the amplification of precipitation extremes. *Science*, **321**, 1481-1484.

Allan, R. P., M. A. Ringer, and A. Slingo, 2003: Evaluation of moisture in the Hadley Centre climate model using simulations of HIRS water-vapour channel radiances. *Quarterly Journal of the Royal Meteorological Society*, **129**, 3371-3389.

Allan, R. P., A. Slingo, S. F. Milton, and M. E. Brooks, 2007: Evaluation of the Met Office global forecast model using Geostationary Earth Radiation Budget (GERB) data. *Quarterly Journal of the Royal Meteorological Society*, **133**, 1993-2010.

Allen, M., P. Stott, J. Mitchell, R. Schnur, and T. Delworth, 2000: Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*. 617-620.

An, S.-I., J.-W. Kim, S.-H. Im, B.-M. Kim, and J.-H. Park, 2011: Recent and future sea surface temperature trends in tropical pacific warm pool and cold tongue regions. *Climate Dynamics*. doi:10.1007/s00382-011-1129-7, 1-11.

Anav, A., et al., Submitted: Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth System Models. *Journal of Climate*.

Anderson, B. T., J. R. Knight, M. A. Ringer, C. Deser, A. S. Phillips, J. H. Yoon, and A. Cherchi, 2010: Climate forcings and climate sensitivities diagnosed from atmospheric global circulation models. *Climate Dynamics*, **35**, 1461-1475.

Andrews, T., and P. M. Forster, 2008: CO2 forcing induces semi-direct effects with consequences for climate feedback interpretations. *Geophysical Research Letters*, **35**.

Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor, 2012a: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters*, **39**.

——, 2012b: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters*, **39**, 7.

Annamalai, H., and K. R. Sperber, 2005: Regional heat sources and the active and break phases of boreal summer intraseasonal (30-50 day) variability. *Journal of the Atmospheric Sciences*, **62**, 2726-2748.

Annan, J., and J. Hargreaves, 2011: Understanding the CMIP3 Multimodel Ensemble. *Journal of Climate*, **24**, 4529-4538.

Annan, J. D., and J. C. Hargreaves, 2010: Reliability of the CMIP3 ensemble. *Geophysical Research Letters*, **37**, 5.

Annan, J. D., D. J. Lunt, J. C. Hargreaves, and P. J. Valdes, 2005: Parameter estimation in an atmospheric GCM using the Ensemble Kalman Filter. *Nonlinear Processes in Geophysics*, **12**.

Anstey, J. A., et al., 2012: Multi model analysis of Northern Hemisphere winter blocking and its relation to the stratosphere. *JGR-Atmos*, **submitted**.

Antonov, J. I., et al., 2010: World Ocean Atlas 2009, Volume 2: Salinity. U. S. G. P. O. NOAA Atlas NESDIS 69, Ed., 184 ..

Archer, D. E., G. Eshel, A. Winguth, W. Broecker, R. Pierrehumbert, M. Tobis, and R. Jacob, 2000: Atmospheric pCO(2) sensitivity to the biological pump in the ocean. *Global Biogeochemical Cycles*, **14**, 1219-1230.

Arneth, A., et al., 2010: From biota to chemistry and climate: towards a comprehensive description of trace gas exchange between the biosphere and atmosphere. *Biogeosciences*, **7**, 121-149.

Arora, V. K., and G. J. Boer, 2005: Fire as an interactive component of dynamic vegetation models. *Journal of Geophysical Research-Biogeosciences*, **110**.

——, 2010: Uncertainties in the 20th century carbon budget associated with land use change. *Global Change Biology*, **16**, 3327-3348.

Arora, V. K., et al., 2011a: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophysical Research Letters*, **38**, L05805, doi:05810.01029/02010GL046270.

——, 2011b: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophysical Research Letters*, **38**.

Arora, V. K., et al., 2009: The Effect of Terrestrial Photosynthesis Down Regulation on the Twentieth-Century Carbon Budget Simulated with the CCCma Earth System Model. *Journal of Climate*, **22**, 6066-6088.

Artale, V., et al., 2010: An atmosphere–ocean regional climate model for the Mediterranean area: assessment of a present climate simulation. *Climate Dynamics*, **35**, 721-740.

Arzhanov, M. M., P. F. Demchenko, A. V. Eliseev, and I. I. Mokhov, 2008: Simulation of Characteristics of Thermal and Hydrologic Soil Regimes in Equilibrium Numerical Experiments with a Climate Model of Intermediate Complexity. *Izvestiya Atmospheric and Oceanic Physics*, **44**, 548-566.

Assmann, K. M., M. Bentsen, J. Segschneider, and C. Heinze, 2010: An isopycnic ocean carbon cycle model. *Geoscientific Model Development*, **3**, 143-167.

Aumont, O., and L. Bopp, 2006: Globalizing results from ocean in situ iron fertilization studies. *Global Biogeochemical Cycles*, **20**.

Aumont, O., E. Maier-Reimer, S. Blain, and P. Monfray, 2003: An ecosystem model of the global ocean including Fe, Si, P colimitations. *Global Biogeochemical Cycles*, **17**.

Austin, J., and R. J. Wilson, 2006: Ensemble simulations of the decline and recovery of stratospheric ozone. *Journal of Geophysical Research-Atmospheres*, **111**.

Axelsson, P., M. Tjernström, S. Söderberg, and G. Svensson, 2011: An Ensemble of Arctic Simulations of the AOE-2001 Field Experiment. *Atmosphere*, **2**, 146-170.

Baehr, J., S. Cunnningham, H. Haak, P. Heimbach, T. Kanzow, and J. Marotzke, 2009: Observed and simulated estimates of the meridional overturning circulation at 26.5 N in the Atlantic. *Ocean Science Discussions*, **6**, 1333-1367.

Bailey, D., M. Holland, E. Hunke, B. Lipscomb, B. Briegleb, B. Bitz, and J. Schramm, 2010: Community ice code (cice) user's guide, version 4.0, released with cesm1.0. National Center for Atmospheric Research.

Balan Sarojini, B., et al., 2011: High frequency variability of the Atlantic meridional overturning circulation. *Ocean Science Discussions*, **8**, 219-246.

Baldocchi, D., et al., 2001: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem–Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities. *Bull. Amer. Meteorol. Soc.*, **82**, 2415-2434.

Baldwin, M. P., et al., 2001: The quasi-biennial oscillation. *Reviews of Geophysics*, **39**, 179-229.

Balmaseda, M. A., A. Vidard, and D. L. T. Anderson, 2008: The ECMWF Ocean Analysis System: ORA-S3. *Monthly Weather Review*, **136**, 3018-3034.

Bao, Q., G. Wu, Y. Liu, J. Yang, Z. Wang, and T. Zhou, 2010: An introduction to the coupled model FGOALS1.1-s and its performance in East Asia. *Advances in Atmospheric Sciences*, **27**.

Bao, Q., et al., 2012: The Flexible Global Ocean-Atmosphere-Land System model Version: FGOALS-s2. *Advances in Atmospheric Sciences*, **submitted**.

Barkstrom, B. R., 1984: The Earth Radiation Budget Experiment (ERBE). *Bull. Amer. Meteorol. Soc.*, **65**, 1170-1185.

Barnes, E. A., and D. L. Hartmann, 2010: Influence of eddy-driven jet latitude on North Atlantic jet persistence and blocking frequency in CMIP3 integrations. *Geophys. Res. Lett.*, **37**, L23802-.

Barnes, E. A., J. Slingo, and T. Woollings, 2012: A methodology for the comparison of blocking climatologies across indices, models and climate scenarios. *Climate Dynamics*, **38**, 2467-2481.

Barnier, B., et al., 2006: Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. *Ocean Dyn.*, **56**, 543-567.

Barriopedro, D., R. Garcia-Herrera, and R. M. Trigo, 2010a: Application of blocking diagnosis methods to General Circulation Models. Part I: a novel detection scheme. *Climate Dynamics*, **35**, 1373-1391.

Barriopedro, D., R. Garcia-Herrera, J. F. Gonzalez-Rouco, and R. M. Trigo, 2010b: Application of blocking diagnosis methods to General Circulation Models. Part II: model simulations. *Climate Dynamics*, **35**, 1393-1409.

Bartlein, P. J., et al., 2010a: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis. *Climate Dynamics*. doi:10.1007/s00382-010-0904-1.

Bartlein, P. J., et al., 2010b: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis. *Climate Dynamics*. doi:10.1007/s00382-010-0904-1.

Bathiany, S., M. Claussen, V. Brovkin, T. Raddatz, and V. Gayler, 2010: Combined biogeophysical and biogeochemical effects of large-scale forest cover changes in the MPI earth system model. *Biogeosciences*, **7**, 1383-1399.

Bauer, H. S., V. Wulfmeyer, and L. Bengtsson, 2008a: The representation of a synoptic-scale weather system in a thermodynamically adjusted version of the ECHAM4 general circulation model. *Meteorology and Atmospheric Physics*, **99**, 129-153.

Bauer, S. E., D. Koch, N. Unger, S. M. Metzger, D. T. Shindell, and D. G. Streets, 2007: Nitrate aerosols today and in 2030: a global simulation including aerosols and tropospheric ozone. *Atmospheric Chemistry and Physics*, **7**.

Bauer, S. E., et al., 2008b: MATRIX (Multiconfiguration Aerosol TRacker of mIXing state): an aerosol microphysical module for global atmospheric models. *Atmospheric Chemistry and Physics*, **8**, 6003-6035.

Beare, R., et al., 2006: An Intercomparison of Large-Eddy Simulations of the Stable Boundary Layer. *Boundary-Layer Meteorology*, **118**, 247-272-272.

Bell, M. J., M. J. Martin, and N. K. Nichols, 2004: Assimilation of data into an ocean model with systematic errors near the equator. *Quarterly Journal of the Royal Meteorological Society*, **130**, 873-893.

Bellassen, V., G. Le Maire, J. F. Dhote, P. Ciais, and N. Viovy, 2010: Modelling forest management within a global vegetation model Part 1: Model structure and general behaviour. *Ecol. Model.*, **221**, 2458-2474.

Bellassen, V., G. le Maire, O. Guin, J. F. Dhote, P. Ciais, and N. Viovy, 2011: Modelling forest management within a global vegetation model-Part 2: Model validation from a tree to a continental scale. *Ecol. Model.*, **222**, 57-75.

Bellouin, N., J. Rae, A. Jones, C. Johnson, J. Haywood, and O. Boucher, 2011: Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate. *J. Geophys. Res.*, **116**, 1–25.

Bellouin, N., O. Boucher, J. Haywood, C. Johnson, A. Jones, J. Rae, and S. Woodward, 2007: Improved representation of aerosols for HadGEM2.

Bellucci, A., S. Gualdi, and A. Navarra, 2010: The Double-ITCZ Syndrome in Coupled General Circulation Models: The Role of Large-Scale Vertical Circulation Regimes. *Journal of Climate*, **23**, 1127-1145.

Bender, F., H. Rodhe, R. Charlson, A. Ekman, and N. Loeb, 2006: 22 views of the global albedo - comparison between 20 GCMs and two satellites. *Tellus Series a-Dynamic Meteorology and Oceanography*, **58**, 320-330.

Bengtsson, L., and K. Hodges, 2011: On the evaluation of temperature trends in the tropical troposphere. *Climate Dynamics*, **36**, 419-430.

Bengtsson, L., K. I. Hodges, and N. Keenlyside, 2009: Will Extratropical Storms Intensify in a Warmer Climate? *Journal of Climate*, **22**, 2276-2301.

Bentsen, M., et al., 2012: The Norwegian Earth System 1 Model, NorESM1-M. Part 1: Description and basic evaluation *Geosci. Model Dev*, **submitted**.

Berckmans, J., T. Woollings, M.-E. Demory, and P.-L. Vidal, 2012: Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. submitted.

Bergengren, J., D. Waliser, and Y. Yung, 2011: Ecological sensitivity: a biospheric view of climate change. *Climatic Change*, **107**, 433-457.

Bergengren, J., S. Thompson, D. Pollard, and R. DeConto, 2001: Modeling global climate-vegetation interactions in a doubled CO2 world. *Climatic Change*, **50**, 31-75.

Bernie, D. J., E. Guilyardi, G. Madec, J. M. Slingo, S. Woolnough, and J. Cole, 2008: Impact od resolving the diurnal cycle in an ocean-atmosphere GCM. Part 2: A fiurnally coupled CGCM. *Climate Dynamics*, **31**, 909-925.

Bi, D., et al., 2012a: ACCESS-OM: the Ocean and Sea ice Core of the ACCESS Coupled Model. *Aust. Met. Oceanog. J.*, **submitted**.

Bi, D., et al., 2012b: The ACCESS Coupled Model: Description, Control Climate and Evaluation. *Aust. Met. Oceanog. J*, **submitted**.

Bitz, C. M., and W. H. Lipscomb, 1999: An energy-conserving thermodynamic sea ice model for climate study. *Journal of Geophysical Research. Oceans*, **104**, 15,669-615,677.

Bitz, C. M., J. C. Fyfe, and G. M. Flato, 2002: Sea ice response to wind forcing from AMIP models. *Journal of Climate*, **15**, 522-536.

Blanke, B., and S. Raynaud, 1997: Kinematics of the Pacific Equatorial Undercurrent: An Eulerian and Lagrangian Approach from GCM Results. *Journal of Physical Oceanography*, **27**, 1038-1053.

Block, K., and T. Mauritsen, 2012: Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO_2. *Journal of Advances in Modeling Earth Systems*. submitted.

Blyth, E., J. Gash, A. Lloyd, M. Pryor, G. P. Weedon, and J. Shuttleworth, 2010: Evaluating the JULES Land Surface Model Energy Fluxes Using FLUXNET Data. *Journal of Hydrometeorology*, **11**, 509-519.

Boberg, F., and J. H. Christensen, 2012: Overestimation of Mediterranean summer temperature projections due to model deficiencies. *Nature Clim. Change*, **2**, 433-436.

Boccaletti, G., R. Ferrari, and B. Fox-Kemper, 2007: Mixed layer instabilities and restratification. *Journal of Physical Oceanography*, **37**, 2228-2250.

Bodas-Salcedo, A., M. Webb, M. Brooks, M. Ringer, K. Williams, S. Milton, and D. Wilson, 2008: Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities. *Journal of Geophysical Research-Atmospheres*. doi:ARTN D00A13, 10.1029/2007JD009620, -.

Bodas-Salcedo, A., et al., 2011: COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteorol. Soc.*, **92**, 1023-1043.

Bodeker, G., H. Shiona, and H. Eskes, 2005: Indicators of Antarctic ozone depletion. *Atmospheric Chemistry and Physics*, **5**, 2603-2615.

Boe, J., A. Hall, and X. Qu, 2009a: September sea-ice cover in the Arctic Ocean projected to vanish by 2100. *Nature Geoscience*. doi:DOI 10.1038/NGEO467, 341-343.

Boe, J., A. Hall, and X. Qu, 2009b: Deep ocean heat uptake as a major source of spread in transient climate change simulations. *Geophysical Research Letters*. doi:ARTN L22701, 10.1029/2009GL040845, -.

Boe, J., L. Terray, F. Habets, and E. Martin, 2007: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, **27**, 1643-1655.

Boe, J. L., A. Hall, and X. Qu, 2009c: September sea-ice cover in the Arctic Ocean projected to vanish by 2100. *Nature Geoscience*, **2**, 341-343.

Boer, G., and S. Lambert, 2008: The energy cycle in atmospheric models. *Climate Dynamics*, **30**, 371-390.

Boer, G. J., and B. Yu, 2003: Climate sensitivity and climate state. *Climate Dynamics*, **21**, 167-176.

Boer, G. J., and K. Hamilton, 2008: QBO influence on extratropical predictive skill. *Climate Dynamics*, **31**, 987-1000.

Bonan, G. B., 2008: Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, **320**, 1444-1449.

Bond, T. C., et al., 2007: Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850-2000. *Global Biogeochemical Cycles*, **21**, 16.

Bondeau, A., P. C. Smith, S. Zaehle, S. Schaphoff, W. Lucht, W. Cramer, and D. Gerten, 2007: Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, **13**, 679-706.

Boning, C. W., A. Dispert, M. Visbeck, S. R. Rintoul, and F. U. Schwarzkopf, 2008: The response of the Antarctic Circumpolar Current to recent climate change. *Nature Geoscience*, **1**, 864-869.

Bony, S., and J. L. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models - art. no. L20806. *Geophysical Research Letters*, **32**, 20806-20806.

Boone, A., et al., 2009: THE AMMA LAND SURFACE MODEL INTERCOMPARISON PROJECT (ALMIP). *Bull. Amer. Meteorol. Soc.*, **90**, 1865-1880.

Boone, A. A., I. Poccard-Leclercq, Y. K. Xue, J. M. Feng, and P. de Rosnay, 2010: Evaluation of the WAMME model surface fluxes using results from the AMMA land-surface model intercomparison project. *Climate Dynamics*, **35**, 127-142.

Booth, Ben B. B., Dunstone, Nick J., Halloran, Paul R., Andrews, Timothy, & Bellouin, Nicolas. (2012). Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. Nature, 484(7393). doi: 10.1038/nature10946

Boschat, G., P. Terray, and S. Masson, 2012: Robustness of SST teleconnections and precursory patterns associated with the Indian summer monsoon. *Climate Dynamics*, **38**, 2143-2165.

Boyle, J., and S. A. Klein, 2010: Impact of horizontal resolution on climate model forecasts of tropical precipitation and diabatic heating for the TWP-ICE period. *J. Geophys. Res.*, **115**, D23113.

Boyle, J., S. Klein, G. Zhang, S. Xie, and X. Wei, 2008: Climate Model Forecast Experiments for TOGA COARE. *Monthly Weather Review*, **136**, 808-832.

Bracegirdle, T., et al., 2012: Assessment of surface winds over the Atlantic, Indian and Pacific Ocean sectors of the Southern Hemisphere in CMIP5 models: historical bias, forcing response, and state dependence., submitted ed., JGR-Atmospheres.

Bracegirdle, T. J., and D. B. Stephenson, 2012a: Higher precision estimates of regional polar warming by ensemble regression of climate model projections. *Climate Dynamics*. doi:10.1007/s00382-012-1330-3.

Bracegirdle, T. J., and D. B. Stephenson, 2012b: On the robustness of emergent constraints used in multi-model climate change projections of Arctic warming. *Journal of Climate*, **submitted**.

Braconnot, P., F. Hourdin, S. Bony, J. Dufresne, J. Grandpeix, and O. Marti, 2007a: Impact of different convective cloud schemes on the simulation of the tropical seasonal cycle in a coupled oceanâ□“atmosphere model. *Climate Dynamics*, **29**, 501-520.

Braconnot, P., et al., 2012: Evaluation of climate models using palaeoclimatic data. *Nature Climate Change*, **2**, 417-424.

Braconnot, P., et al., 2007b: Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum - Part 2: feedbacks with emphasis on the location of the ITCZ and mid- and high latitudes heat budget. *Climate of the Past*, **3**, 279-296.

Braconnot, P., et al., 2007c: Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum - Part 1: experiments and large-scale features. *Climate of the Past*, **3**, 261-277.

Brands, S., J. Taboada, A. Cofino, T. Sauter, and C. Schneider, 2011: Statistical downscaling of daily temperatures in the NW Iberian Peninsula from global climate models: validation and future scenarios. *Climate Research*, **48**, 163-176.

Bresson, R., and R. Laprise, 2011: Scale-decomposed atmospheric water budget over North America as simulated by the Canadian Regional Climate Model for current and future climates. *Climate Dynamics*, **36**, 365-384.

Breugem, W. P., W. Hazeleger, and R. J. Haarsma, 2006: Multimodel study of tropical Atlantic variability and change. *Geophysical Research Letters*, **33**.

Brewer, S., J. Guiot, and F. Torre, 2007: Mid-Holocene climate change in Europe: a data-model comparison. *Climate of the Past*, **3**, 499-512.

Briegleb, B. P., and B. Light, 2007: A Delta-Eddington multiple scattering parameterization for solar radiation in the sea ice component of the Community Climate System Model. National Center for Atmospheric Research.

Briegleb, B. P., C. M. Blitz, E. C. Hunke, W. H. Lipscomb, M. M. Holland, J. L. Schramm, and R. E. Moritz, 2004: Scientific description of the sea ice component in the Community Climate System Model, Version 3.

Brierley, C. M., M. Collins, and A. J. Thorpe, 2010: The impact of perturbations to ocean-model parameters on climate and climate change in a coupled model. *Climate Dynamics*, **34**, 325-343.

Brogniez, H., and R. T. Pierrehumbert, 2007: Intercomparison of tropical tropospheric humidity in GCMs with AMSU-B water vapor data. *Geophysical Research Letters*, **34**.

Brogniez, H., R. Roca, and L. Picon, 2005: Evaluation of the distribution of subtropical free tropospheric humidity in AMIP-2 simulations using METEOSAT water vapor channel data. *Geophys. Res. Lett.*, **32**, L19708.

Bromwich, D., J. Nicolas, and A. Monaghan, 2011: An Assessment of Precipitation Changes over Antarctica and the Southern Ocean since 1989 in Contemporary Global Reanalyses. *Journal of Climate*, **24**, 4189-4209.

Brovkin, V., J. Bendtsen, M. Claussen, A. Ganopolski, C. Kubatzki, V. Petoukhov, and A. Andreev, 2002: Carbon cycle, vegetation, and climate dynamics in the Holocene: Experiments with the CLIMBER-2 model. *Global Biogeochemical Cycles*, **16**.

Brown, J., A. Fedorov, and E. Guilyardi, 2010a: How well do coupled models replicate ocean energetics relevant to ENSO? *Climate Dynamics*. 1-12.

Brown, J., O. J. Ferrians, J. A. Heginbottom, and E. S. E.S. Melnikov, 1997: *International Permafrost Association Circum-Arctic Map of Permafrost and Ground Ice Conditions*.

——, 1998: Digital circum-arctic map of permafrost and ground ice conditions. In Circumpolar Active-Layer Permafrost System (CAPS) CD-ROM. 1.0 ed., U. o. C. a. B. National Snow and Ice Data Center, Ed.

Brown, J. R., C. Jakob, and J. M. Haynes, 2010b: An Evaluation of Rainfall Frequency and Intensity over the Australian Region in a Global Climate Model. *Journal of Climate*, **23**, 6504-6525.

Brutel-Vuilmet, C., M. Menegoz, and G. Krinner, 2012: An analysis of present and future seasonal Northern Hemisphere land snow cover simulated by CMIP5 coupled climate models. *The Cryosphere*.

Bryan, F. O., M. W. Hecht, and R. D. Smith, 2007: Resolution convergence and sensitivity studies with North Atlantic circulation models. Part I: The western boundary current system. *Ocean Modelling*, **16**, 141-159.

Bryan, F. O., R. Tomas, J. M. Dennis, D. B. Chelton, N. G. Loeb, and J. L. McClean, 2010: Frontal Scale Air-Sea Interaction in High-Resolution Coupled Climate Models. *Journal of Climate*, **23**, 6277-6291.

Bryan, K., and L. J. Lewis, 1979: WATER MASS MODEL OF THE WORLD OCEAN. *Journal of Geophysical Research-Oceans and Atmospheres*, **84**, 2503-2517.

Bryden, H. L., H. R. Longworth, and S. A. Cunningham, 2005: Slowing of the Atlantic meridional overturning circulation at 25N. *Nature*, **438**, 655-657.

Buehler, T., C. C. Raible, and T. F. Stocker, 2011: The relationship of winter season North Atlantic blocking frequencies to extreme cold or dry spells in the ERA-40. *Tellus Series a-Dynamic Meteorology and Oceanography*, **63**, 212-222.

Buser, C. M., H. R. Kunsch, D. Luthi, M. Wild, and C. Schar, 2009: Bayesian multi-model projection of climate: bias assumptions and interannual variability. *Climate Dynamics*, **33**, 849-868.

Butchart, N., A. A. Scaife, J. Austin, S. H. E. Hare, and J. R. Knight, 2003: Quasi-biennial oscillation in ozone in a coupled chemistry-climate model. *Journal of Geophysical Research*, **108**, ACL14-11-ACL14-ACL14-10.

Cadule, P., et al., 2010: Benchmarking coupled climate-carbon models against long-term atmospheric CO2 measurements. *Global Biogeochemical Cycles*, **24**.

Cai, W., A. Sullivan, and T. Cowan, 2011: Interactions of ENSO, the IOD, and the SAM in CMIP3 Models. *Journal of Climate*, **24**, 1688-1704.

Cai, W. J., A. Sullivan, and T. Cowan, 2009: Rainfall Teleconnections with Indo-Pacific Variability in the WCRP CMIP3 Models. *Journal of Climate*, **22**, 5046-5071.

Calov, R., A. Ganopolski, V. Petoukhov, M. Claussen, and R. Greve, 2002: Large-scale instabilities of the Laurentide ice sheet simulated in a fully coupled climate-system model. *Geophysical Research Letters*, **29**.

Cameron-Smith, P., J. F. Lamarque, P. Connell, C. Chuang, and F. Vitt, 2006: Toward an Earth system model: atmospheric chemistry, coupling, and petascale computing. *Scidac 2006: Scientific Discovery through Advanced Computing*, W. M. Tang, Ed.

Capotondi, A., A. Wittenberg, and S. Masina, 2006: Spatial and temporal structure of Tropical Pacific interannual variability in 20th century coupled simulations. *Ocean Modelling*, **15**, 274–298.

Capotondi, A., M. A. Alexander, N. A. Bond, E. N. Curchitser, and J. D. Scott, 2012: Enhanced upper ocean stratification with climate change in the CMIP3 models. *Journal of Geophysical Research-Oceans*, **117**.

Carslaw, K. S., O. Boucher, D. V. Spracklen, G. W. Mann, J. G. L. Rae, S. Woodward, and M. Kulmala, 2010: A review of natural aerosol interactions and feedbacks within the Earth system. *Atmospheric Chemistry and Physics*, **10**, 1701-1737.

Cassou, C., 2008: Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation. *Nature*, **455**, 523-527.

Cattiaux, J., H. Douville, and Y. Peings, 2012: European temperatures in CMIP5: origins of present-day biases and future uncertainties. *Climate Dynamics*. Submitted.

Catto, J., N. Nicholls, and C. Jakob, 2012a: North Australian sea surface temperatures and the El Niño Southern Oscillation in observations and models. in press ed., J. Climate.

——, 2012b: North Australian sea surface temperatures and the El Niño Southern Oscillation in the CMIP5 models. in press ed., J. Climate.

Catto, J. L., L. C. Shaffrey, and K. I. Hodges, 2010: Can Climate Models Capture the Structure of Extratropical Cyclones? *Journal of Climate*, **23**, 1621-1635.

——, 2011: Northern Hemisphere Extratropical Cyclones in a Warming Climate in the HiGEM High-Resolution Climate Model. *Journal of Climate*, **24**, 5336-5352.

Cavicchia, L., and H. Storch, 2011: The simulation of medicanes in a high-resolution regional climate model. *Climate Dynamics*. doi:10.1007/s00382-011-1220-0, 1-18.

Cha, D., D. Lee, and S. Hong, 2008: Impact of boundary layer processes on seasonal simulation of the East Asian summer monsoon using a Regional Climate Model. *Meteorology and Atmospheric Physics*, **100**, 53-72.

Chahine, M. T., et al., 2006: AIRS: Improving Weather Forecasting and Providing New Data on Greenhouse Gases. *Bull. Amer. Meteorol. Soc.*, **87**, 911-926.

Champion, A. J., K. I. Hodges, L. O. Bengtsson, N. S. Keenlyside, and M. Esch, 2011: Impact of increasing resolution and a warmer climate on extreme weather from Northern Hemisphere extratropical cyclones. *Tellus Series a-Dynamic Meteorology and Oceanography*, **63**, 893-906.

Chan, S. C., E. J. Kendon, H. J. Fowler, S. Blenkinsop, C. A. T. Ferro, and D. B. Stephenson, 2012: Does increasing resolution improve the simulation of United Kingdom daily precipitation in a regional climate model? *Climate Dynamics*. Submitted.

Chang, C. Y., S. Nigam, and J. A. Carton, 2008: Origin of the springtime westerly bias in equatorial Atlantic surface winds in the Community Atmosphere Model version 3 (CAM3) simulation. *Journal of Climate*, **21**, 4766-4778.

Chang, C. Y., J. A. Carton, S. A. Grodsky, and S. Nigam, 2007: Seasonal climate of the tropical Atlantic sector in the NCAR community climate system model 3: Error structure and probable causes of errors. *Journal of Climate*, **20**, 1053-1070.

Chang, E. K. M., Y. Guo, and X. Xia, 2012: CMIP5 Multi-Model Ensemble Projection of Storm Track Change Under Global Warming. *Journal of Geophysical Research*. doi:papers2://publication/uuid/22DFC6A5-CB39-4108-B773-722ACBE305F7, submitted.

Charlton-Perez, A. J., and e. al, *2012*: Mean climate and variability of the stratosphere in CMIP5 models. *J. Geophys. Res.*, **submitted**.

Charlton-Perez, A. J., and Coauthors, 2012: Mean climate and variability of the stratosphere in CMIP5 models. *J. Geophys. Res.*, submitted.

Chen, C. T., and T. Knutson, 2008: On the verification and comparison of extreme rainfall indices from climate models. *Journal of Climate*, **21**, 1605-1621.

Chen, H. M., T. J. Zhou, R. B. Neale, X. Q. Wu, and G. J. Zhang, 2010: Performance of the New NCAR CAM3.5 in East Asian Summer Monsoon Simulations: Sensitivity to Modifications of the Convection Scheme. *Journal of Climate*, **23**, 3657-3675.

Chen, J., F. P. Brissette, D. Chaumont, and M. Braun, 2012: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resources Research*, **Submitted**.

Chen, Y. H., and A. D. Del Genio, 2009: Evaluation of tropical cloud regimes in observations and a general circulation model. *Climate Dynamics*, **32**, 355-369.

Chiang, J. C. H., and A. H. Sobel, 2002: Tropical tropospheric temperature variations caused by ENSO and their influence on the remote tropical climate. *Journal of Climate*, **15**.

Chiang, J. C. H., and D. J. Vimont, 2004: Analogous Pacific and Atlantic meridional modes of tropical atmosphere-ocean variability. *Journal of Climate*, **17**, 4143-4158.

Chou, C., and J. Y. Tu, 2008: Hemispherical asymmetry of tropical precipitation in ECHAM5/MPI-OM during El Nino and under global warming. *Journal of Climate*, **21**, 1309-1332.

Chou, C., J. D. Neelin, J. Y. Tu, and C. T. Chen, 2006: Regional tropical precipitation change mechanisms in ECHAM4/OPYC3 under global warming. *Journal of Climate*, **19**, 4207-4223.

Chou, S., et al., 2012: Downscaling of South America present climate driven by 4-member HadCM3 runs. *Climate Dynamics*, **38**, 635-653.

Christensen, J., F. Boberg, O. Christensen, and P. Lucas-Picher, 2008: On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, **35**, -.

Christensen, J., E. Kjellstrom, F. Giorgi, G. Lenderink, and M. Rummukainen, 2010: Weight assignment in regional climate models. *Climate Research*. doi:DOI 10.3354/cr00916, 179-194.

Christensen, J., B. Machenhauer, R. Jones, C. Schar, P. Ruti, M. Castro, and G. Visconti, 1997: Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions. *Climate Dynamics*, **13**, 489-506.

Christensen, J. H., and F. Boberg, 2012: Temperature dependent model deficiencies affect CMIP5 multi-model mean climate projections. *Geophysical Research Letters*, **Submitted**.

Christian, J. R., et al., 2010: The global carbon cycle in the Canadian Earth system model (CanESM1): Preindustrial control simulation. *Journal of Geophysical Research-Biogeosciences*, **115**.

Christidis, N., P. A. Stott, and S. J. Brown, 2011: The Role of Human Activity in the Recent Warming of Extremely Warm Daytime Temperatures. *Journal of Climate*, **24**, 1922-1930.

Christy, J. R., et al., 2010: What Do Observational Datasets Say about Modeled Tropospheric Temperature Trends since 1979? *Remote Sensing*, **2**, 2148-2169.

Chung, E. S., B. J. Soden, and B. J. Sohn, 2010a: Revisiting the determination of climate sensitivity from relationships between surface temperature and radiative fluxes. *Geophysical Research Letters*, **37**.

Chung, E. S., D. Yeomans, and B. J. Soden, 2010b: An assessment of climate feedback processes using satellite observations of clear-sky OLR. *Geophysical Research Letters*, **37**.

Cionni, I., et al., 2011: Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing. *Atmos. Chem. Phys.*, **11**, 11267-11292.

Clark, D. B., et al., 2011: The Joint UK Land Environment Simulator (JULES), model description - Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, **4**, 701-722.

Claussen, M., et al., 2002: Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models. *Climate Dynamics*, **18**, 579-586.

Clement, A. C., R. Burgman, and J. R. Norris, 2009: Observational and Model Evidence for Positive Low-Level Cloud Feedback. *Science*, **325**, 460-464.

Clement, A. C., R. Seager, M. A. Cane, and S. E. Zebiak, 1996: An ocean dynamical thermostat. *Journal of Climate*, **9**, 2190-2196.

CLIMAP, 1981: Seasonal reconstructions of the Earth's surface at the last glacial maximum. geological Society of America, Boulder, Colorado. Map Series Technical Report MC-36.

Coelho, C. A. S., and L. Goddard, 2009: El Nino-Induced Tropical Droughts in Climate Change Projections. *Journal of Climate*, **22**, 6456-6476.

Collatz, G. J., M. Ribas-Carbo, and J. A. Berry, 1992: COUPLED PHOTOSYNTHESIS-STOMATAL CONDUCTANCE MODEL FOR LEAVES OF C4 PLANTS. *Australian Journal of Plant Physiology*, **19**.

Collatz, G. J., J. T. Ball, C. Grivet, and J. A. Berry, 1991: PHYSIOLOGICAL AND ENVIRONMENTAL-REGULATION OF STOMATAL CONDUCTANCE, PHOTOSYNTHESIS AND TRANSPIRATION - A MODEL THAT INCLUDES A LAMINAR BOUNDARY-LAYER. *Agricultural and Forest Meteorology*, **54**.

Collins, M., S. Tett, and C. Cooper, 2001: The Internal Climate Variability of HadCM3, a Version of the Hadley Centre Coupled Model without Flux Adjustments. *Climate Dynamics*, **17**, 61-81.

Collins, M., C. M. Brierley, M. MacVean, B. B. B. Booth, and G. R. Harris, 2007: The sensitivity of the rate of transient climate change to ocean physics perturbations. *Journal of Climate*, **20**, 2315-2320.

Collins, M., B. B. B. Booth, G. R. Harris, J. M. Murphy, D. M. H. Sexton, and M. J. Webb, 2006a: Towards quantifying uncertainty in transient climate change. *Climate Dynamics*, **27**, 127-147.

Collins, M., R. E. Chandler, P. M. Cox, J. M. Huthnance, J. Rougier, and D. B. Stephenson, 2012a: Quantifying future climate change. *Nature Climate Change*, **2**, 403-409.

Collins, M., R. Chandler, P. Cox, J. Huthnance, J. Rougier, and D. Stephenson, 2012b: Quantifying future climate change. *Nature Climate Change*, **2**, 403-409.

Collins, M., B. Booth, B. Bhaskaran, G. Harris, J. Murphy, D. Sexton, and M. Webb, 2010: Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. *Climate Dynamics*. doi:10.1007/s00382-010-0808-0, 1-30.

Collins, M., B. Booth, B. Bhaskaran, G. Harris, J. Murphy, D. Sexton, and M. Webb, 2011a: Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. *Climate Dynamics*. doi:DOI 10.1007/s00382-010-0808-0, 1737-1766.

Collins, W. D., et al., 2006b: The formulation and atmospheric simulation of the Community Atmosphere Model version 3 (CAM3). *Journal of Climate*, **19**, 2144-2161.

Collins, W. D., et al., 2006c: The Community Climate System Model version 3 (CCSM3). *Journal of Climate*, **19**, 2122-2143.

Collins, W. J., et al., 2008: Evaluation of the HadGEM2 model.

Collins, W. J., et al., 2011b: Development and evaluation of an Earth-System model-HadGEM2. *Geoscientific Model Development*, **4**, 1051-1075.

Colman, R. A., and B. J. McAvaney, 2011: On tropospheric adjustment to forcing and climate feedbacks. *Climate Dynamics*, **36**, 1649-1658.

Colman, R. A., A. F. Moise, and L. I. Hanson, 2011: Tropical Australian climate and the Australian monsoon as simulated by 23 CMIP3 models. *Journal of Geophysical Research-Atmospheres*, **116**.

Comiso, J. C., 2008: Bootstrap sea ice concentrations from nimbus-7 smmr and dmsp ssm/i.

Compo, G. P., and P. D. Sardeshmukh, 2009: Oceanic influences on recent continental warming. *Climate Dynamics*, **32**, 333-342.

Connolley, W., and T. Bracegirdle, 2007: An antarctic assessment of IPCC AR4 coupled models. *Geophysical Research Letters*. doi:ARTN L22505, 10.1029/2007GL031648, -.

Coon, M., R. Kwok, G. Levy, M. Pruis, H. Schreyer, and D. Sulsky, 2007: Arctic Ice Dynamics Joint Experiment (AIDJEX) assumptions revisited and found inadequate. *Journal of Geophysical Research*, **112, C11S90**.

Coppola, E., F. Giorgi, S. Rauscher, and C. Piani, 2010: Model weighting based on mesoscale structures in precipitation and temperature in an ensemble of regional climate models. *Climate Research*. doi:DOI 10.3354/cr00940, 121-134.

Cordero, E. C., and P. M. D. Forster, 2006: Stratospheric variability and trends in models used for the IPCC AR4. *Atmospheric Chemistry and Physics*, **6**, 5369-5380.

Costa, A. A., D. C. Sales, M. M. Coutinho, F. das Chagas Vasconcelos Jr, L. M. de Araújo Jr, and E. M. da Silva, 2012: Climate change simulations over the Tropical Americas using a regional model. *Climate Dynamics*, **Submitted**.

Covey, C., A. G. Dai, D. Marsh, and R. S. Lindzen, 2011: The Surface-Pressure Signature of Atmospheric Tides in Modern Climate Models. *Journal of the Atmospheric Sciences*, **68**, 495-514.

Cox, P., 2001a: Description of the “TRIFFID” Dynamic Global Vegetation Model. Hadley Centre, United Kingdom Meteorological Office.

Cox, P., 2001b: Description of the "TRIFFID" Dynamic Global Vegetation Model

Cox, P. M., R. A. Betts, C. D. Jones, S. A. Spall, and I. J. Totterdell, 2000: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**, 184-187.

Cox, P. M., R. A. Betts, C. B. Bunton, R. L. H. Essery, P. R. Rowntree, and J. Smith, 1999a: The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. *Climate Dynamics*, **15**, 183-203.

——, 1999b: The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. *Climate Dynamics*, **15**, 183-203.

Cox, P. M., D. Pearson, B. B. B. Booth, P. Friedlingstein, C. Huntingford, C. D. Jones, and C. M. Luke, 2012: Carbon dioxide variability constrains the sensitivity of tropical carbon to climate change. *Nature*, **submitted**.

Cramer, W., et al., 2001: Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models. *Global Change Biology*, **7**, 357-373.

Crétat, J., B. Pohl, Y. Richard, and P. Drobinski, 2012: Uncertainties in simulating regional climate of Southern Africa: sensitivity to physical parameterizations using WRF. *Climate Dynamics*, **38**, 613-634.

Croft, B., U. Lohmann, and K. von Salzen, 2005: Black carbon ageing in the Canadian Centre for Climate modelling and analysis atmospheric general circulation model. *Atmospheric Chemistry and Physics*, **5**.

Crook, J., N. Gillett, and S. Keeley, 2008: Sensitivity of Southern Hemisphere climate to zonal asymmetry in ozone. *Geophysical Research Letters*, **35**.

Crucifix, M., 2006: Does the Last Glacial Maximum constrain climate sensitivity? *Geophysical Research Letters*, **33**, -.

Cunningham, S., et al., 2010: The present and future system for measuring the Atlantic meridional overturning circulation and heat transport. *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2)*, Venice, Italy, 21-25 September 2009, ESA Publication.

Cunningham, S. A., S. G. Alderson, B. A. King, and M. A. Brandon, 2003: Transport and variability of the Antarctic Circumpolar Current in Drake Passage. *Journal of Geophysical Research-Oceans*, **108**.

Cunningham, S. A., et al., 2007: Temporal Variability of the Atlantic Meridional Overturning Circulation at 26.5°N. *Science*, **317**, 935-938.

Curry, W. B., and D. W. Oppo, 2005: Glacial water mass geometry and the distribution of delta C-13 of Sigma CO2 in the western Atlantic Ocean. *Paleoceanography*, **20**, -.

Cuxart, J., et al., 2006: Single-Column Model Intercomparison for a Stably Stratified Atmospheric Boundary Layer. *Boundary-Layer Meteorology*, **118**, 273-303-303.

Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *Journal of Climate*, **19**, 4605-4630.

Dai, A., and C. Deser, 1999: Diurnal and semidiurnal variations in global surface wind and divergence fields. *Journal of Geophysical Research-Atmospheres*, **104**.

Dai, A., and K. E. Trenberth, 2004: The diurnal cycle and its depiction in the Community Climate System Model. *Journal of Climate*, **17**, 930-951.

Dai, A., K. E. Trenberth, and T. T. Qian, 2004a: A global dataset of Palmer Drought Severity Index for 1870-2002: Relationship with soil moisture and effects of surface warming. *Journal of Hydrometeorology*, **5**.

Dai, Y. J., R. E. Dickinson, and Y. P. Wang, 2004b: A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *Journal of Climate*, **17**.

Dai, Y. J., et al., 2003: The Common Land Model. *Bull. Amer. Meteorol. Soc.*, **84**.

Dallmeyer, A., M. Claussen, and J. Otto, 2010: Contribution of oceanic and vegetation feedbacks to Holocene climate change in monsoonal Asia. *Climate of the Past*, **6**, 195-218.

Danabasoglu, G., and P. R. Gent, 2009: Equilibrium Climate Sensitivity: Is It Accurate to Use a Slab Ocean Model? *Journal of Climate*, **22**, 2494-2499.

Danabasoglu, G., R. Ferrari, and J. C. McWilliams, 2008: Sensitivity of an ocean general circulation model to a parameterization of near-surface eddy fluxes. *Journal of Climate*, **21**, 1192-1208.

Danabasoglu, G., W. G. Large, and B. P. Briegleb, 2010: Climate impacts of parameterized Nordic Sea overflows. *Journal of Geophysical Research-Oceans*, **115**.

Danabasoglu, G., W. G. Large, J. J. Tribbia, P. R. Gent, B. P. Briegleb, and J. C. McWilliams, 2006: Diurnal coupling in the tropical oceans of CCSM3. *Journal of Climate*, **19**, 2347-2365.

Danabasoglu, G., et al., 2012: The CCSM4 Ocean Component. *Journal of Climate*, **25**.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood, 2005: A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **131**.

Davis, B. A. S., and S. Brewer, 2009: Orbital forcing and role of the latitudinal insolation/temperature gradient. *Climate Dynamics*, **32**, 143-165.

Dawson, A., T. N. Palmer, and S. Corti, 2012: Simulating regime structures in weather and climate prediction models. *Geophysical Research Letters*. Submitted.

de Elia, R., and H. Cote, 2010: Climate and climate change sensitivity to model configuration in the Canadian RCM over North America. *Meteorologische Zeitschrift*, **19**, 325-339.

de Elia, R., S. Biner, and A. Frigon, 2012: Climate variability and expected regional climate change over North America. *Climate Dynamics*, **Submitted**.

de Jong, M. F., S. S. Drijfhout, W. Hazeleger, H. M. van Aken, and C. A. Severijns, 2009: Simulations of Hydrographic Properties in the Northwestern North Atlantic Ocean in Coupled Climate Models. *Journal of Climate*, **22**, 1767-1786.

De Szoeke, S. P., and S. P. Xie, 2008: The tropical eastern Pacific seasonal cycle: Assessment of errors and mechanisms in IPCC AR4 coupled ocean - Atmosphere general circulation models. *Journal of Climate*, **21**, 2573-2590.

Dee, D. P., et al., 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137**, 553-597.

DelSole, T., and J. Shukla, 2009: Artificial Skill due to Predictor Screening. *Journal of Climate*, **22**, 331-345.

Delworth, T. L., et al., 2011: Simulated Climate and Climate Change in the GFDL CM2.5 High-Resolution Coupled Climate Model. *Journal of Climate*, **25**, 2755-2781.

Delworth, T. L., et al., 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *Journal of Climate*, **19**.

Deque, M., 2007: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global and Planetary Change*, **57**, 16-26.

Deque, M., 2010: Regional climate simulation with a mosaic of RCMs. *Meteorologische Zeitschrift*, **19**, 259-266.

Deque, M., C. Dreveton, A. Braun, and D. Cariolle, 1994: THE ARPEGE/IFS ATMOSPHERE MODEL - A CONTRIBUTION TO THE FRENCH COMMUNITY CLIMATE MODELING. *Climate Dynamics*, **10**.

Deque, M., S. Somot, E. Sanchez-Gomez, C. Goodess, D. Jacob, G. Lenderink, and O. Christensen, 2012: The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. *Climate Dynamics*, **38**, 951-964.

Déqué, M., et al., 2007: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. 53-70.

Deser, C., A. S. Phillips, and J. W. Hurrell, 2004: Pacific interdecadal climate variability: Linkages between the tropics and the North Pacific during boreal winter since 1900. *Journal of Climate*, **17**, 3109-3124.

Deser, C., A. S. Phillips, V. Bourdette, and H. Teng, 2011a: Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*. doi:10.1007/s00382-010-0977-x.

Deser, C., et al., 2011b: ENSO and Pacific Decadal Variability in Community Climate System Model Version 4. *J. Climate (in press)*.

Dessler, A. E., 2010: A Determination of the Cloud Feedback from Climate Variations over the Past Decade. *Science*, **330**, 1523-1527.

Dessler, A. E., and S. Wong, 2009: Estimates of the Water Vapor Climate Feedback during El Nino-Southern Oscillation. *Journal of Climate*, **22**, 6404-6412.

Deushi, M., and K. Shibata, 2011: Development of a Meteorological Research Institute Chemistry-Climate Model version 2 for the Study of Tropospheric and Stratospheric Chemistry.

Di Luca, A., R. Elía, and R. Laprise, 2012: Potential for small scale added value of RCM's downscaled climate change signal. *Climate Dynamics*. doi:10.1007/s00382-012-1415-z, 1-18.

Diaconescu, E. P., R. Laprise, and A. Zadra, 2012: Can added value be expected in RCM-simulated large scales? *Climate Dynamics*, **Submitted**.

Diffenbaugh, N., M. Ashfaq, and M. Scherer, 2011: Transient regional climate change: Analysis of the summer climate response in a high-resolution, century-scale ensemble experiment over the continental United States. *Journal of Geophysical Research-Atmospheres*, **116**.

DiNezio, P. N., A. C. Clement, G. A. Vecchi, B. J. Soden, and B. P. Kirtman, 2009: Climate Response of the Equatorial Pacific to Global Warming. *Journal of Climate*, **22**, 4873-4892.

Dix, M., et al., 2012: The ACCESS Coupled Model: Documentation of core CMIP5 simulations and initial results. *Aust. Met. Oceanog. J.*, **submitted**.

Dokken, T. M., and E. Jansen, 1999: Rapid changes in the mechanism of ocean convection during the last glacial period. *Nature*, **401**, 458-461.

Domingues, C., J. Church, N. White, P. Gleckler, S. Wijffels, P. Barker, and J. Dunn, 2008: Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature*. doi:DOI 10.1038/nature07080, 1090-U1096.

Donat, M., G. Leckebusch, S. Wild, and U. Ulbrich, 2010: Benefits and limitations of regional multi-model ensembles for storm loss estimations. *Climate Research*. doi:DOI 10.3354/cr00891, 211-225.

Doney, S. C., et al., 2009: Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data. *Journal of Marine Systems*, **76**, 95-112.

Donner, L. J., et al., 2011: The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. *Journal of Climate*, **24**, 3484-3519.

Donohoe, A., and D. S. Battisti, 2011: Atmospheric and Surface Contributions to Planetary Albedo. *Journal of Climate*, **24**, 4402-4418.

Donohoe, A., and D. S. Battisti, 2012: What Determines Meridional Heat Transport in Climate Models? *Journal of Climate*, **25**, 3832-3850.

Dorn, W., K. Dethloff, and A. Rinke, 2009: Improved simulation of feedbacks between atmosphere and sea ice over the Arctic Ocean in a coupled regional climate model. *Ocean Modelling*, **29**, 103-114.

Doscher, R., K. Wyser, H. E. M. Meier, M. W. Qian, and R. Redler, 2010: Quantifying Arctic contributions to climate predictability in a regional coupled ocean-ice-atmosphere model. *Climate Dynamics*, **34**, 1157-1176.

Dosio, A., and P. Paruolo, 2011: Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate. *Journal of Geophysical Research-Atmospheres*, **116**.

Douglass, D., J. Christy, B. Pearson, and S. Singer, 2008: A comparison of tropical temperature trends with model predictions. *International Journal of Climatology*. doi:DOI 10.1002/joc.1651, 1693-1701.

Driouech, F., M. Deque, and E. Sanchez-Gomez, 2010: Weather regimes-Moroccan precipitation link in a regional climate change simulation. *Global and Planetary Change*, **72**, 1-10.

Druyan, L. M., et al., 2010: The WAMME regional model intercomparison study. *Climate Dynamics*, **35**, 175-192.

Du, Y., S.-P. Xie, G. Huang, and K. Hu, 2009: Role of Air-Sea Interaction in the Long Persistence of El Nino-Induced North Indian Ocean Warming. *Journal of Climate*, **22**, 2023-2038.

Ducet, N., P. Y. Le Traon, and G. Reverdin, 2000: Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and-2. *Journal of Geophysical Research-Oceans*, **105**, 19477-19498.

Dufresne, J.-L., and a. co-authors, 2011: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. Submitted ed., Clim. Dyn.

Dufresne, J.-L., et al., 2012: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim. Dynamics.*, **submitted**.

Dufresne, J. L., and S. Bony, 2008: An assessment of the primary sources of spread of global warming estimates from coupled atmosphere-ocean models. *Journal of Climate*, **21**, 5135-5144.

Dunkerton, T. J., 1991: NONLINEAR PROPAGATION OF ZONAL WINDS IN AN ATMOSPHERE WITH NEWTONIAN COOLING AND EQUATORIAL WAVEDRIVING. *Journal of the Atmospheric Sciences*, **48**, 236-263.

Dunkerton, T. J., 1997: The role of gravity waves in the quasi-biennial oscillation. *Journal of Geophysical Research*, **102**, 26053-26076.

Duplessy, J. C., N. J. Shackleton, R. Fairbanks, L. Labeyrie, D. Oppo, and N. Kallel, 1988: Deep water source variation during the last climatic cycle and their impact on th global deep water circulation. *Paleoceanography*, **3**, 343-360.

Durack, P. J., S. E. Wijffels, and R. J. Matear, 2012: Ocean Salinities Reveal Strong Global Water Cycle Intensification During 1950 to 2000. *Science*, **336**.

Eby, M., K. Zickfeld, A. Montenegro, D. Archer, K. J. Meissner, and A. J. Weaver, 2009: Lifetime of Anthropogenic Climate Change: Millennial Time Scales of Potential CO(2) and Surface Temperature Perturbations. *Journal of Climate*, **22**, 2501-2511.

Eby, M., et al., 2012: Historical and Idealized Climate Model Experiments: An EMIC Intercomparison. *Climate of the Past*. submitted.

——, submitted: **Historical and Idealized Climate Model Experiments: An EMIC Intercomparison** *Climate of the Past*.

Edwards, N., and R. Marsh, 2005: Uncertainties due to transport-parameter sensitivity in an efficient 3-D ocean-climate model. *Climate Dynamics*, **24**, 415-433.

Edwards, N. R., D. Cameron, and J. Rougier, 2011: Precalibrating an intermediate complexity climate model. *Climate Dynamics*, **37**, 1469-1482.

Ek, M. B., et al., 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research-Atmospheres*, **108**.

Eliseev, A. V., and I. I. Mokhov, 2011: Uncertainty of climate response to natural and anthropogenic forcings due to different land use scenarios. *Advances in Atmospheric Sciences*, **28**, 1215-1232.

Emanuel, K., R. Sundararajan, and J. Williams, 2008: Hurricanes and global warming - Results from downscaling IPCC AR4 simulations. *Bull. Amer. Meteorol. Soc.*, **89**, 347-+.

Ern, M., and P. Preusse, 2009: Wave fluxes of equatorial Kelvin waves and QBO zonal wind forcing derived from SABER and ECMWF temperature space time spectra. *Atmospheric Chemistry and Physics*, **9**, 3957–3986.

Essery, R. L. H., M. J. Best, R. A. Betts, P. M. Cox, and C. M. Taylor, 2003: Explicit representation of subgrid heterogeneity in a GCM land surface scheme. *Journal of Hydrometeorology*, **4**.

Eum, H., P. Gachon, R. Laprise, and T. Ouarda, 2012: Evaluation of regional climate model simulations versus gridded observed and regional reanalysis products using a combined weighting scheme. *Climate Dynamics*, **38**, 1433-1457.

Evan, A. T., D. J. Vimont, A. K. Heidinger, J. P. Kossin, and R. Bennartz, 2009: The Role of Aerosols in the Evolution of Tropical North Atlantic Ocean Temperature Anomalies. *Science*, **324**, 778-781.

Eyring, V., et al., 2010: Transport impacts on atmosphere and climate: Shipping. *Atmos. Environ.*, **44**, 4735-4771.

Eyring, V., et al., 2005: A strategy for process-oriented validation of coupled chemistry-climate models. *Bull. Amer. Meteorol. Soc.* doi:DOI 10.1175/BAMS-86-8-1117, 1117-+.

Eyring, V., et al., 2012a: Long-term changes in tropospheric and stratospheric ozone and associated climate impacts in CMIP5 simulations. *J. Geophys. Res.,* , **submitted**.

Eyring, V., et al., 2012b: Long-term changes in tropospheric and stratospheric ozone and associated climate impacts in CMIP5 simulations. *J. Geophys. Res.*, **submitted**.

Eyring, V., et al., 2007: Multimodel projections of stratospheric ozone in the 21st century. *Journal of Geophysical Research-Atmospheres*. doi:ARTN D16303, 10.1029/2006JD008332, -.

Faloona, I., 2009: Sulfur processing in the marine atmospheric boundary layer: A review and critical assessment of modeling uncertainties. *Atmos. Environ.*, **43**, 2841-2854.

Fan, F. X., M. E. Mann, S. Lee, and J. L. Evans, 2010: Observed and Modeled Changes in the South Asian Summer Monsoon over the Historical Period. *Journal of Climate*, **23**, 5193-5205.

Fangli, Q., Y. Yeli, Y. Yongzeng, Q. Zheng, X. Changshui, and M. Jian, 2004: Wave-induced mixing in the upper ocean: distribution and application to a global ocean circulation model. *Geophysical Research Letters*, **31**.

Fanning, A. F., and A. J. Weaver, 1996: An atmospheric energy-moisture balance model: Climatology, interpentadal climate change, and coupling to an ocean general circulation model. *Journal of Geophysical Research-Atmospheres*, **101**, 15111-15128.

Farneti, R., and P. R. Gent, 2011: The effects of the eddy-induced advection coefficient in a coarse-resolution coupled climate model. *Ocean Modelling*, **39**, 135-145.

Farneti, R., T. L. Delworth, A. J. Rosati, S. M. Griffies, and F. R. Zeng, 2010: The Role of Mesoscale Eddies in the Rectification of the Southern Ocean Response to Climate Change. *Journal of Physical Oceanography*, **40**, 1539-1557.

Fasullo, J., and K. E. Trenberth, 2012: Outlook Less Cloudy: The Role of Subsiding Air Teleconnections in Climate Sensitivity. *Science*. submitted.

Fauchereau, N., S. Trzaska, Y. Richard, P. Roucou, and P. Camberlin, 2003: Sea-surface temperature co-variability in the southern Atlantic and Indian Oceans and its connections with the atmospheric circulation in the southern hemisphere. *International Journal of Climatology*, **23**, 663-677.

Felzer, B., D. Kicklighter, J. Melillo, C. Wang, Q. Zhuang, and R. Prinn, 2004: Effects of ozone on net primary production and carbon sequestration in the conterminous United States using a biogeochemistry model. *Tellus Series B-Chemical and Physical Meteorology*, **56**, 230-248.

Feng, J., and C. Fu, 2006: Inter-comparison of 10-year precipitation simulated by several RCMs for Asia. *Advances in Atmospheric Sciences*. doi:DOI 10.1007/s00376-006-0531-2, 531-542.

Feng, J., et al., 2011: Comparison of four ensemble methods combining regional climate simulations over Asia. *Meteorology and Atmospheric Physics*, **111**, 41-53.

Fernandez-Donado, L., et al., 2012: Temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*.

Ferrari, R., J. C. McWilliams, V. M. Canuto, and M. Dubovikov, 2008: Parameterization of eddy fluxes near oceanic boundaries. *Journal of Climate*, **21**, 2770-2789.

Ferrari, R., S. M. Griffies, A. J. G. Nurser, and G. K. Vallis, 2010: A boundary-value problem for the parameterized mesoscale eddy transport. *Ocean Modelling*, **32**, 143-156.

Feser, F., 2006: Enhanced Detectability of Added Value in Limited-Area Model Results Separated into Different Spatial Scales. *Monthly Weather Review*, **134**, 2180-2190.

Feser, F., and M. Barcikowska, 2012: The influence of spectral nudging on typhoon formation in regional climate models. *Environmental Research Letters*, **7**.

Feser, F., B. Rockel, H. von Storch, J. Winterfeldt, and M. Zahn, 2011: REGIONAL CLIMATE MODELS ADD VALUE TO GLOBAL MODEL DATA A Review and Selected Examples. *Bull. Amer. Meteorol. Soc.*, **92**, 1181-1192.

Fetterer, F., K. Knowles, W. Meier, and M. Savoie, 2002: Sea Ice Index. National Snow and Ice data Center.

Fichefet, T., and M. A. Morales Maqueda, 1997: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *J. Geophys. Res.*, **102**, 12,646.

Fichefet, T., and M. A. M. Maqueda, 1997: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *Journal of Geophysical Research-Oceans*, **102**, 12609-12646.

——, 1999: Modelling the influence of snow accumulation and snow-ice formation on the seasonal cycle of the Antarctic sea-ice cover. *Climate Dynamics*, **15**.

Field, P. R., A. Gettelman, R. B. Neale, R. Wood, P. J. Rasch, and H. Morrison, 2008: Midlatitude Cyclone Compositing to Constrain Climate Model Behavior Using Satellite Observations. *Journal of Climate*, **21**, 5887-5903.

Fioletov, V., G. Bodeker, A. Miller, R. McPeters, and R. Stolarski, 2002: Global and zonal total ozone variations estimated from ground-based and satellite measurements: 1964-2000. *Journal of Geophysical Research-Atmospheres*, **107**.

Fischer, E. M., D. M. Lawrence, and B. M. Sanderson, 2011: Quantifying uncertainties in projections of extremes-a perturbed land surface parameter experiment. *Climate Dynamics*, **37**, 1381-1398.

Fischer, E. M., S. I. Seneviratne, D. Luthi, and C. Schar, 2007a: Contribution of land-atmosphere coupling to recent European summer heat waves. *Geophysical Research Letters*, **34**.

Fischer, E. M., S. I. Seneviratne, D. Lüthi, and C. Schär, 2007b: Contribution of land-atmosphere coupling to recent European summer heat waves. *Geophys. Res. Lett.*, **34**, L06707.

Flanner, M. G., K. M. Shell, M. Barlage, D. K. Perovich, and M. A. Tschudi, 2011: Radiative forcing and albedo feedback from the Northern Hemisphere cryosphere between 1979 and 2008. *Nature Geoscience*, **4**, 151-155.

Flato, G., 2011: Earth system models: an overview. John Wiley and Sons, 783-800.

Fletcher, C. G., H. X. Zhao, P. J. Kushner, and R. Fernandes, 2012: Using models and satellite observations to evaluate the strength of snow albedo feedback. *Journal of Geophysical Research-Atmospheres*, **117**, 9.

Fogli, P. G., et al., 2009: INGV-CMCC Carbon (ICC): A Carbon Cycle Earth System Model. *CMCC Research Papers*.

Fogt, R. L., J. Perlwitz, A. J. Monaghan, D. H. Bromwich, J. M. Jones, and G. J. Marshall, 2009: Historical SAM Variability. Part II: Twentieth-Century Variability and Trends from Reconstructions, Observations, and the IPCC AR4 Models. *Journal of Climate*, **22**, 5346-5365.

Foltz, G. R., and M. J. McPhaden, 2008: Trends in Saharan dust and tropical Atlantic climate during 1980-2006. *Geophysical Research Letters*, **35**.

Fontaine, B., and S. Janicot, 1996: Sea Surface Temperature Fields Associated with West African Rainfall Anomaly Types. *Journal of Climate*, **9**, 2935-2940.

Forest, C. E., P. H. Stone, and A. P. Sokolov, 2006: Estimated PDFs of climate system properties including natural and anthropogenic forcings. *Geophysical Research Letters*, **33**.

——, 2008: Constraining climate model parameters from observed 20th century changes. *Tellus Series a-Dynamic Meteorology and Oceanography*, **60**, 911-920.

Forest, C. E., P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster, 2002: Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science*, **295**, 113-117.

Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka, 2012: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *TBD*.

Fowler, H., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, **27**, 1547-1578.

Fox-Kemper, B., R. Ferrari, and R. Hallberg, 2008: Parameterization of mixed layer eddies. Part I: Theory and diagnosis. *Journal of Physical Oceanography*, **38**, 1145-1165.

Fox-Kemper, B., et al., 2011: Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations. *Ocean Modelling*, **39**, 61-78.

Fox-Rabinovitz, M., J. Cote, B. Dugas, M. Deque, J. McGregor, and A. Belochitski, 2008: Stretched-grid Model Intercomparison Project: decadal regional climate simulations with enhanced variable and uniform-resolution GCMs. *Meteorology and Atmospheric Physics*. doi:DOI 10.1007/s00703-008-0301-z, 159-177.

Frame, D., B. Booth, J. Kettleborough, D. Stainforth, J. Gregory, M. Collins, and M. Allen, 2005: Constraining climate forecasts: The role of prior assumptions. *Geophysical Research Letters*, **32**.

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. *Journal of Climate*, **23**, 3626-3638.

Free, M., and J. Lanzante, 2009: Effect of Volcanic Eruptions on the Vertical Temperature Profile in Radiosonde Data and Climate Models. *Journal of Climate*, **22**, 2925-2939.

Friedlingstein, P., et al., 2001: Positive feedback between future climate change and the carbon cycle. *Geophysical Research Letters*, **28**, 1543-1546.

Friedlingstein, P., et al., 2006: Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison. *Journal of Climate*, **19**, 3337-3353.

Friedrichs, M. A. M., et al., 2009: Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean. *Journal of Marine Systems*, **76**, 113-133.

Friend, A. D., et al., 2007: FLUXNET and modelling the global carbon cycle. *Global Change Biology*, **13**, 610-633.

Frierson, D. M. W., J. Lu, and G. Chen, 2007: Width of the Hadley cell in simple and comprehensive general circulation models. *Geophysical Research Letters*, **34**.

Fu, Q., S. Manabe, and C. M. Johanson, 2011: On the warming in the tropical upper troposphere: Models versus observations. *Geophys. Res. Lett.*, **38**, L15704.

Furrer, R., S. R. Sain, D. Nychka, and G. A. Meehl, 2007a: Multivariate Bayesian analysis of atmosphere - Ocean general circulation models. *Environmental and Ecological Statistics*, **14**, 249-266.

Furrer, R., R. Knutti, S. Sain, D. Nychka, and G. Meehl, 2007b: Spatial patterns of probabilistic temperature change projections from a multivariate Bayesian analysis. *Geophysical Research Letters*. doi:ARTN L06711, 10.1029/2006GL027754, -.

Furtado, J., E. Di Lorenzo, N. Schneider, and N. A. Bond, 2011a: North Pacific Decadal Variability and Climate Change in the IPCC AR4 Models. *Journal of Climate*, **24**, 18.

Furtado, J. C., E. Di Lorenzo, N. Schneider, and N. A. Bond, 2011b: North Pacific Decadal Variability and Climate Change in the IPCC AR4 Models. *Journal of Climate*, **24**, 3049-3067.

Fyfe, J. C., N. P. Gillett, and D. W. J. Thompson, 2010: Comparing variability and trends in observed and modelled global-mean surface temperature. *Geophysical Research Letters*, **37**.

Fyke, J. G., A. J. Weaver, D. Pollard, M. Eby, L. Carter, and A. Mackintosh, 2011: A new coupled ice sheet/climate model: description and sensitivity to model physics under Eemian, Last Glacial Maximum, late Holocene and modern climate conditions. *Geoscientific Model Development*, **4**, 117-136.

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. *New Phytologist*, **187**, 647-665.

Ganachaud, A., and C. Wunsch, 2003: Large-scale ocean heat and freshwater transports during the World Ocean Circulation Experiment. *Journal of Climate*, **16**, 696-705.

Gangsto, R., F. Joos, and M. Gehlen, 2011: Sensitivity of pelagic calcification to ocean acidification. *Biogeosciences*, **8**, 433-458.

Gates, W. L., et al., 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteorol. Soc.*, **80**, 29-55.

Gbobaniyi, E. O., B. J. Abiodun, M. A. Tadross, B. C. Hewitson, and W. J. Gutowski, 2011: The coupling of cloud base height and surface fluxes: a transferability intercomparison. *Theoretical and Applied Climatology*, **106**, 189-210.

Gedney, N., and P. M. Cox, 2003: The Sensitivity of Global Climate Model Simulations to the Representation of Soil Moisture Heterogeneity. *Journal of Hydrometeorology*, **4**, 1265-1275.

Gehlen, M., R. Gangsto, B. Schneider, L. Bopp, O. Aumont, and C. Ethe, 2007: The fate of pelagic CaCO3 production in a high CO2 ocean: a model study. *Biogeosciences*, **4**, 505-519.

Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, **20**, 150-155.

Gent, P. R., and G. Danabasoglu, 2011: Response to Increasing Southern Hemisphere Winds in CCSM4. *Journal of Climate*, **24**, 4992-4998.

Gent, P. R., J. Willebrand, T. J. McDougall, and J. C. McWilliams, 1995: PARAMETERIZING EDDY-INDUCED TRACER TRANSPORTS IN OCEAN CIRCULATION MODELS. *Journal of Physical Oceanography*, **25**, 463-474.

Gent, P. R., et al., 2011: The Community Climate System Model Version 4. *Journal of Climate*, **24**, 4973-4991.

Gerber, E. P., L. M. Polvani, and D. Ancukiewicz, 2008: Annular mode time scales in the Intergovernmental Panel on Climate Change Fourth Assessment Report models. *Geophys. Res. Lett.*, **35**, L22707-.

Gettelman, A., and Q. Fu, 2008: Observed and simulated upper-tropospheric water vapor feedback. *Journal of Climate*, **21**, 3282-3289.

Gillett, N. P., 2005: Climate modelling - Northern Hemisphere circulation. *Nature*, **437**, 496-496.

Giorgetta, M. A., E. Manzini, and E. Roeckner, 2002: Forcing of the Quasi-Biennial Oscillation from a broad spectrum of atmospheric waves. *Geophysical Research Letters*, **29**.

Giorgetta, M. A., et al., 2012: Climate change from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project 5 *JOURNAL OF ADVANCES IN MODELING EARTH SYSTEMS,*, **submitted**.

Giorgi, F., and E. Coppola, 2010: Does the model regional bias affect the projected regional climate change? An analysis of global model projections. *Climatic Change*, **100**, 787-795.

Giorgi, F., C. Jones, and G. Asrar, 2009: Addressing climate information at the regional level: The CORDEX framework. *WMO Bulletin*, World Meteorological Organisation, 175-183.

Girard, L., J. Weiss, J. M. Molines, B. Barnier, and S. Bouillon, 2009: Evaluation of high resolution sea ice models on the basis of statistical and scaling properties of Arctic sea ice drift and deformation. *Journal of Geophysical Research*, **114, C08015**.

Gleckler, P., K. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models. *Journal of Geophysical Research-Atmospheres*. doi:ARTN D06104, 10.1029/2007JD008972, -.

Gleckler, P., K. AchutaRao, J. Gregory, B. Santer, K. Taylor, and T. Wigley, 2006: Krakatoa lives: The effect of volcanic eruptions on ocean heat content and thermal expansion. *Geophysical Research Letters*. doi:ARTN L17702, 10.1029/2006GL026771, -.

Gleckler, P. J., et al., 2012a: Human-induced global ocean warming on multidecadal timescales. *NATURE CLIMATE CHANGE*, **2**, 524-529.

——, 2012b: Human-induced global ocean warming on multidecadal timescales, **2**, 524-529.

Gnanadesikan, A., S. M. Griffies, and B. L. Samuels, 2007: Effects in a climate model of slope tapering in neutral physics schemes. *Ocean Modelling*, **16**, 1-16.

Goddard, L., and S. J. Mason, 2002: Sensitivity of seasonal climate forecasts to persisted SST anomalies. *Climate Dynamics*, **19**, 619-631.

Goosse, H., and T. Fichefet, 1999: Importance of ice-ocean interactions for the global ocean circulation: A model study. *Journal of Geophysical Research-Oceans*, **104**, 23337-23355.

Goosse, H., et al., 2010: Description of the Earth system model of intermediate complexity LOVECLIM version 1.2. *Geoscientific Model Development*, **3**, 603-633.

Gordon, C., et al., 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, **16**, 147-168.

Gordon, H., et al., 2010: The CSIRO Mk3.5 Climate Model. *CAWCR Technical Report* **21**, 1-74.

Gordon, H. B., et al., 2002: The CSIRO Mk3 Climate System Model, Technical Paper No. 60, CSIRO Atmospheric Research. CSIRO Atmospheric Research, Aspendale, Vic., Australia, .

Greene, A. M., L. Goddard, and U. Lall, 2006: Probabilistic multimodel regional temperature change projections. *Journal of Climate*, **19**, 4326-4343.

Greeves, C. Z., V. D. Pope, R. A. Stratton, and G. M. Martin, 2007: Representation of Northern Hemisphere winter storm tracks in climate models. *Climate Dynamics*, **28**, 683-702.

Gregory, D., G. J. Shutts, and J. R. Mitchell, 1998: A new gravity-wave-drag scheme incorporating anisotropic orography and low-level wave breaking: Impact upon the climate of the UK Meteorological Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, **124**, 463-493.

Gregory, J., and M. Webb, 2008: Tropospheric adjustment induces a cloud component in CO2 forcing. *Journal of Climate*, **21**, 58-71.

Gregory, J. M., et al., 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, **31**.

Gregory, J. M., et al., 2005: A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO2 concentration. *Geophysical Research Letters*, **32**.

Griffies, Stephen M., & Greatbatch, R. J. (2012). Physical processes that impact the evolution of global mean sea level in ocean climate models. . *Ocean Modelling*, **51**. doi: 10.1016/j.ocemod.2012.04.003

Griffies, S. M., M. J. Harrison, R. C. Pacanowski, and A. Rosati, 2004: Technical guide to MOM4, 337 pp. pp.

Griffies, S. M., et al., 2005: Formulation of an ocean model for global climate simulations. *Ocean Sci.*, **1**, 45-79.

Griffies, S. M., et al., 2009: Coordinated Ocean-ice Reference Experiments (COREs). *Ocean Modelling*, **26**, 1-46.

Grose, M., M. Pook, P. McIntosh, J. Risbey, and N. Bindoff, 2012: The simulation of cutoff lows in a regional climate model: reliability and future trends. *Climate Dynamics*, **39**, 445-459.

Güttler, I., I. Stepanov, G. Nikulin, C. Jones, and C. Brankovic, 2012: Precipitation in complex orography simulated by the regional climate model RCA3. *Tellus*. Submitted.

Guilyardi, E., 2006: El Niño - mean state - seasonal cycle interactions in a multi-model ensemble. *Clim. Dyn.*, **26**, 229-348.

Guilyardi, E., P. Braconnot, F. F. Jin, S. T. Kim, M. Kolasinski, T. Li, and I. Musat, 2009a: Atmosphere Feedbacks during ENSO in a Coupled GCM with a Modified Atmospheric Convection Scheme. *Journal of Climate*, **22**, 5698-5718.

Guilyardi, E., et al., 2009b: UNDERSTANDING EL NINO IN OCEAN-ATMOSPHERE GENERAL CIRCULATION MODELS Progress and Challenges. *Bull. Amer. Meteorol. Soc.*, **90**, 325-+.

Gupta, A. S., A. Santoso, A. S. Taschetto, C. C. Ummenhofer, J. Trevena, and M. H. England, 2009: Projected changes to the southern hemisphere ocean and sea ice in the IPCC AR4 climate models. *Journal of Climate*. doi:10.1175/2008jcli2827.1, 3047-3078.

Gurney, K. R., et al., 2003: TransCom 3 CO2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information. *Tellus Series B-Chemical and Physical Meteorology*, **55**, 555-579.

Gutowski, W., et al., 2010: Regional Extreme Monthly Precipitation Simulated by NARCCAP RCMs. *Journal of Hydrometeorology*. doi:DOI 10.1175/2010JHM1297.1, 1373-1379.

Hall, A., and X. Qu, 2006: Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophysical Research Letters*. doi:ARTN L03502, 10.1029/2005GL025127, -.

Hallberg, R., and A. Gnanadesikan, 2006: The role of eddies in determining the structure and response of the wind-driven southern hemisphere overturning: Results from the Modeling Eddies in the Southern Ocean (MESO) project. *Journal of Physical Oceanography*, **36**, 2232-2252.

Halloran, P. R., 2012: Does atmospheric CO2 seasonality play an important role in governing the air-sea flux of CO2? *Biogeosciences*, **9**.

Ham, Y.-G., J. S. Kug, I. S. Kang, F. F. Jin, and A. Timmermann, 2010: Impact of diurnal atmospher-ocean coupling on tropical climate simulations using a coupled GCM. *Climate Dynamics*, **34**, 905-917.

Hamilton, K., 1998: Effects of an imposed Quasi-Biennial Oscillation in a comprehensive troposphere-stratosphere-mesosphere general circulation model. *Journal of the Atmospheric Sciences*, **55**, 2393–2418.

Hannart, A., J. L. Dufresne, and P. Naveau, 2009: Why climate sensitivity may not be so unpredictable. *Geophysical Research Letters*, **36**.

Hannay, C., et al., 2009: Evaluation of Forecasted Southeast Pacific Stratocumulus in the NCAR, GFDL, and ECMWF Models. *Journal of Climate*, **22**, 2871-2889.

Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: GLOBAL SURFACE TEMPERATURE CHANGE. *Reviews of Geophysics*, **48**.

Hansen, J., M. Sato, P. Kharecha, and K. von Schuckmann, 2011: Earth's energy imbalance and implications. *Atmospheric Chemistry and Physics*, **11**, 13421-13449.

Hansen, J., et al., 1983: Efficient Three-Dimensional Global Models for Climate Studies: Models I and II. *Mon. Weath. Rev.*, **111**, -662.

Hansen, J., et al., 1984: Climate Sensitivity: Analysis of Feedback Mechanisms. *Climate Processes and Climate Sensitivity, Geophysical Monograph*, **29**, 130-163.

Hansen, J., et al., 2005: Efficacy of climate forcings. *Journal of Geophysical Research-Atmospheres*, **110**.

Hargreaves, J., 2010: Skill and uncertainty in climate models. *Wiley Interdisciplinary Reviews-Climate Change*, **1**, 556-564.

Hargreaves, J. C., A. Abe-Ouchi, and J. D. Annan, 2007: Linking glacial and future climates through an ensemble of GCM simulations. *Climate of the Past*, **3**, 77-87.

Hargreaves, J. C., A. Paul, R. Ohgaito, A. Abe-Ouchi, and J. D. Annan, 2011: Are paleoclimate model ensembles consistent with the MARGO data synthesis? *Climate of the Past*, **7**, 917-933.

Hargreaves, J. C., J. D. Annan, R. Ohgaito, A. Paul, and A. Abe-Ouchi, 2012: Skill and reliability of climate model ensembles at the Last Glacial Maximum and mid Holocene. *Climate of the Past*. submitted.

Harrison, S. P., et al., 2012: Model benchmarking with glacial and mid-Holocene climates. *Science*. submitted.

Hasegawa, A., and S. Emori, 2007: Effect of air-sea coupling in the assessment of CO2-induced intensification of tropical cyclone activity. *Geophysical Research Letters*, **34**.

Hasumi, H., 2006: CCSR Ocean Component Model (COCO) Version 4.0. Centre for Climate System Research.

Hasumi, H., and S. Emori, 2004: K-1 Coupled GCM (MIROC)Description. Center for Climate System Research, University of Tokyo.

Hawkins, E., and R. Sutton, 2009: The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bull. Amer. Meteorol. Soc.*, **90**, 1095-1107.

Haynes, J. M., C. Jakob, W. B. Rossow, G. Tselioudis, and J. Brown, 2011: Major Characteristics of Southern Ocean Cloud Regimes and Their Effects on the Energy Budget. *Journal of Climate*, **24**, 5061-5080.

Haynes, P. H., 2006: The Latitudinal Structure of the QBO. *Quarterly Journal of the Royal Meteorological Society*, **124**, 2645-2670.

Haywood, J. M., N. Bellouin, A. Jones, O. Boucher, M. Wild, and K. P. Shine, 2011: The roles of aerosol, water vapor and cloud in future global dimming/brightening. *Journal of Geophysical Research-Atmospheres*, **116**.

Hazeleger, W., and R. J. Haarsma, 2005: Sensitivity of tropical Atlantic climate to mixing in a coupled ocean-atmosphere model. *Climate Dynamics*, **25**, 387-399.

Hazeleger, W., et al., 2011: EC-Earth V2.2: description and validation of a new seamless earth system prediction model. *Climate Dynamics*. doi:papers2://publication/doi/10.1007/s00382-011-1228-5, 1-19.

Hegerl, G., and F. Zwiers, 2011: Use of models in detection and attribution of climate change. *Wiley Interdisciplinary Reviews-Climate Change*, **2**.

Hegerl, G. C., et al., 2007: Understanding and Attributing Climate Change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to theFourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis,K.B. Averyt, M. Tignor and H.L. Miller, Ed., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 665-775.

Heinze, C., 2004: Simulating oceanic CaCO3 export production in the greenhouse. *Geophysical Research Letters*, **31**.

Heinze, C., I. Kriest, and E. Maier-Reimer, 2009: Age offsets among different biogenic and lithogenic components of sediment cores revealed by numerical modeling. *Paleoceanography*, **24**.

Held, I. M., 2005: The gap between simulation and understanding in climate modeling. *Bull. Amer. Meteorol. Soc.*, **86**, 1609-+.

Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *Journal of Climate*, **19**, 5686-5699.

Held, I. M., and K. M. Shell, 2012: Using Relative Humidity as a State Variable in Climate Feedback Analysis. *Journal of Climate*, **25**, 2578-2582.

Henson, S. A., D. Raitsos, J. P. Dunne, and A. McQuatters-Gollop, 2009: Decadal variability in biogeochemical models: Comparison with a 50-year ocean colour dataset. *Geophysical Research Letters*, **36**.

Hermes, J. C., and C. J. C. Reason, 2005: Ocean model diagnosis of interannual coevolving SST variability in the South Indian and South Atlantic Oceans. *Journal of Climate*, **18**, 2864-2882.

Hernández-Díaz, L., R. Laprise, L. Sushama, A. Martynov, K. Winger, and B. Dugas: Climate simulation over CORDEX Africa domain using the fifth-generation Canadian Regional Climate Model (CRCM5). *Climate Dynamics*. doi:10.1007/s00382-012-1387-z, 1-19.

Hewitt, H. T., et al., 2011: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system. *Geoscientific Model Development*, **4**.

Hibbard, K. A., G. A. Meehl, P. M. Cox, and P. Friedlingstein, 2007: A Strategy for Climate Change Stabilization Experiments. *Eos Trans.* , **88**.

Hibler, W. D., 1979: A dynamic thermodynamic sea ice model. *Journal of Physical Oceanography*, **9**, 815-846.

Hirai, M., T. Sakashita, H. Kitagawa, T. Tsuyuki, M. Hosaka, and M. Oh'Izumi, 2007: Development and validation of a new land surface model for JMA's operational global model using the CEOP observation dataset. *Journal of the Meteorological Society of Japan*, **85A**.

Hirschi, M., et al., 2011a: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience*, **4**, 17-21.

Hirschi, M., et al., 2011b: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience*. doi:DOI 10.1038/NGEO1032, 17-21.

Hofmann, M., and M. A. Morales Maqueda, 2011: The response of Southern Ocean eddies to increased midlatitude westerlies: A non eddy resolving model study. *Geophysical Research Letters*, **38**, L03605, doi:03610.01029/02010GL045972.

Holden, P. B., N.R. Edwards, K.I.C. Oliver, T.M. Lenton, and R.D. Wilkinson, 2010: A probabilistic calibration of climate sensitivity and terrestrial carbon change in GENIE-1, *Wilkinson Climate Dynamics*, doi: 10.1007/s00382-009-0630-8. http://www.genie.ac.uk/index.htm.

Holian, G. L., A. P. Sokolov, and R. G. Prinn, 2001: Uncertainty in atmospheric CO_2 predictions from a global ocean carbon cycle model. MIT Joint Program on the Science and Policy of Global Change, 25 pp.

Holland, M., D. Bailey, B. Briegleb, B. Light, and E. Hunke, 2012a: Improved sea ice shortwave radiation physics in CCSM4: The impact of melt ponds and aerosols on arctic aea ice. *Journal of Climate*, **25**, 1413-1430.

Holland, M. M., M. C. Serreze, and J. Stroeve, 2010: The sea ice mass budget of the Arctic and its future change as simulated by coupled climate models. *Climate Dynamics*. doi:10.1007/s00382-008-0493-4.

Holland, M. M., D. A. Bailey, B. P. Briegleb, B. Light, and E. Hunke, 2012b: Improved Sea Ice Shortwave Radiation Physics in CCSM4: The Impact of Melt Ponds and Aerosols on Arctic Sea Ice. *Journal of Climate*, **25**.

Holland, P. R., A. Jenkins, and D. M. Holland, 2008: The response of ice shelf basal melting to variations in ocean temperature. *Journal of Climate*, **21**, 2558-2572.

Holton, J. R., 1983: THE INFLUENCE OF GRAVITY-WAVE BREAKING ON THE GENERAL-CIRCULATION OF THE MIDDLE ATMOSPHERE. *Journal of the Atmospheric Sciences*, **40**, 2497-2507.

Holton, J. R., and H. C. Tan, 1980: THE INFLUENCE OF THE EQUATORIAL QUASI-BIENNIAL OSCILLATION ON THE GLOBAL CIRCULATION AT 50 MB. *Journal of the Atmospheric Sciences*, **37**, 2200-2208.

Horowitz, L. W., et al., 2003: A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2. *Journal of Geophysical Research-Atmospheres*, **108**.

Hourdin, F., et al., 2012: LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*. doi:DOI: 10.1007/s00382-012-1343-y.

Hourdin, F., et al., in press: From LMDZ5A to LMDZ5B : revisiting the parameterizations of clouds and convection in the atmospheric component of the IPSL-CM5 climate model. *Climate Dynamics*.

Hourdin, F., et al., 2010: AMMA-MODEL INTERCOMPARISON PROJECT. *Bull. Amer. Meteorol. Soc.*, **91**, 95-+.

Hourdin F., et al., 2012: Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. *Climate Dynamics*. doi:10.1007/s00382-012-1343-y.

Hu, Z.-Z., B. Huang, Y.-T. Hou, W. Wang, F. Yang, C. Stan, and E. Schneider, 2011: Sensitivity of tropical climate to low-level clouds in the NCEP climate forecast system. *Climate Dynamics*, **36**, 1795-1811.

Huber, M., I. Mahlstein, M. Wild, J. Fasullo, and R. Knutti, 2011: Constraints on Climate Sensitivity from Radiation Patterns in Climate Models. *Journal of Climate*, **24**, 1034-1052.

Hunke, E. C., and J. K. Dukowicz, 1997: An elastic-viscous-plastic model for sea ice dynamics. *Journal of Physical Oceanography*, **27**, 1849-1867.

Hunke, E. C., and W. H. Lipscomb, 2011: CICE: the Los Alamos Sea Ice ModelDocumentation and Software User's ManualVersion 4.1. Los Alamos National Laboratory.

Hunke, E. C., W. H. Lipscomb, and A. K. Turner, 2010: Sea Ice Models for Climate Study: Retorspective and New Directions. *Journal of Glaciology*, **56**, 1162-1172.

——, 2011a: Sea Ice Models for Climate Study: Retorspective and New Directions. *Journal of Glaciology*, **56**, 1162-1172.

Hunke, E. C., D. Notz, A. K. Turner, and M. Vancoppenolle, 2011b: The multiphase physics of sea ice : A review for model developers. *The Cryosphere*, **5**, 989-1009.

Hurrell, J., G. A. Meehl, D. Bader, T. L. Delworth, B. Kirtman, and B. Wielicki, 2009: A UNIFIED MODELING APPROACH TO CLIMATE SYSTEM PREDICTION. *Bull. Amer. Meteorol. Soc.*, **90**, 1819-1832.

Hurrell, J., et al., 2012: The Community Earth System Model: A framework for collaborative research. *Bull. Amer. Meteorol. Soc.*, **submitted**.

Hurrell, J. W., et al., 2010: Decadal Climate Prediction: Opportunities and Challenges. *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2)*, Venice, Italy, 21-25 September 2009, ESA Publication

Hurtt, G. C., et al., 2009: Harmonization of global land-use scenarios for the period 1500-2100 for IPCC-AR5. *iLEAPS Newsletter*, **7**, 6-8.

Hutchings, J. K., A. Roberts, C. A. Geiger, and J. Richter-Menge, 2011: Spatial and temporal characterization of sea-ice deformation. *Annals of Glaciology*, **52**, 360-368.

Huybrechts, P., 2002: Sea-level changes at the LGM from ice-dynamic reconstructions of the Greenland and Antarctic ice sheets during the glacial cycles. *Quat. Sci. Rev.*, **21**, 203-231.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, and S. A. Clough, 2003: Evaluation of upper tropospheric water vapor in the NCAR Community Climate Model (CCM3) using modeled and observed HIRS radiances. *J. Geophys. Res.*, **108**, 4037.

Illingworth, A. J., et al., 2007: Cloudnet. *Bull. Amer. Meteor. Soc.*, **88**, 883-898.

Ilyina, T., R. E. Zeebe, E. Maier-Reimer, and C. Heinze, 2009: Early detection of ocean acidification effects on marine calcification. *Global Biogeochemical Cycles*, **23**.

Ilyina, T., K. Six, J. Segschneider, J. Maier-Reimer, H. Li, and I. Nunez-Riboni, 2012: The global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth System Model in different CMIP5 experimental realizations. *JOURNAL OF ADVANCES IN MODELING EARTH SYSTEMS,*.

Inatsu, M., and M. Kimoto, 2009: A Scale Interaction Study on East Asian Cyclogenesis Using a General Circulation Model Coupled with an Interactively Nested Regional Model. *Monthly Weather Review*. doi:DOI 10.1175/2009MWR2825.1, 2851-2868.

Inatsu, M., Y. Satake, M. Kimoto, and N. Yasutomi, 2012: GCM Bias of the Western Pacific Summer Monsoon and Its Correction by Two-Way Nesting System. *Journal of the Meteorological Society of Japan*, **90B**, 1-10.

Ingram, W., 2010: A very simple model for the water vapour feedback on climate change. *Quarterly Journal of the Royal Meteorological Society*, **136**, 30-40.

Inoue, J., J. P. Liu, J. O. Pinto, and J. A. Curry, 2006: Intercomparison of Arctic Regional Climate Models: Modeling clouds and radiation for SHEBA in May 1998. *Journal of Climate*, **19**, 4167-4178.

IPCC, 2007: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC).* Cambridge University Press, 996 pp pp.

——, 2012: *IPCC WGI/WGII Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX).* in press pp.

Ishii, M., and M. Kimoto, 2009: Reevaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections. *Journal of Oceanography*, **65**, 287-299.

Ito, A., and T. Oikawa, 2002: A simulation model of the carbon cycle in land ecosystems (Sim-CYCLE): a description based on dry-matter production theory and plot-scale validation. *Ecol. Model.*, **151**, 143-176.

Iversen, T., et al., 2012: The Norwegian Earth System Model, NorESM1-M. Part 2: Climate Response and Scenario Projections *Geosci. Model Dev.*. **submitted**.

Jackson, C. S., M. K. Sen, G. Huerta, Y. Deng, and K. P. Bowman, 2008a: Error Reduction and Convergence in Climate Prediction. *Journal of Climate*, **21**, 6698-6709.

Jackson, L., R. Hallberg, and S. Legg, 2008b: A parameterization of shear-driven turbulence for ocean climate models. *Journal of Physical Oceanography*, **38**, 1033-1053.

Jacob, D., et al., 2012: Assessing the Transferability of the Regional Climate Model REMO to Different COordinated Regional Climate Downscaling EXperiment (CORDEX) Regions. *Atmosphere*, **3**, 181-199.

Jakob, C., 2010: ACCELERATING PROGRESS IN GLOBAL ATMOSPHERIC MODEL DEVELOPMENT THROUGH IMPROVED PARAMETERIZATIONS Challenges, Opportunities, and Strategies. *Bull. Amer. Meteorol. Soc.*, **91**, 869-+.

Jang, C. J., J. Park, T. Park, and S. Yoo, 2011: Response of the ocean mixed layer depth to global warming and its impact on primary production: a case for the North Pacific Ocean. *Ices Journal of Marine Science*, **68**, 996-1007.

Jansen, E., et al., 2007: Paleoclimate. *Climate Chane 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessement Report of the Intergovernmental Panel on Climate Change*, S. Solomon, et al., Eds., Cambridge University Press.

Jayne, S. R., 2009: The Impact of Abyssal Mixing Parameterizations in an Ocean General Circulation Model. *Journal of Physical Oceanography*, **39**, 1756-1775.

Jiang, J. H., et al., 2012a: Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "A-Train" satellite observations. *Journal of Geophysical Research*. doi:papers2://publication/uuid/EFC082ED-E93E-4838-8497-FF33970F2B1E, submitted.

——, 2012b: Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "A-Train" satellite observations. *Journal of Geophysical Research*, **117**.

Jin, F. F., S. T. Kim, and L. Bejarano, 2006: A coupled-stability index for ENSO. *Geophys. Res. Let.*, **33**, L23708.

Jochum, M., and R. Murtugudde, 2006: Temperature advection by tropical instability waves. *Journal of Physical Oceanography*, **36**, 592-605.

Johanson, C. M., and Q. Fu, 2009: Hadley Cell Widening: Model Simulations versus Observations. *Journal of Climate*, **22**.

John, V., and B. Soden, 2007: Temperature and humidity biases in global climate models and their impact on climate feedbacks. *Geophysical Research Letters*. doi:ARTN L18704, 10.1029/2007GL030429, -.

Johns, T. C., et al., 2003: Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. *Climate Dynamics*, **20**.

Johns, T. C., et al., 2006: The new Hadley Centre Climate Model (HadGEM1): Evaluation of coupled simulations. *Journal of Climate*, **19**.

Johnson, N. C., and S. B. Feldstein, 2010: The Continuum of North Pacific Sea Level Pressure Patterns: Intraseasonal, Interannual, and Interdecadal Variability. *Journal of Climate*, **23**, 851-867.

Joly, M., A. Voldoire, H. Douville, P. Terray, and J. F. Royer, 2007: African monsoon teleconnections with tropical SSTs in a set of IPCC4 coupled models. *Clim. Dyn.*, 1-32.

Jones, A., D. L. Roberts, M. J. Woodage, and C. E. Johnson, 2001: Indirect sulphate aerosol forcing in a climate model with an interactive sulphur cycle. *Journal of Geophysical Research-Atmospheres*, **106**.

Jones, G. S., P. A. Stott, and N. Christidis, 2012: Attribution of observed historical near surface temperature variations to anthropogenic and natural causes using CMIP5 simulations. *J. Geophys. Res.*, submitted.

Jones, P. D., M. New, D. E. Parker, S. Martin, and I. G. Rigor, 1999: Surface air 	temperature and its variations over the last 150 years. *Reviews of Geophysics*, **37**, 173-199.

Joseph, S., A. K. Sahai, B. N. Goswami, P. Terray, S. Masson, and J. J. Luo, 2012: Possible role of warm SST bias in the simulation of boreal summer monsoon in SINTEX-F2 coupled model. *Climate Dynamics*, **38**, 1561-1576.

Joussaume, S., and K. E. Taylor, 1995: Status of the Paleoclimate Modeling Intercomparison Project. *in Proceedings of the first international AMIP scientific conference, WCRP-92, Monterey, USA*. 425-430.

Jun, M., R. Knutti, and D. W. Nychka, 2008a: Spatial Analysis to Quantify Numerical Model Bias and Dependence: How Many Climate Models Are There? *Journal of the American Statistical Association*, **103**, 934-947.

Jun, M., R. Knutti, and D. Nychka, 2008b: Spatial Analysis to Quantify Numerical Model Bias and Dependence: How Many Climate Models Are There? *Journal of the American Statistical Association*. doi:DOI 10.1198/016214507000001265, 934-947.

Jung, T., et al., 2012: High-Resolution Global Climate Simulations with the ECMWF Model in Project Athena: Experimental Design, Model Climate, and Seasonal Forecast Skill. *Journal of Climate*, **25**, 3155-3172.

Jungclaus, J. H., et al., 2006: Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM. *Journal of Climate*, **19**, 3952-3972.

Jungclaus, J. H., et al., 2012: Characteristics of the ocean simulations in MPIOM, the ocean componentof the MPI-Earth System Model. *JOURNAL OF ADVANCES IN MODELING EARTH SYSTEMS,*, **submitted**.

Jungclaus, J. H., et al., 2010: Climate and carbon-cycle variability over the last millennium. *Climate of the Past*, **6**, 723-737.

Kageyama, M., et al., 2006: Last Glacial Maximum temperatures over the North Atlantic, Europe and western Siberia: a comparison between PMIP models, MARGO sea-surface temperatures and pollen-based reconstructions. *Quat. Sci. Rev.*, **25**, 2082-2102.

Kakitha, K., Y. Peings, P. Terray, and H. Douville, 2011: Indian Summer Monsoon and its relationship with ENSO as simulated by the two French CMIP5 coupled GCMs. submitted ed.

Kalnay, E., et al., 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteorol. Soc.*, **77**, 437-471.

Kanada, S., M. Nakano, S. Hayashi, T. Kato, M. Nakamura, K. Kurihara, and A. Kitoh, 2008: Reproducibility of Maximum Daily Precipitation Amount over Japan by a High-resolution Non-hydrostatic Model. *SOLA*, **4**, 105-108.

Kanamaru, H., and M. Kanamitsu, 2007: Fifty-seven-year California Reanalysis Downscaling at 10 km (CaRD10). Part II: Comparison with North American Regional Reanalysis. *Journal of Climate*, **20**, 5572-5592.

Kanamitsu, M., and L. DeHaan, 2011: The Added Value Index: A new metric to quantify the added value of regional models. *Journal of Geophysical Research-Atmospheres*, **116**, -.

Kanzow, T., et al., 2010: Seasonal variability of the Atlantic meridional overturning circulation at 26.5°N. *Journal of Climate*, **23**, 5678–5698.

Karlsson, J., and G. Svensson, 2010: The simulation of Arctic clouds and their influence on the winter surface temperature in present-day climate in the CMIP3 multi-model dataset. *Clim. Dyn.*, DOI 10.1007/s00382-00010-00758-00386.

Karlsson, J., G. Svensson, and H. Rodhe, 2008: Cloud radiative forcing of subtropical low level clouds in global models. *Climate Dynamics*, **30**, 779-788.

Karnauskas, K., G. Johnson, and R. Murtugudde, 2012: An Equatorial Ocean Bottleneck in Global Climate Models. J. Climate, 343-349.

Karpechko, A., N. Gillett, G. Marshall, and A. Scaife, 2008: Stratospheric influence on circulation changes in the Southern Hemisphere troposphere in coupled climate models. *Geophysical Research Letters*. doi:ARTN L20806, 10.1029/2008GL035354, -.

Karpechko, A. Y., N. P. Gillett, G. J. Marshall, and J. A. Screen, 2009: Climate impacts of the southern annular mode simulated by the CMIP3 models. (vol 22, pg 3751, 2009). *Journal of Climate*, **22**, 6149-6150.

Kattsov, V. M., et al., 2010: Arctic sea-ice change: a grand challenge of climate science. *Journal of Glaciology*, **56**, 1115-1121.

Kawatani, Y., K. Hamilton, and S. Watanabe, 2011: The quasi-biennial oscillation in a double CO2 climate. *Journal of the Atmospheric Sciences*. doi:doi: 10.1175/2010JAS3623.1.

Kawazoe, S., and W. Gutowski, 2012: Regional, extreme daily precipitation in NARCCAP simulations, **Submitted**.

Keeley, S. P. E., R. T. Sutton, and L. C. Shaffrey, 2012: The impact of North Atlantic sea surface temperature errors on the simulation of North Atlantic European region climate. *Quarterly Journal of the Royal Meteorological Society*, **in press**.

Kendon, E. J., N. M. Roberts, C. A. Senior, and M. J. Roberts, 2012: Realism of rainfall in a very high resolution regional climate model. *Journal of Climate*. doi:10.1175/jcli-d-11-00562.1.

Khairoutdinov, M. F., D. A. Randall, and C. DeMott, 2005: Simulations of the Atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *Journal of the Atmospheric Sciences*, **62**, 2136-2154.

Kharin, V. V., F. W. Zwiers, X. B. Zhang, and G. C. Hegerl, 2007: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate*, **20**, 1419-1444.

Kharin, V. V., F. W. Zwiers, X. Zhang, and M. Wehner, 2012: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*. submitted.

Khvorostyanov, D. V., G. Krinner, P. Ciais, M. Heimann, and S. A. Zimov, 2008a: Vulnerability of permafrost carbon to global warming. Part I: model description and role of heat generated by organic matter decomposition. *Tellus Series B-Chemical and Physical Meteorology*, **60**, 250-264.

Khvorostyanov, D. V., P. Ciais, G. Krinner, S. A. Zimov, C. Corradi, and G. Guggenberger, 2008b: Vulnerability of permafrost carbon to global warming. Part II: sensitivity of permafrost carbon stock to global warming. *Tellus Series B-Chemical and Physical Meteorology*, **60**, 265-275.

Kidston, J., and E. P. Gerber, 2010: Intermodel variability of the poleward shift of the austral jet stream in the CMIP3 integrations linked to biases in 20th century climatology. *Geophys. Res. Lett.*, **37**, L09708-.

Kiehl, J. T., 2007: Twentieth century climate model response and climate sensitivity. *Geophysical Research Letters*, **34**.

Kim, D., et al., 2012a: The tropical subseasonal variability simulated in the NASA GISS general circulation model. *Journal of Climate*, **25**, 4641-4659.

Kim, D., et al., 2009: Application of MJO Simulation Diagnostics to Climate Models. *Journal of Climate*, **22**, 6413-6436.

Kim, H.-J., K. Takata, B. Wang, M. Watanabe, M. Kimoto, T. Yokohata, and T. Yasunari, 2011: Global Monsoon, El Niño, and Their Interannual Linkage Simulated by MIROC5 and the CMIP3 CGCMs. *Journal of Climate*, **24**, 5604-5618.

Kim, J., et al., 2012b: Evaluation of the CORDEX-Africa multi-RCM hindcast experiment using the Regional Climate Model Evaluation System (RCMES). *Climatic Change*, **Submitted**.

Kim, S., and F.-F. Jin, 2010a: An ENSO stability analysis. Part I: results from a hybrid coupled model. *Climate Dynamics*. 1-15.

——, 2010b: An ENSO stability analysis. Part II: results from the twentieth and twenty-first century simulations of the CMIP3 models. *Climate Dynamics*. 1-19.

Kim, S. T., and J.-Y. Yu, 2012: The two types of ENSO in CMIP5 models. *Geophysical Research Letters*, **39**.

Kinne, S., et al., 2006: An AeroCom initial assessment - optical properties in aerosol component modules of global models. *Atmospheric Chemistry and Physics*, **6**, 1815-1834.

Kirkevåg, K., et al., 2012: Aerosol-climate interactions in the Norwegian Earth System Model – NorESM. *Geosci. Model Dev*, **submitted**.

Kistler, R., et al., 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteorol. Soc.*, **82**, 247-267.

Kjellstrom, E., G. Nikulin, U. Hansson, G. Strandberg, and A. Ullerstig, 2011: 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations. *Tellus Series a-Dynamic Meteorology and Oceanography*, **63**, 24-40.

Kjellstrom, E., F. Boberg, M. Castro, J. Christensen, G. Nikulin, and E. Sanchez, 2010: Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. *Climate Research*. doi:DOI 10.3354/cr00932, 135-150.

Klein, P., and G. Lapeyre, 2009: The Oceanic Vertical Pump Induced by Mesoscale and Submesoscale Turbulence. *Annual Review of Marine Science*, **1**, 351-375.

Klein, S. A., and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Monthly Weather Review*, **127**, 2514-2531.

Klein, S. A., B. J. Soden, and N. C. Lau, 1999: Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. *Journal of Climate*, **12**, 917-932.

Klein, S. A., X. Jiang, J. Boyle, S. Malyshev, and S. Xie, 2006: Diagnosis of the summertime warm and dry bias over the U.S. Southern Great Plains in the GFDL climate model using a weather forecasting approach. *Geophys. Res. Lett.*, **33**, L18805.

Klein, S. A., Y. Zhang, M. D. Zelinka, R. Pincus, J. S. Boyle, and P. J. Glecker, 2012: Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator. *Journal of Geophysical Research*. doi:papers2://publication/uuid/8DF5BFDD-8734-4DED-9565-9F1C783FF391, submitted.

Klocke, D., R. Pincus, and J. Quaas, 2011a: On constraining estimates of climate sensitivity with present-day observations through model weighting. *Journal of Climate*. in press.

——, 2011b: On Constraining Estimates of Climate Sensitivity with Present-Day Observations through Model Weighting. *Journal of Climate*, **24**, 6092-6099.

——, 2011c: On constraining estimates of climate sensitivity. *Journal of Climate*, **accepted**.

Kloster, S., et al., 2010: Fire dynamics during the 20th century simulated by the Community Land Model. *Biogeosciences*, **7**, 1877-1902.

Knight, C., et al., 2007: Association of parameter, software, and hardware variation with large-scale behavior across 57,000 climate models. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 12259-12264.

Knight, J. R., 2009: The Atlantic Multidecadal Oscillation Inferred from the Forced Climate Response in Coupled General Circulation Models. *Journal of Climate*, **22**, 1610-1625.

Knutson, T. R., et al., 2006: Assessment of twentieth-century regional surface temperature trends using the GFDL CM2 coupled models. *Journal of Climate*, **19**, 1624-1651.

Knutti, R., 2008: Why are climate models reproducing the observed global surface warming so well? *Geophysical Research Letters*, **35**.

——, 2010: The end of model democracy? *Climatic Change*, **102**, 395-404.

Knutti, R., and L. Tomassini, 2008: Constraints on the transient climate response from observed global temperature and ocean heat uptake. *Geophysical Research Letters*, **35**.

Knutti, R., and G. C. Hegerl, 2008: The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nature Geoscience*, **1**, 735-743.

Knutti, R., D. Masson, and A. Gettelman, 2012: Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, **submitted**.

Knutti, R., G. A. Meehl, M. R. Allen, and D. A. Stainforth, 2006: Constraining climate sensitivity from the seasonal cycle in surface temperature. *Journal of Climate*, **19**, 4224-4233.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, 2010a: Challenges in Combining Projections from Multiple Climate Models. *Journal of Climate*, **23**, 2739-2758.

Knutti, R., G. Abramowitz, M. Collins, V. Eyring, P. J. Gleckler, B. Hewitson, and L. Mearns, 2010b: Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections. *In: Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, and P.M. Midgley (eds.)]. IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland.*

Koch, D., et al., 2011: Coupled Aerosol-Chemistry-Climate Twentieth-Century Transient Model Investigation: Trends in Short-Lived Species and Climate Responses. *Journal of Climate*, **24**.

Koldunov, N. V., D. Stammer, and J. Marotzke, 2010: Present-day Arctic sea ice variability in the coupled ECHAM5/MPI-OM model. *Journal of Climate*, **23**, 2520–2543.

Koltzow, M., T. Iversen, and J. Haugen, 2008: Extended Big-Brother experiments: the role of lateral boundary data quality and size of integration domain in regional climate modelling. *Tellus Series a-Dynamic Meteorology and Oceanography*, **60**, 398-410.

Komuro, Y., et al., 2012: Sea-Ice in Twentieth-Century Simulations by New MIROC Coupled Models: A Comparison between Models with High Resolution and with Ice Thickness Distribution. *Journal of the Meteorological Society of Japan*, **90A**.

Kostopoulou, E., K. Tolika, I. Tegoulias, C. Giannakopoulos, S. Somot, C. Anagnostopoulou, and P. Maheras, 2009: Evaluation of a regional climate model using in situ temperature observations over the Balkan Peninsula. *Tellus Series a-Dynamic Meteorology and Oceanography*. doi:DOI 10.1111/j.1600-0870.2009.00389.x, 357-370.

Koven, C. J., et al., 2011: Permafrost carbon-climate feedbacks accelerate global warming. *Proceedings of the National Academy of Sciences*, **108**, 14769-14774.

Kowalczyk, E. A., Y. P. Wang, R. M. Law, H. L. Davies, J. L. McGregor, and G. Abramowitz 2006: The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model

Kowalczyk, E. A., et al., 2012: The land surface model component of ACCESS: description and impact on the simulated surface climatology. *Aust. Met. Oceanog. J.,* , **submitted**.

Kravtsov, S., and C. Spannagle, 2008: Multidecadal climate variability in observed and modeled surface temperatures. *Journal of Climate*, **21**, 1104-1121.

Krinner, G., et al., 2005: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, **19**.

Krüger, L., R. da Rocha, M. Reboita, and T. Ambrizzi, 2012: RegCM3 nested in HadAM3 scenarios A2 and B2: projected changes in extratropical cyclogenesis, temperature and precipitation over the South Atlantic Ocean. *Climatic Change*, **113**, 599-621.

Kusaka, H., T. Takata, and Y. Takane, 2010: Reproducibility of Regional Climate in Central Japan Using the 4-km Resolution WRF Model. *Sola*, **6**, 113-116.

Kusunoki, S., R. Mizuta, and M. Matsueda, 2011: Future changes in the East Asian rain band projected by global atmospheric models with 20-km and 60-km grid size. *Climate Dynamics*, **37**, 2481-2493.

Kwon, Y. O., and S. C. Riser, 2004: North Atlantic Subtropical Mode Water: A history of ocean-atmosphere interaction 1961–2000. *Geophysical Research Letters*, **31**, L19307-L19307.

Kwon, Y. O., M. A. Alexander, N. A. Bond, C. Frankignoul, H. Nakamura, B. Qiu, and L. Thompson, 2010: Role of the Gulf Stream and Kuroshio-Oyashio Systems in Large-Scale Atmosphere-Ocean Interaction: A Review. *Journal of Climate*, **23**, 3249-3281.

L'Ecuyer, T., and G. Stephens, 2007: The Tropical Atmospheric Energy Budget from the TRMM Perspective. Part II: Evaluating GCM Representations of the Sensitivity of Regional Energy and Water Cycles to the 1998–99 ENSO Cycle. *J. Climate*, **20**, 4548-4571.

Laepple, T., and P. Huybers, 2012: Consistend proxy demonstration of highly variable Holocene marine temperatures. submitted.

Laine, A., G. Lapeyre, and G. Riviere, 2011: A Quasigeostrophic Model for Moist Storm Tracks. *Journal of the Atmospheric Sciences*, **68**, 1306-1322.

Laine, A., M. Kageyama, P. Braconnot, and R. Alkama, 2009: Impact of Greenhouse Gas Concentration Changes on Surface Energetics in IPSL-CM4: Regional Warming Patterns, Land-Sea Warming Ratios, and Glacial-Interglacial Differences. *Journal of Climate*, **22**, 4621-4635.

Lamarque, J. F., et al., 2012: CAM-chem: description and evaluation of interactive atmospheric chemistry in the Community Earth System Model. *Geoscientific Model Development*, **5**.

Lamarque, J. F., et al., 2010: Historical (1850-2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmospheric Chemistry and Physics*, **10**, 7017-7039.

Lambert, F. H., G. R. Harris, M. Collins, J. M. Murphy, D. M. H. Sexton, and B. B. B. Booth, 2012: Interactions between perturbations to different Earth system components simulated by a fully-coupled climate model. *Climate Dynamics*. submitted.

Lambert, S., and G. Boer, 2001: CMIP1 evaluation and intercomparison of coupled climate models. *Climate Dynamics*, **17**, 83-106.

Landerer, F., W., P. Gleckler, J., and L. Tong, submitted: Evaluation of dynamic sea surface height in CMIP3 and CMIP5 models against satellite observations. *Journal of Geophysical Research*.

Landrum, L., M. M. Holland, D. P. Schneider, and E. Hunke, 2012: Antarctic sea ice climatology, variability and late 20th century change in CCSM4. *Journal of Climate*.

Langenbrunner, B., and J. D. Neelin, 2012: Analyzing ENSO teleconnections in CMIP models as a measure of model fidelity in simulating precipitation. submitted ed., J. Climate.

Laprise, R., 2008: Regional climate modelling. *Journal of Computational Physics*, **227**, 3641-3666.

Laprise, R., et al., 2008: Challenging some tenets of Regional Climate Modelling. *Meteorology and Atmospheric Physics*, **100**, 3-22.

Large, W., and S. Yeager, 2009: The global climatology of an interannually varying airâ“sea flux data set. *Climate Dynamics*, **33**, 341-364.

Larow, T. E., Y. K. Lim, D. W. Shin, E. P. Chassignet, and S. Cocke, 2008: Atlantic basin seasonal hurricane simulations. *Journal of Climate*, **21**, 3191-3206.

Lawrence, D. M., et al., 2012: The CCSM4 Land Simulation, 1850-2005: Assessment of Surface Climate and New Capabilities. *Journal of Climate*, **25**.

Lawrence, D. M., et al., 2011: Parameterization Improvements and Functional and Structural Advances in Version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, **3**, 2011MS000045.

Le Quere, C., et al., 2005: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, **11**, 2016-2040.

Le Quere, C., et al., 2009: Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, **2**, 831-836.

Leduc, M., and R. Laprise, 2009: Regional climate model sensitivity to domain size. *Climate Dynamics*, **32**, 833-854.

Lee, D. S., et al., 2009: Aviation and global climate change in the 21st century. *Atmos. Environ.*, **43**, 3520-3537.

Legg, S., L. Jackson, and R. W. Hallberg, 2008: Eddy-resolving modeling of overflows. *Eddy resolving ocean models*, 177 ed., M. Hecht, and H. Hasumi, Eds., American Geophysical Union, 63-82.

Legg, S., et al., 2009: IMPROVING OCEANIC OVERFLOW REPRESENTATION IN CLIMATE MODELS The Gravity Current Entrainment Climate Process Team. *Bull. Amer. Meteorol. Soc.*, **90**, 657-+.

LeGrande, A. N., et al., 2006: Consistent simulations of multiple proxy responses to an abrupt climate change event. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 837-842.

Leloup, J., M. Lengaigne, and J.-P. Boulanger, 2008: Twentieth century ENSO characteristics in the IPCC database. *Clim. Dyn.*, **30**, 277-291.

Lemoine, D. M., 2010: Climate Sensitivity Distributions Dependence on the Possibility that Models Share Biases. *Journal of Climate*, **23**, 4395-4415.

Lenaerts, J., M. van den Broeke, S. Dery, E. van Meijgaard, W. van de Berg, S. Palm, and J. Rodrigo, 2012: Modeling drifting snow in Antarctica with a regional climate model: 1. Methods and model evaluation. *Journal of Geophysical Research-Atmospheres*, **117**.

Lenderink, G., 2010: Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations. *Climate Research*. doi:DOI 10.3354/cr00946, 151-166.

Lenderink, G., and E. Van Meijgaard, 2008: Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geoscience*, **1**, 511-514.

Levine, R. C., and A. G. Turner, 2012: Dependence of Indian monsoon rainfall on moisture fluxes across the Arabian Sea and the impact of coupled model sea surface temperature biases. *Climate Dynamics*, **38**, 2167-2190.

Levis, S., 2010: Modeling vegetation and land use in models of the Earth System. *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 840–856.

Levis, S., G. B. Bonan, and P. J. Lawrence, 2007: Present-day springtime high-latitude surface albedo as a predictor of simulated climate sensitivity. *Geophysical Research Letters*, **34**, 4.

Levitus, S., J. I. Antonov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, and A. V. Mishonov, 2009: Global ocean heat content 1955-2008 in light of recently revealed instrumentation problems. *Geophysical Research Letters*, **36**.

Levy, H., L. W. Horowitz, M. D. Schwarzkopf, Y. Ming, J.-C. Golaz, V. Naik, and V. Ramaswamy, 2012: The Roles of Aerosol Direct and Indirect Effects in Past and Future Climate Change,. *Journal of Geophysical Research*, **submitted**.

Lewis, S. L., P. M. Brando, O. L. Phillips, G. M. F. van der Heijden, and D. Nepstad, 2011: The 2010 Amazon Drought. *Science*, **331**, 554-554.

Lewis, T., and S. Lamoureux, 2010: Twenty-first century discharge and sediment yield predictions in a small high Arctic watershed. *Global and Planetary Change*, **71**, 27-41.

Li, C., J.-S. von Storch, and J. Marotzke, 2012a: Deep-ocean heat uptake and equilibrium climate response. *Climate Dynamics*. 10.1007/s00382-00012-01350-z.

Li, H. B., A. Robock, and M. Wild, 2007: Evaluation of Intergovernmental Panel on Climate Change Fourth Assessment soil moisture simulations for the second half of the twentieth century. *Journal of Geophysical Research-Atmospheres*, **112**.

Li, J.-L. F., D. E. Waliser, and J. H. Jiang, 2011a: Correction to "Comparisons of satellites liquid water estimates to ECMWF and GMAO analyses, 20th century IPCC AR4 climate simulations, and GCM simulations". *Geophys. Res. Lett.*, **38**, L24807.

Li, J., S.-P. X. and A. Mestas-Nunez, E. R. C. and Gang Huang, R. D'Arrigo, F. Liu, J. Ma, and X. Zheng, 2011b: Interdecadal Modulation of ENSO Amplitude During the Last Millennium. *Nature Climate Change*.

Li, J. L. F., et al., 2012b: An observationally-based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary reanalyses using contemporary satellite data. *Journal of Geophysical Research*. doi:papers2://publication/uuid/F0065C08-7313-42EE-B950-A0A57CA600CB, submitted.

——, 2012c: An observationally-based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary reanalyses using contemporary satellite data. *Journal of Geophysical Research*. doi:10.1029/2012JD017640, in press.

Li, T., and G. H. Philander, 1996: On the annual cycle in the eastern equatorial Pacific. *J. Climate*, **9**, 2986-2998.

Lienert, f., J. C. Fyfe, and W. J. Merryfield, 2011: Do Climate Models Capture the Tropical Influences on North Pacific sea surface temperature variability? *Journal of Climate*. doi:10.1175/JCLI-D-11-00205.1.).

Lin, A. L., and T. Li, 2008: Energy Spectrum Characteristics of Boreal Summer Intraseasonal Oscillations: Climatology and Variations during the ENSO Developing and Decaying Phases. *Journal of Climate*, **21**, 6304-6320.

Lin, J.-L., 2007: The Double-ITCZ Problem in IPCC AR4 Coupled GCMs: Ocean-Atmosphere Feedback Analysis. *Journal of Climate*, **20**, 4497-4525.

Lin, J. L., et al., 2006: Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: Convective signals. *Journal of Climate*, **19**, 2665-2690.

Lin, P., Y. Yongqiang, and H. Liu, 2012a: Long-term Stability and Oceanic Mean State Simulated by the Coupled Model FGOALS-s2. *Advances in Atmospheric Sciences*. doi:doi: 10.1007/s00376-012-2042-7.

Lin, Y., et al., 2012b: TWP-ICE global atmospheric model intercomparison: Convection responsiveness and resolution impact. *J. Geophys. Res.*, **117**, D09111.

Lindsay, K., et al., 2012: Preindustrial Control and 20th Century Carbon Cycle Experiments with the Earth System Model CESM1-(BGC). *J. Climate*, **submitted**.

Lindvall, J., G. Svensson, and C. Hannay, 2012: Evaluation of near-surface parameters in the two versions of the atmospheric model in CESM1 using flux station observations. *Journal of Climate*. doi:10.1175/jcli-d-12-00020.1.

Linkin, M., and S. Nigam, 2008: The north pacific oscillation-west Pacific teleconnection pattern: Mature-phase structure and winter impacts. *Journal of Climate*, **21**, 1979-1997.

Lipscomb, W. H., et al., 2012: Implementation and initial evaluation of the Glimmer Community Ice Sheet Model in the Community Earth System Model. *J. Climate*, **submitted**.

Liu, J., 2010: Sensitivity of sea ice and ocean simulations to sea ice salinity in a coupled global climate model. *Science China-Earth Sciences*, **53**.

Liu, L., W. Yu, and T. Li, 2011: Dynamic and Thermodynamic Air–Sea Coupling Associated with the Indian Ocean Dipole Diagnosed from 23 WCRP CMIP3 Models. *Journal of Climate*, **24**, 4941–4958.

Liu, S. C., C. B. Fu, C. J. Shiu, J. P. Chen, and F. T. Wu, 2009: Temperature dependence of global precipitation extremes. *Geophysical Research Letters*, **36**, 4.

Liu, X. H., et al., 2007: Uncertainties in global aerosol simulations: Assessment using three meteorological data sets. *Journal of Geophysical Research-Atmospheres*, **112**.

Liu, Y., 1996: Modeling the emissions of nitrous oxide and methane from the terrestrial biosphere to the atmosphere. MIT Joint Program on the Science and Policy of Global Change, 219 pp.

Lloyd, J., E. Guilyardi, and H. Weller, 2010: The role of atmosphere feedbacks during ENSO in the CMIP3 models. Part II: using AMIP runs to understand the heat flux feedback mechanisms. *Climate Dynamics*. 1-22.

Lloyd, J., E. Guilyardi, H. Weller, and J. Slingo, 2009: The role of atmosphere feedbacks during ENSO in the CMIP3 models. *Atmospheric Science Letters*, **10**, 170-176.

Loeb, N. G., et al., 2009: Toward Optimal Closure of the Earth's Top-of-Atmosphere Radiation Budget. *Journal of Climate*, **22**, 748-766.

Loeb, N. G., et al., 2012: Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty. *Nature Geoscience*, **in press**.

Lohmann, U., K. von Salzen, N. McFarlane, H. G. Leighton, and J. Feichter, 1999: Tropospheric sulfur cycle in the Canadian general circulation model. *Journal of Geophysical Research-Atmospheres*, **104**.

Long, M. C., K. Lindsay, S. Peacock, J. K. Moore, and S. C. Doney, 2012: 20th century oceanic carbon uptake and storage in CESM1-BGC. *J. Climate*, **submitted**.

Loptien, U., O. Zolina, S. Gulev, M. Latif, and V. Soloviov, 2008: Cyclone life cycle characteristics over the Northern Hemisphere in coupled GCMs. *Climate Dynamics*, **31**, 507-532.

Lorenz, P., and D. Jacob, 2005: Influence of regional scale information on the global circulation: A two-way nesting climate simulation. *Geophysical Research Letters*, **32**, -.

Losch, M., D. Menemenlis, J.-M. Campin, P. Heimbach, and C. Hill, 2011: On the formulation of sea-ice models. Part 1: Effects of different solver implementations and parameterizations. *Ocean Modelling*.

Lott, F., and M. J. Miller, 1997: A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quarterly Journal of the Royal Meteorological Society*, **123**, 101-127.

Loutre, M. F., A. Mouchet, T. Fichefet, H. Goosse, H. Goelzer, and P. Huybrechts, 2011: Evaluating climate model performance with various parameter sets using observations over the recent past. *Climate of the Past*, **7**, 511-526.

Loyola, D., and M. Coldewey-Egbers, 2012: Multi-sensor data merging with stacked neural networks for the creation of satellite long-term climate data records. *Eurasip Journal on Advances in Signal Processing*. doi:10.1186/1687-6180-2012-91.

Loyola, D., et al., 2009: Global long-term monitoring of the ozone layer - a prerequisite for predictions. *International Journal of Remote Sensing*, **30**, 4295-4318.

Lu, J., G. A. Vecchi, and T. Reichler, 2007: Expansion of the Hadley cell under global warming. *Geophysical Research Letters*, **34**.

Lucarini, V., and F. Ragone, 2011: ENERGETICS OF CLIMATE MODELS: NET ENERGY BALANCE AND MERIDIONAL ENTHALPY TRANSPORT. *Reviews of Geophysics*, **49**.

Lucas-Picher, P., S. Somot, M. Déqué, B. Decharme, and A. Alias, 2012a: Evaluation of the regional climate model ALADIN to simulate the climate over North America in the CORDEX framework. *Climate Dynamics*, **Submitted**.

Lucas-Picher, P., M. Wulff-Nielsen, J. Christensen, G. Adalgeirsdottir, R. Mottram, and S. Simonsen, 2012b: Very high resolution regional climate model simulations over Greenland: Identifying added value. *Journal of Geophysical Research-Atmospheres*, **117**.

Lumpkin, R., K. G. Speer, and K. P. Koltermann, 2008: Transport across 48Â°N in the Atlantic Ocean. *Journal of Physical Oceanography*, **38**, 733-752.

Luo, J. J., S. Masson, E. Roeckner, G. Madec, and T. Yamagata, 2005: Reducing climatology bias in an ocean-atmosphere CGCM with improved coupling physics. *Journal of Climate*, **18**, 2344-2360.

Lynn, B., R. Healy, and L. Druyan, 2009: Quantifying the sensitivity of simulated climate change to model configuration. *Climatic Change*, **92**, 275-298.

MacKinnon, J., et al., 2009: Using global arrays to investigate internal-waves and mixing. *OceanObs09: Sustained Ocean Observations and Information for Society*, Venice, Italy, ESA.

Madden, R. A., and P. R. Julian, 1972: Description of Global-Scale Circulation Cells in Tropics with a 40-50 Day Period. *Journal of the Atmospheric Sciences*, **29**, 1109-&.

——, 1994: Observations of the 40-50-Day Tropical Oscillation - a Review. *Monthly Weather Review*, **122**, 814-837.

Madec, G., 2008: NEMO ocean engine. IPSL.

Madec, G., P. Delecluse, M. Imbard, and C. Levy, 1998: OPA 8.1 ocean general circulation model reference manual, 91 pp. pp.

Mahlstein, I., and R. Knutti, 2010: Regional climate change patterns identified by cluster analysis. *Climate Dynamics*, **35**, 587-600.

Mahlstein, I., P. R. Gent, and S. Solomon, 2012: Historical Antarctic mean sea ice area, sea ice trends, and winds in CMIP5 simulations. *Journal of Climate*.

Maier-Reimer, E., I. Kriest, J. Segschneider, and P. Wetze, 2005: The HAMburg Ocean Carbon Cycle Model HAMOCC 5.1.

Manabe, S., 1969: CLIMATE AND THE OCEAN CIRCULATION1. *Monthly Weather Review*, **97**, 739-774.

Manabe, S., and K. Bryan, 1985: CO2-induced change in a coupled ocean-atmosphere model and its paleoclimatic implications. *Journal of Geophysical Research*, **90**, 1689-1707.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteorol. Soc.*, **78**, 1069-1079.

Manzini, E., and e. al., 2012: Surface Climate Change in the CMIP5 simulations: Northern winter response and the role of Stratospheric Variability. *Journal of Geophysical Research-Atmospheres*, **submitted**.

Manzini, E., M. A. Giorgetta, M. Esch, L. Kornblueh, and E. Roeckner, 2006: The influence of sea surface temperatures on the northern winter stratosphere: Ensemble simulations with the MAECHAM5 model. *Journal of Climate*, **19**, 3863-3881.

Maraun, D., 2012: Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. *Geophys. Res. Lett.*, **39**, L06706.

Maraun, D., H. Rust, and T. Osborn, 2010a: Synoptic airflow and UK daily precipitation extremes Development and validation of a vector generalised linear model. *Extremes*, **13**, 133-153.

Maraun, D., et al., 2010b: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**, RG3003.

Marchand, R., N. Beagley, and T. P. Ackerman, 2009: Evaluation of Hydrometeor Occurrence Profiles in the Multiscale Modeling Framework Climate Model Using Atmospheric Classification. *Journal of Climate*, **22**, 4557-4573.

Maris, M., B. de Boer, and J. Oerlemans, 2012: A climate model intercomparison for the Antarctic region: present and past. *Climate of the Past*, **8**, 803-814.

Markovic, M., H. Lin, and K. Winger, 2010: Simulating Global and North American Climate Using the Global Environmental Multiscale Model with a Variable-Resolution Modeling Approach. *Monthly Weather Review*, **138**, 3967-3987.

Marsh, D. R., M. J. Mills, D. E. Kinnison, J.-F. Lamarque, N. Calvo, and L. M. Polvani, 2012: Climate change from 1850 to 2005 simulated in CESM1(WACCM). *J. Climate*, **submitted**.

Marsh, R., S. A. Mueller, A. Yool, and N. R. Edwards, 2011: Incorporation of the C-GOLDSTEIN efficient climate model into the GENIE framework: "eb_go_gs" configurations of GENIE. *Geoscientific Model Development*, **4**, 957-992.

Marsh, R., et al., 2009: Recent changes in the North Atlantic circulation simulated with eddy-permitting and eddy-resolving ocean models. *Ocean Modelling*, **28**, 226-239.

Marshall, A., and A. A. Scaife, 2009: Impact of the Quasi-Biennial Oscillation on seasonal forecasts. *Journal of Geophysical Research*, **114**, D18110.

Marsland, S. J., et al., 2012: Evaluation of ACCESS Climate Model ocean diagnostics in CMIP5 simulations. *Aust. Met. Oceanog. J.*, **submitted**.

Marti, O., et al., 2010: Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution. *Climate Dynamics*, **34**, 1-26.

Martin, G. M., et al., 2011: The HadGEM2 family of Met Office Unified Model climate configurations. *Geoscientific Model Development*, **4**.

Martius, O., L. M. Polvani, and H. C. Davies, 2009: Blocking precursors to stratospheric sudden warming events. *Geophysical Research Letters*, **36**.

Masarie, K. A., and P. P. Tans, 1995: EXTENSION AND INTEGRATION OF ATMOSPHERIC CARBON-DIOXIDE DATA INTO A GLOBALLY CONSISTENT MEASUREMENT RECORD. *Journal of Geophysical Research-Atmospheres*, **100**.

Masato, G., B. Hoskins, and T. Woollings, 2012: Winter and summer Northern Hemisphere blocking in CMIP5 models. *Journal of Climate*. submitted.

Masson-Delmotte, V., et al., 2010: EPICA Dome C record of glacial and interglacial intensities. *Quat. Sci. Rev.*, **29**, 113-128.

Masson-Delmotte, V., et al., 2006: Past and future polar amplification of climate change: climate model intercomparisons and ice-core constraints (vol 26, pg 513, 2006). *Climate Dynamics*, **27**, 437-440.

Masson, D., and R. Knutti, 2011a: Spatial-Scale Dependence of Climate Model Performance in the CMIP3 Ensemble. *Journal of Climate*, **24**, 2680-2692.

Masson, D., and R. Knutti, 2011b: Climate model genealogy. *Geophysical Research Letters*, **38**.

——, 2012: Predictor Screening, Calibration and Observational Constraints in Climate Model Ensembles: an Illustration using Climate Sensitivity. *Journal of Climate*. submitted.

Massonnet, F., T. Fichefet, H. Goosse, C. M. Bitz, G. Philippon-Berthier, M. M. Holland, and P.-Y. Barriat, 2012: Constraining projections of summer Arctic sea ice. *The Cryosphere*, **submitted**.

Mastrandrea, M. D., et al., 2011: Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties. Intergovernmental Panel on Climate Change (IPCC).

Materia, S., P. A. Dirmeyer, Z. C. Guo, A. Alessandri, and A. Navarra, 2010: The Sensitivity of Simulated River Discharge to Land Surface Representation and Meteorological Forcings. *Journal of Hydrometeorology*, **11**, 334-351.

Matsueda, M., 2009: Blocking Predictability in Operational Medium-Range Ensemble Forecasts. *Sola*, **5**, 113-116.

Matsueda, M., R. Mizuta, and S. Kusunoki, 2009a: Future change in wintertime atmospheric blocking simulated using a 20-km-mesh atmospheric global circulation model. *J. Geophys. Res.*, **114**, D12114-.

Matsueda, M., R. Mizuta, and S. Kusunoki, 2009b: Future change in wintertime atmospheric blocking simulated using a 20-km-mesh atmospheric global circulation model. *Journal of Geophysical Research-Atmospheres*, **114**.

Matsueda, M., H. Endo, and R. Mizuta, 2010a: Future change in Southern Hemisphere summertime and wintertime atmospheric blockings simulated using a 20-km-mesh AGCM. *Geophys. Res. Lett.*, **37**, L02803-.

Matsueda, M., H. Endo, and R. Mizuta, 2010b: Future change in Southern Hemisphere summertime and wintertime atmospheric blockings simulated using a 20-km-mesh AGCM. *Geophysical Research Letters*, **37**.

Matsumoto, K., K. S. Tokos, A. R. Price, and S. J. Cox, 2008: First description of the Minnesota Earth System Model for Ocean biogeochemistry (MESMO 1.0). *Geoscientific Model Development*, **1**, 1-15.

Maurer, E., and H. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences*, **12**, 551-563.

Mauritsen, T., et al., 2012: Tuning the climate of a global model. *Journal of Advances in Modeling Earth Systems*. doi:papers2://publication/uuid/DC747E24-EBCD-4F54-9A13-8900A8BD26EF, submitted.

May, P. T., J. H. Mather, G. Vaughan, K. N. Bower, C. Jakob, G. M. McFarquhar, and G. G. Mace, 2008: The Tropical Warm Pool International Cloud Experiment. *Bull. Amer. Meteorol. Soc.*, **89**, 629-645.

May, W., 2007: The simulation of the variability and extremes of daily precipitation over Europe by the HIRHAM regional climate model. *Global and Planetary Change*, **57**, 59-82.

Maykut, G. A., and N. Untersteiner, 1971: Some results from a time dependent thermodynamic model of sea ice. *Journal of Geophysical Research*, **76**, 1550-1575.

McClean, J., M. Maltrud, and F. Bryan, 2006a: Measures of the fidelity of eddying ocean models. *Oceanography*, **19**.

McClean, J. L., and J. C. Carman, 2011: Investigation of IPCC AR4 coupled climate model North Atlantic modewater formation. *Ocean Modelling*, **40**, 14–34.

McClean, J. L., M. E. Maltrud, and F. O. Bryan, 2006b: Measures of the fidelity of eddying ocean models. *Oceanography*, **19**, 104-117.

McClean, J. L., et al., 2011: A prototype two-decade fully-coupled fine-resolution CCSM simulation. *Ocean Modelling*, **39**, 10-30.

McCormack, J. P., S. D. Eckermann, D. E. Siskind, and T. J. McGee, 2006: CHEM2D-OPP: A new linearized gas-phase ozone photochemistry parameterization for high-altitude NWP and climate models. *Atmospheric Chemistry and Physics*, **6**.

McCrary, R. R., and D. A. Randall, 2010: Great Plains Drought in Simulations of the Twentieth Century. *Journal of Climate*, **23**, 2178-2196.

McCreary, J. P., and P. Lu, 1994: Interaction between the Subtropical and Equatorial Ocean Circulations: The Subtropical Cell. *Journal of Physical Oceanography*, **24**, 466-497.

McDonald, R. E., 2011: Understanding the impact of climate change on Northern Hemisphere extra-tropical cyclones. *Climate Dynamics*, **37**, 1399-1425.

McDougall, T. J., and P. C. McIntosh, 2001: The temporal-residual-mean velocity. Part II: Isopycnal interpretation and the tracer and momentum equations. *Journal of Physical Oceanography*, **31**, 1222-1246.

McFarlane, N. A., 1987: THE EFFECT OF OROGRAPHICALLY EXCITED GRAVITY-WAVE DRAG ON THE GENERAL-CIRCULATION OF THE LOWER STRATOSPHERE AND TROPOSPHERE. *Journal of the Atmospheric Sciences*, **44**, 1775-1800.

McKitrick, R., S. McIntyre, and C. Herman, 2010: Panel and multivariate methods for tests of trend equivalence in climate data series. *Atmospheric Science Letters*, **11**, 270-277.

McLandress, C., 2002: Interannual Variations of the Diurnal Tide in the mesosphere induced by a zonal mean wind oscillation in the tropics. *Geophysical Research Letters*, **29**.

McLandress, C., T. Shepherd, J. Scinocca, D. Plummer, M. Sigmond, A. Jonsson, and M. Reader, 2011: Separating the Dynamical Effects of Climate Change and Ozone Depletion. Part II Southern Hemisphere Troposphere. *Journal of Climate*, **24**, 1850-1868.

McLaren, A. J., et al., 2006: Evaluation of the sea ice simulation in a new coupled atmosphere-ocean climate model (HadGEM1). *Journal of Geophysical Research-Oceans*, **111**.

McManus, J. F., R. Francois, J. M. Gherardi, L. D. Keigwin, and S. Brown-Leger, 2004: Collapse and rapid resumption of Atlantic meridional circulation linked to deglacial climate changes. *Nature*, **428**, 834-837.

McPhaden, M. J., and D. X. Zhang, 2002: Slowdown of the meridional overturning circulation in the upper Pacific Ocean. *Nature*, **415**, 603-608.

McWilliams, J. C., 2008: The nature and consequences of oceanic eddies. *Ocean Modeling in an eddying regime*, M. Hecht, and H. Hasumi, Eds., American Geophysical Union, 5-15.

Mearns, L. O., et al., 2012: The North American Regional Climate Change Assessment Program: Overview of Phase I Results. *Bull. Amer. Meteorol. Soc.* doi:10.1175/bams-d-11-00223.1.

Mears, C. A., F. J. Wentz, P. Thorne, and D. Bernie, 2011: Assessing uncertainty in estimates of atmospheric temperature changes from MSU and AMSU using a Monte-Carlo estimation technique. *J. Geophys. Res.*, **116**, D08112.

Mears, C. A., B. D. Santer, F. J. Wentz, K. E. Taylor, and M. F. Wehner, 2007: Relationship between temperature and precipitable water changes over tropical oceans. *Geophysical Research Letters*, **34**, 10.1029/2007GL031936.

Meehl, G. A., P. R. Gent, J. M. Arblaster, B. L. Otto-Bliesner, E. C. Brady, and A. Craig, 2001: Factors that affect the amplitude of El Niño in global coupled climate models. *Clim. Dyn.*, **17**, 515.

Meehl, G. A., J. M. Arblaster, J. M. Caron, H. Annamalai, M. Jochum, A. Chakraborty, and R. Murtugudde, 2012: Monsoon Regimes and Processes in CCSM4. Part I: The Asian-Australian Monsoon. *Journal of Climate*, **25**, 2583-2608.

Meehl, G. A., et al., 2007a: The WCRP CMIP3 multimodel dataset - A new era in climate change research. *Bull. Amer. Meteorol. Soc.*, **88**, 1383-+.

Meehl, G. A., et al., 2007b: Global Climate Projections. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, et al., Eds., Cambridge University Press, 747-845.

Meehl, G. A., et al., 2009: DECADAL PREDICTION Can It Be Skillful? *Bull. Amer. Meteorol. Soc.*, **90**, 1467-1485.

Meijers, A., E. Shuckburgh, N. Bruneau, J.-B. Sallee, T. Bracegirdle, and Z. Wang, 2012: Representation of the Antarctic Circumpolar Current in the CMIP5 climate models and future changes under warming scenarios. submitted ed., JGR Oceans.

Meinshausen, M., et al., 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213-241.

Meissner, K. J., A. J. Weaver, H. D. Matthews, and P. M. Cox, 2003: The role of land surface dynamics in glacial inception: a study with the UVic Earth System Model. *Climate Dynamics*, **21**, 515-537.

Melia, D. S., 2002: A global coupled sea ice-ocean model. *Ocean Modelling*, **4**.

Melillo, J. M., A. D. McGuire, D. W. Kicklighter, B. Moore, C. J. Vorosmarty, and A. L. Schloss, 1993: GLOBAL CLIMATE-CHANGE AND TERRESTRIAL NET PRIMARY PRODUCTION. *Nature*, **363**, 234-240.

Melsom, A., V. Lien, and W. P. Budgell, 2009: Using the Regional Ocean Modeling System (ROMS) to improve the ocean circulation from a GCM 20th century simulation. *Ocean Dyn.*, **59**, 969-981.

Menary, M., W. Park, K. Lohmann, M. Vellinga, D. Palmer, M. Latif, and J. H. Jungclaus, 2011: A multimodel comparison of centennial Atlantic meridional overturning circulation variability. *Climate Dynamics*. doi:10.1007/s00382-011-1172-4.

Menendez, C., M. de Castro, A. Sorensson, J. Boulanger, and C. M. Grp, 2010: CLARIS Project: towards climate downscaling in South America. *Meteorologische Zeitschrift*. doi:DOI 10.1127/0941-2948/2010/0459, 357-362.

Menon, S., D. Koch, G. Beig, S. Sahu, J. Fasullo, and D. Orlikowski, 2010: Black carbon aerosols and the third polar ice cap. *Atmospheric Chemistry and Physics*, **10**.

Mercado, L. M., C. Huntingford, J. H. C. Gash, P. M. Cox, and V. Jogireddy, 2007: Improving the representation of radiation interception and photosynthesis for climate model applications. *Tellus Series B-Chemical and Physical Meteorology*, **59**.

Merrifield, M. A., and M. E. Maltrud, 2011: Regional sea level trends due to a Pacific trade wind intensification. *Geophysical Research Letters*, **38**.

Merryfield, W. J., et al., 2012: The Canadian Seasonal to Interannual Prediction System. Part I: Models and Initialization. *Monthly Weather Review*, **submitted**.

Météo-France, 2011: ARPEGE-Climate Version 5.2. Météo-France.

Mieville, A., et al., 2010: Emissions of gases and particles from biomass burning during the 20th century using satellite data and an historical reconstruction. *Atmos. Environ.*, **44**, 1469-1477.

Miller, A., et al., 2002: A cohesive total ozone data set from the SBUV(/2) satellite system. *Journal of Geophysical Research-Atmospheres*, **107**.

Miller, R. L., G. A. Schmidt, and D. T. Shindell, 2006: Forced annular variations in the 20th century intergovernmental panel on climate change fourth assessment report models. *Journal of Geophysical Research-Atmospheres*, **111**.

Milliff, R., A. Bonazzi, C. Wikle, N. Pinardi, and L. Berliner, 2011: Ocean ensemble forecasting. Part I: Ensemble Mediterranean winds from a Bayesian hierarchical model. *Quarterly Journal of the Royal Meteorological Society*, **137**, 858-878.

Milly, P. C. D., and A. B. Shmakin, 2002: Global modeling of land water and energy balances. Part I: the land dynamics (LaD) model. *Journal of Hydrometeorology*, **3**.

Min, S. K., X. B. Zhang, F. W. Zwiers, and G. C. Hegerl, 2011: Human contribution to more-intense precipitation extremes. *Nature*, **470**, 376-379.

Minobe, S., 1997: A 50-70 year climatic oscillation over the North Pacific and North America. *Geophysical Research Letters*, **24**, 683-686.

——, 1999: Resonance in bidecadal and pentadecadal climate oscillations over the North Pacific: Role in climatic regime shifts. *Geophysical Research Letters*, **26**, 855-858.

Misra, V., 2007: Addressing the issue of systematic errors in a regional climate model. *Journal of Climate*, **20**, 801-818.

Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, **25**.

Miyama, T., and M. Kawamiya, 2009: Estimating allowable carbon emission for CO(2) concentration stabilization using a GCM-based Earth system model. *Geophysical Research Letters*, **36**.

Mizuta, R., et al., 2012: Climate Simulations Using MRI-AGCM3.2 with 20-km Grid. *Journal of the Meteorological Society of Japan*, **90A**, 233-258.

Mohr, M., 2008: New routines for gridding of temperature and precipitation observations for "seNorge.no". Norwegian Meteorological Institute, 40 pp.

Montoya, M., A. Griesel, A. Levermann, J. Mignot, M. Hofmann, A. Ganopolski, and S. Rahmstorf, 2005: The earth system model of intermediate complexity CLIMBER-3 alpha. Part 1: description and performance for present-day conditions. *Climate Dynamics*, **25**, 237-263.

Moore, J. K., K. Lindsay, S. C. Doney, M. Long, and K. Misumi, 2012: Marine Ecosystem Dynamics and Biogeochemical Cycling in the Community Earth System Model (CESM1-BGC). . *J. Climate*, **submitted**.

Morgenstern, O., et al., 2010a: Anthropogenic forcing of the Northern Annular Mode in CCMVal-2 models. *J. Geophys. Res.*, **115**, D00M03-.

Morgenstern, O., et al., 2010b: Anthropogenic forcing of the Northern Annular Mode in CCMVal-2 models. *Journal of Geophysical Research-Atmospheres*, **115**.

Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones, 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research-Atmospheres*, **117**.

Moss, R. H., et al., 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747-756.

Mouchet, A., and L. M. François, 1996: Sensitivity of a global oceanic carbon cycle model to the circulation and the fate of organic matter: preliminary results. *Physics and Chemistry of the Earth*, **21**, 511-516.

Msadek, R., and C. Frankignoul, 2009: Atlantic multidecadal oceanic variability and its influence on the atmosphere in a climate model. *Climate Dynamics*, **33**, 45-62.

Muller, S. A., F. Joos, N. R. Edwards, and T. F. Stocker, 2006: Water mass distribution and ventilation time scales in a cost-efficient, three-dimensional ocean model. *Journal of Climate*, **19**, 5479-5499.

Murakami, H., and M. Sugi, 2010: Effect of Model Resolution on Tropical Cyclone Climate Projections. *SOLA*, **6**, 73-76.

Murakami, H., et al., 2012: Future Changes in Tropical Cyclone Activity Projected by the New High-Resolution MRI-AGCM. *Journal of Climate*, **25**, 3237-3260.

Murphy, J., B. Booth, M. Collins, G. Harris, D. Sexton, and M. Webb, 2007: A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*. doi:DOI 10.1098/rsta.2007.2077, 1993-2028.

Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, -772.

Murtugudde, R., J. Beauchamp, C. R. McClain, M. Lewis, and A. J. Busalacchi, 2002: Effects of penetrative radiation on the upper tropical ocean circulation. *Journal of Climate*, **15**, 470-486.

Muryshev, K. E., A. V. Eliseev, I. I. Mokhov, and N. A. Diansky, 2009: Validating and assessing the sensitivity of the climate model with an ocean general circulation model developed at the Institute of Atmospheric Physics, Russian Academy of Sciences. *Izvestiya Atmospheric and Oceanic Physics*, **45**, 416-433.

Najjar, R. G., et al., 2007: Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from Phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2). *Global Biogeochemical Cycles*, **21**.

Nakano, H., H. Tsujino, M. Hirabara, T. Yasuda, T. Motoi, M. Ishii, and G. Yamanaka, 2011: Uptake mechanism of anthropogenic CO2 in the Kuroshio Extension region in an ocean general circulation model. *Journal of Oceanography*, **67**.

Naoe, H., and K. Shibata, 2010: Equatorial quasi-biennial oscillation influence on northern winter extratropical circulation. *Journal of Geophysical Research-Atmospheres*, **115**.

Neale, R. B., J. H. Richter, and M. Jochum, 2008: The Impact of Convection on ENSO: From a Delayed Oscillator to a Series of Events. *Journal of Climate*, **21**, 5904-5924.

Neale, R. B., J. Richter, S. Park, P. H. Lauritzen, S. J. Vavrus, P. J. Rasch, and M. Zhang, 2012: The Mean Climate of the Community Atmosphere Model (CAM4) in Forced SST and Fully Coupled Experiments. *J. Climate*.

Neale, R. B., et al., 2010: Description of the NCAR Community Atmosphere Model (CAM 4.0).

Neelin, J. D., 2007: Moist dynamics of tropical convection zones in monsoons, teleconnections and global warming., In The Global Circulation of the Atmosphere, T. Schneider and A. Sobel, Eds, Princeton University Press, Princeton, 385pp.

Neelin, J. D., and N. Zeng, 2000: A quasi-equilibrium tropical circulation model - Formulation. *Journal of the Atmospheric Sciences*, **57**, 1741-1766.

Neelin, J. D., C. Chou, and H. Su, 2003: Tropical drought regions in global warming and El Nino teleconnections. *Geophysical Research Letters*, **30**.

Neelin, J. D., A. Bracco, H. Luo, J. C. McWilliams, and J. E. Meyerson, 2010: Considerations for parameter optimization and sensitivity in climate models. *Proceedings of the National Academy of Sciences*, **107**, 21349-21354.

Neggers, R. A. J., 2009: A Dual Mass Flux Framework for Boundary Layer Convection. Part II: Clouds. *Journal of the Atmospheric Sciences*, **66**, 1489-1506.

Neggers, R. A. J., M. Kohler, and A. C. M. Beljaars, 2009: A Dual Mass Flux Framework for Boundary Layer Convection. Part I: Transport. *Journal of the Atmospheric Sciences*, **66**, 1465-1487.

Nicolsky, D. J., V. E. Romanovsky, V. A. Alexeev, and D. M. Lawrence, 2007: Improved modeling of permafrost dynamics in a GCM land-surface scheme. *Journal of Geophysical Research*, **34**, L08501.

Nikiema, O., and R. Laprise, 2010: Diagnostic budget study of the internal variability in ensemble simulations of the Canadian RCM. *Climate Dynamics*. doi:10.1007/s00382-010-0834-y, 1-25.

Nikulin, G., et al., 2012: Precipitation Climatology in An Ensemble of CORDEX-Africa Regional Climate Simulations. *Journal of Climate*. doi:10.1175/jcli-d-11-00375.1.

Ning, L., M. E. Mann, R. Crane, and T. Wagener, 2011: Probabilistic Projections of Climate Change for the Mid-Atlantic Region of the United States - Validation of Precipitation Downscaling During the Historical Era. *Journal of Climate*, **Early online release**.

Nishii, K., et al., 2011: Relationship of the reproducibility of multiple variables among global climate models. *Journal of the Meteorological Society of Japan*. submitted.

Nishii, K., et al., 2012: Relationship of the Reproducibility of Multiple Variables among Global Climate Models, **90A**, 87-100.

Nomura, D., H. Yoshikawa-Inoue, T. Toyota, and K. Shirasawa, 2010: Effects of snow, snowmelting and refreezing processes on air-sea-ice CO2 flux. *Journal of Glaciology*, **56**, 262-270.

Notz, D., 2012: Sea-ice extent provides a limited metric of model performance. *Journal of Geophysical Research*.

Notz, D., F. A. Haumann, H. Haak, J. H. Jungclaus, and J. Marotzke, 2012: Sea-ice evolution in the Arctic as modeled by MPI-ESM. *Journal of Advances in Modeling Earth Systems*, **submitted**.

O'Connor, F. M., C. E. Johnson, O. Morgenstern, and W. J. Collins, 2009: Interactions between tropospheric chemistry and climate model temperature and humidity biases. *Geophysical Research Letters*, **36**.

O'Farrell, S. P., 1998: Investigation of the dynamic sea ice component of a coupled atmosphere sea ice general circulation model. *Journal of Geophysical Research-Oceans*, **103**.

O'ishi, R., and A. Abe-Ouchi, 2011: Polar amplification in the mid-Holocene derived from dynamical vegetation change with a GCM. *Geophysical Research Letters*, **38**.

Oden, J. T., and S. Prudhomme, 2002: Estimation of modeling error in computational mechanics. *Journal of Computational Physics*, **182**, 496-515.

Ogura, T., M. J. Webb, A. Bodas-Salcedo, K. D. Williams, T. Yokohata, and D. R. Wilson, 2008a: Comparison of Cloud Response to CO(2) Doubling in Two GCMs. *Sola*, **4**, 29-32.

Ogura, T., S. Emori, M. J. Webb, Y. Tsushima, T. Yokohata, A. Abe-Ouchi, and M. Kimoto, 2008b: Towards understanding cloud response in atmospheric GCMs: The use of tendency diagnostics. *Journal of the Meteorological Society of Japan*, **86**, 69-79.

Ohba, M., D. Nohara, and H. Ueda, 2010: Simulation of Asymmetric ENSO Transition in WCRP CMIP3 Multimodel Experiments. *Journal of Climate*, **23**, 6051-6067.

Ohgaito, R., and A. Abe-Ouchi, 2009: The effect of sea surface temperature bias in the PMIP2 AOGCMs on mid-Holocene Asian monsoon enhancement. *Climate Dynamics*, **33**, 975-983.

Oka, A., E. Tajika, A. Abe-Ouchi, and K. Kubota, 2011: Role of the ocean in controlling atmospheric CO2 concentration in the course of global glaciations. *Climate Dynamics*, **37**, 1755-1770.

Oki, T., T. Nishimura, and P. Dirmeyer, 1999: Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP). *J. Meteorol. Soc. Jap.*, **77**, 235-255.

Oleson, K. W., 2004: Technical Description of the <st1:place w:st="on"><st1:placename w:st="on">Community <st1:placetype w:st="on">Land Model (CLM). National Center for Atmospheric Research.

Oleson, K. W., G. B. Bonan, J. Feddema, M. Vertenstein, and C. S. B. Grimmond, 2008a: An urban parameterization for a global climate model. Part I: Formulation and evaluation for two cities. *Journal of Applied Meteorology and Climatology*, **47**, 1038-1060.

Oleson, K. W., et al., 2010: Technical Description of version 4.0 of the Community Land Model (CLM) NATIONAL CENTER FOR ATMOSPHERIC RESEARCH.

Oleson, K. W., et al., 2008b: Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research-Biogeosciences*, **113**.

Oleson, K. W., et al., 2008c: Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research-Biogeosciences*, **113**.

Opsteegh, J. D., R. J. Haarsma, F. M. Selten, and A. Kattenberg, 1998: ECBILT: a dynamic alternative to mixed boundary conditions in ocean models. *Tellus Series a-Dynamic Meteorology and Oceanography*, **50**, 348-367.

Oshima, K., and Y. Tanimoto, 2009: An Evaluation of Reproducibility of the Pacific Decadal Oscillation in the CMIP3 Simulations. *Journal of the Meteorological Society of Japan*, **87**, 755-770.

Ottera, O. H., M. Bentsen, H. Drange, and L. Suo, 2010: External forcing as a metronome for Atlantic multidecadal variability. *Nature Geoscience*, **3**, 688-694.

Otto-Bliesner, B. L., et al., 2007a: Last Glacial Maximum ocean thermohaline circulation: PMIP2 model intercomparisons and data constraints. *Geophysical Research Letters*, **34**, -.

——, 2007b: Last Glacial Maximum ocean thermohaline circulation: PMIP2 model intercomparisons and data constraints. *Geophys. Res. Lett.*, **34**, L12706, doi:12710.11029/12007GL029475.

Otto-Bliesner, B. L., et al., 2009: A comparison of PMIP2 model simulations and the MARGO proxy reconstruction for tropical sea surface temperatures at last glacial maximum. *Climate Dynamics*, **32**, 799-815.

Otto, J., T. Raddatz, M. Claussen, V. Brovkin, and V. Gayler, 2009: Separation of atmosphere-ocean-vegetation feedbacks and synergies for mid-Holocene climate. *Global Biogeochemical Cycles*, **23**, -.

Ozturk, T., H. Altinsoy, M. Turkes, and M. Kurnaz, 2012: Simulation of temperature and precipitation climatology for the Central Asia CORDEX domain using RegCM 4.0. *Climate Research*, **52**, 63-76.

Paeth, H., 2011: Postprocessing of simulated precipitation for impact research in West Africa. Part I: model output statistics for monthly data. *Climate Dynamics*, **36**, 1321-1336.

Paeth, H., M. Rauthe, and S.-K. Min, 2008: Multi-model Bayesian assessment of climate change in the northern annular mode. *Global and Planetary Change*, **60**, 193-206.

Paeth, H., et al., 2011: Progress in regional downscaling of west African precipitation. *Atmospheric Science Letters*. doi:10.1002/asl.306, n/a-n/a.

Palmer, J. R., and I. J. Totterdell, 2001: Production and export in a global ocean ecosystem model. *Deep-Sea Research Part I-Oceanographic Research Papers*, **48**, 1169-1198.

Palmer, T. N., G. J. Shutts, and R. Swinbank, 1986: ALLEVIATION OF A SYSTEMATIC WESTERLY BIAS IN GENERAL-CIRCULATION AND NUMERICAL WEATHER PREDICTION MODELS THROUGH AN OROGRAPHIC GRAVITY-WAVE DRAG PARAMETRIZATION. *Quarterly Journal of the Royal Meteorological Society*, **112**, 1001-1039.

Pan, L. L., and T. Li, 2008: Interactions between the tropical ISO and midlatitude low-frequency flow. *Climate Dynamics*, **31**, 375-388.

Parekh, P., F. Joos, and S. A. Muller, 2008: A modeling assessment of the interplay between aeolian iron fluxes and iron-binding ligands in controlling carbon dioxide fluctuations during Antarctic warm events. *Paleoceanography*, **23**.

Park, S., and C. S. Bretherton, 2009: The University of Washington Shallow Convection and Moist Turbulence Schemes and Their Impact on Climate Simulations with the Community Atmosphere Model. *Journal of Climate*, **22**, 3449-3469.

Park, W., and M. Latif, 2010: Pacific and Atlantic multidecadal variability in the Kiel Climate Model. *Geophysical Research Letters*, **37**.

Parker, D., C. Folland, A. Scaife, J. Knight, A. Colman, P. Baines, and B. W. Dong, 2007: Decadal to multidecadal variability and the climate change background. *Journal of Geophysical Research-Atmospheres*, **112**.

Patricola, C. M., Li, M., Xu, Y., Chang, P., Saravana, R., & Hsieh, J-S. (2012). An Investigation of Tropical Atlantic Bias in a High-Resolution Coupled Regional Climate Model. *Climate Dynamics*. doi: 10.1007/s00382-012-1320-5

Pavlova, T. V., V. M. Kattsov, and V. A. Govorkova, 2011: Sea ice in CMIP5 models: closer to reality? *Trudy GGO (MGO Proceedings, in Russian)*, **564**, 7-18.

Pavlova, T. V., V. M. Kattsov, Y. D. Nadyozhina, P. V. Sporyshev, and V. A. Govorkova, 2007: Terrestrial cryosphere evolution through the 20th and 21st centuries as simulated with the new generation of global climate models. *Earth Cryosphere (in Russian)*, **11**, 3-13.

Pechony, O., and D. T. Shindell, 2009: Fire parameterization on a global scale. *Journal of Geophysical Research-Atmospheres*, **114**.

Pedersen, C. A., E. Roeckner, M. Lüthje, and J. Winther, 2009: A new sea ice albedo scheme including melt ponds for ECHAM5 general circulation model. *Journal of Geophysical Research*, **114, D08101**.

Pennell, C., and *T. Reichler*, 2011a: **On the Effective Number of Climate Models**. *Journal of Climate*. doi:10.1175/2010JCLI3814.1.

Pennell, C., and T. Reichler, 2011b: On the Effective Number of Climate Models. *Journal of Climate*, **24**, 2358-2367.

Perlwitz, J., S. Pawson, R. Fogt, J. Nielsen, and W. Neff, 2008: Impact of stratospheric ozone hole recovery on Antarctic climate. *Geophysical Research Letters*. doi:ARTN L08714, 10.1029/2008GL033317, -.

Peterson, T. C., et al., 2009: State of the Climate in 2008. *Bull. Amer. Meteorol. Soc.*, **90**, S1-S196.

Petoukhov, V., I. I. Mokhov, A. V. Eliseev, and V. A. Semenov, 1998: The IAP RAS global climate model.

Petoukhov, V., A. Ganopolski, V. Brovkin, M. Claussen, A. Eliseev, C. Kubatzki, and S. Rahmstorf, 2000: CLIMBER-2: a climate system model of intermediate complexity. Part I: model description and performance for present climate. *Climate Dynamics*, **16**, 1-17.

Petoukhov, V., et al., 2005: EMIC Intercomparison Project (EMIP-CO2): comparative analysis of EMIC simulations of climate, and of equilibrium and transient responses to atmospheric CO2 doubling. *Climate Dynamics*, **25**, 363-385.

Pfeiffer, A., and G. Zängl, 2010: Validation of climate-mode MM5-simulations for the European Alpine Region. *Theoretical and Applied Climatology*, **101**, 93-108.

Phillips, T. J., et al., 2004: Evaluating Parameterizations in General Circulation Models: Climate Simulation Meets Weather Prediction. *Bull. Amer. Meteorol. Soc.*, **85**, 1903-1915.

Piani, C., D. J. Frame, D. A. Stainforth, and M. R. Allen, 2005: Constraints on climate change from a multi-thousand member ensemble of simulations. *Geophysical Research Letters*, **32**, L23825.

Picard, G., F. Domine, G. Krinner, L. Arnaud, and E. Lefebvre, 2012: Inhibition of the positive snow-albedo feedback by precipitation in interior Antarctica. *Nature Climate Change*. doi:10.1038/NCLIMATE1590.

Pierce, D. W., 2001: Distinguishing coupled ocean-atmosphere interactions from background noise in the North Pacific. *Progress in Oceanography*, **49**, 331-352.

Pierce, D. W., T. P. Barnett, B. D. Santer, and P. J. Gleckler, 2009: Selecting global climate models for regional climate change studies. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 8441-8446.

Pincus, R., C. P. Batstone, R. J. P. Hofmann, K. E. Taylor, and P. J. Glecker, 2008a: Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models. *Journal of Geophysical Research-Atmospheres*, **113**.

Pincus, R., C. Batstone, R. Hofmann, K. Taylor, and P. Glecker, 2008b: Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models. *Journal of Geophysical Research-Atmospheres*. doi:ARTN D14209

DOI 10.1029/2007JD009334, -.

Pincus, R., S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. P. Hofmann, 2012: Reconciling GCM-simulated and satellite-observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators. *Journal of Climate*, **Early Online Release**.

Pinto, J. G., T. Spangehl, U. Ulbrich, and P. Speth, 2006: Assessment of winter cyclone activity in a transient ECHAM4-OPYC3 GHG experiment. *Meteorologische Zeitschrift*, **15**, 279-291.

Piot, M., and R. von Glasow, 2008a: The potential importance of frost flowers, recycling on snow, and open leads for ozone depletion events. *Atmospheric Chemistry and Physics*, **8**, 2437–2467.

——, 2008b: The potential importance of frost flowers, recycling on snow, and open leads for ozone depletion events. *Atmospheric Chemistry and Physics*, **8**, 2437–2467.

Pitman, A., A. Arneth, and L. Ganzeveld, 2010: Review. Regionalizing global climate models. *International Journal of Climatology*, **Published online**.

Pitman, A. J., 2003: The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, **23**, 479-510.

Pitman, A. J., et al., 2009: Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study. *Geophysical Research Letters*, **36**.

Plattner, G. K., et al., 2008: Long-term climate commitments projected with climate-carbon cycle models. *Journal of Climate*, **21**, 2721-2751.

Ploshay, J. J., and N.-C. Lau, 2010: Simulation of the Diurnal Cycle in Tropical Rainfall and Circulation during Boreal Summer with a High-Resolution GCM. *Monthly Weather Review*, **138**, 3434-3453.

Po-Chedley, S., and Q. Fu, 2012: Discrepancies in tropical upper tropospheric warming between atmospheric circulation models and satellites. *Environmental Research Letters*. submitted.

Polvani, L., D. Waugh, G. Correa, and S. Son, 2011: Stratospheric Ozone Depletion: The Main Driver of Twentieth-Century Atmospheric Circulation Changes in the Southern Hemisphere. *Journal of Climate*, **24**, 795-812.

Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton, 2000: The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics*, **16**, 123-146.

Power, S., and R. Colman, 2006: Multi-year predictability in a coupled general circulation model. *Climate Dynamics*, **26**.

Power, S., M. Haylock, R. Colman, and X. D. Wang, 2006: The predictability of interdecadal changes in ENSO activity and ENSO teleconnections. *Journal of Climate*, **19**, 4755-4771.

Prentice, I. C., S. P. Harrison, D. Jolly, and J. Guiot, 1998: The climate and biomes of Europe at 6000 yr BP: Comparison of model simulations and pollen-based reconstructions. *Quat. Sci. Rev.*, **17**, 659-668.

Probst, P., R. Rizzi, E. Tosi, V. Lucarini, and T. Maestri, 2012: Total cloud cover from satellite observations and climate models. *Atmospheric Research*, **107**, 161-170.

Prudhomme, C., and H. Davies, 2009: Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: baseline climate. *Climatic Change*, **93**, 177-195.

Pryor, S., G. Nikulin, and C. Jones, 2012: Influence of spatial resolution on regional climate model derived wind climates. *Journal of Geophysical Research-Atmospheres*, **117**.

Qian, T. T., A. Dai, K. E. Trenberth, and K. W. Oleson, 2006: Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations. *Journal of Hydrometeorology*, **7**, 953-975.

Qu, X., and A. Hall, 2012: On the Persistent Spread in Snow-Albedo Feedback. *Climate Dynamics*, **submitted**.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski, 2005: Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, **133**, 1155-1174.

Raible, C. C., M. Yoshimori, T. F. Stocker, and C. Casty, 2007: Extreme midlatitude cyclones and their implications for precipitation and wind speed extremes in simulations of the Maunder Minimum versus present day conditions. *Climate Dynamics*, **28**, 409-423.

Raible, C. C., T. F. Stocker, M. Yoshimori, M. Renold, U. Beyerle, C. Casty, and J. Luterbacher, 2005: Northern Hemispheric trends of pressure indices and atmospheric circulation patterns in observations, reconstructions, and coupled GCM simulations. *Journal of Climate*.

Raisanen, J., 2007: How reliable are climate models? *Tellus Series a-Dynamic Meteorology and Oceanography*, **59**, 2-29.

Raisanen, J., and J. S. Ylhaisi, 2011: How Much Should Climate Model Output Be Smoothed in Space? *Journal of Climate*, **24**, 867-880.

Raisanen, J., L. Ruokolainen, and J. Ylhaisi, 2010: Weighting of model results for improving best estimates of climate change. *Climate Dynamics*, **35**, 407-422.

Rammig, A., et al., 2010: Estimating the risk of Amazonian forest dieback. *New Phytologist*, **187**, 694-706.

Ramstein, G., M. Kageyama, J. Guiot, H. Wu, C. Hely, G. Krinner, and S. Brewer, 2007: How cold was Europe at the Last Glacial Maximum? A synthesis of the progress achieved since the first PMIP model-data comparison. *Climate of the Past*, **3**, 331-339.

Randall, D. A., M. F. Khairoutdinov, A. Arakawa, and W. W. Grabowski, 2003: Breaking the Cloud Parameterization Deadlock. *Bull. Amer. Meteorol. Soc.*, **84**, 1547-1564.

Randall, D. A., et al., 2007: Climate Models and Their Evaluation. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, et al., Eds., Cambridge University Press,, 589-662.

Randel, W., and F. Wu, 2007: A stratospheric ozone profile data set for 1979-2005: Variability, trends, and comparisons with column ozone data. *Journal of Geophysical Research-Atmospheres*. doi:ARTN D06313, 10.1029/2006JD007339, -.

Randel, W. J., et al., 2009: An update of observed stratospheric temperature trends. *Journal of Geophysical Research-Atmospheres*, **114**, 21.

Randerson, J. T., et al., 2009: Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. *Global Change Biology*, **15**, 2462-2484.

Rapaić, M., M. Leduc, and R. Laprise, 2010: Evaluation of the internal variability and estimation of the downscaling ability of the Canadian Regional Climate Model for different domain sizes over the north Atlantic region using the Big-Brother experimental approach. *Climate Dynamics*. doi:10.1007/s00382-010-0845-8, 1-23.

Raphael, M. N., and M. M. Holland, 2006: Twentieth century simulation of the southern hemisphere climate in coupled models. Part 1: large scale circulation variability. *Climate Dynamics*, **26**, 217-228.

Rashid, H. A., A. C. Hirst, and M. Dix, 2012: Atmospheric circulations in the ACCESS model simulations for CMIP5: Present-- day simulations and future projections. *Aust. Met. Oceanog. J.*, **submitted**.

Rauscher, S. A., E. Coppola, C. Piani, and F. Giorgi, 2010: Resolution effects on regional climate model simulations of seasonal precipitation over Europe. *Climate Dynamics*, **35**, 685-711.

Rauser, F., P. Korn, and J. Marotzke, 2011: Predicting goal error evolution from near-initial information: a learning algorithm. *Journal of Computational Physics*, **230**, 7284-7299.

Rayner, N. A., et al., 2003a: Global analysis of sea surface temperature, sea ice, and night marine air temperature since the late ninteeth century. *Journal of Geophysical Research*, **108**, doi:10,1029/2002JD002670.

Rayner, N. A., et al., 2003b: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, **108**.

Redelsperger, J.-L., C. D. Thorncroft, A. Diedhiou, T. Lebel, D. J. Parker, and J. Polcher, 2006: African Monsoon Multidisciplinary Analysis: An International Research Project and Field Campaign. *Bull. Amer. Meteorol. Soc.*, **87**, 1739-1746.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *Journal of Physical Oceanography*, **12**, 1154-1158.

Reichler, T., and J. Kim, 2008a: How well do coupled models simulate today's climate? *Bull. Amer. Meteorol. Soc.*, **89**, 303-311.

Reichler, T., and J. Kim, 2008b: How well do coupled models simulate today's climate? *Bulletin of the American Meteorological Society*. doi:DOI 10.1175/BAMS-89-3-303, 303-311.

Reick, C. H., T. Raddatz, V. Brovkin, and V. Gayler, 2012: The representation of natural and anthropogenic land cover change in MPI-ESM *JOURNAL OF ADVANCES IN MODELING EARTH SYSTEMS,*, **submitted**.

Richter, I., and S.-P. Xie, 2008: On the origin of equatorial Atlantic biases in coupled general circulation models. *Climate Dynamics*, **31**, 587-598.

Ridgwell, A., and J. C. Hargreaves, 2007: Regulation of atmospheric CO(2) by deep-sea sediments in an Earth system model. *Global Biogeochemical Cycles*, **21**.

Ridgwell, A., I. Zondervan, J. C. Hargreaves, J. Bijma, and T. M. Lenton, 2007a: Assessing the potential long-term increase of oceanic fossil fuel CO2 uptake due to CO2-calcification feedback. *Biogeosciences*, **4**, 481-492.

Ridgwell, A., et al., 2007b: Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling. *Biogeosciences*, **4**, 87-104.

Ringer, M. A., J. M. Edwards, and A. Slingo, 2003: Simulation of satellite channel radiances in the Met Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, **129**, 1169-1190.

Risien, C. M., and D. B. Chelton, 2008: A Global Climatology of Surface Wind and Wind Stress Fields from Eight Years of QuikSCAT Scatterometer Data. *Journal of Physical Oceanography*, **38**, 2379-2413.

Ritz, S. P., T. F. Stocker, and F. Joos, 2011: A Coupled Dynamical Ocean-Energy Balance Atmosphere Model for Paleoclimate Studies. *Journal of Climate*, **24**, 349-375.

Roberts, M. J., et al., 2004: Impact of an eddy-permitting ocean resolution on control and climate change simulations with a global coupled GCM. *Journal of Climate*, **17**, 3-20.

Roberts, M. J., et al., 2009: Impact of Resolution on the Tropical Pacific Circulation in a Matrix of Coupled Models. *Journal of Climate*, **22**, 2541-2556.

Robinson, D. A., and A. Frei, 2000a: Seasonal variability of northern hemisphere snow extent using visible satellite data. *Professional Geographer*, **51**, 307–314.

Robinson, D. A., and A. Frei, 2000b: Seasonal variability of Northern Hemisphere snow extent using visible satellite data. *Professional Geographer*, **52**.

Rodwell, M., and T. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quarterly Journal of the Royal Meteorological Society*, **133**, 129-146.

Roe, G. H., and M. B. Baker, 2007: Why is climate sensitivity so unpredictable? *Science*, **318**, 629-632.

Roe, G. H., and K. C. Armour, 2011: How sensitive is climate sensitivity? *Geophysical Research Letters*, **38**.

Roe, G. H., and M. B. Baker, 2011: Comment on "Another look at climate sensitivity" by Zaliapin and Ghil (2010). *Nonlinear Processes in Geophysics*, **18**, 125-127.

Roeckner, E., et al., 2006: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *Journal of Climate*, **19**, 3771-3791.

Rogers, J. C., 1981: The North Pacific Oscillation. *Journal of Climatology*, **1**, 19.

Rojas, M., 2006: Multiply Nested Regional Climate Simulation for Southern South America: Sensitivity to Model Resolution. *Monthly Weather Review*, **134**, 2208-2223.

Rojas, M., and P. I. Moreno, 2011: Atmospheric circulation changes and neoglacial conditions in the Southern Hemisphere mid-latitudes: insights from PMIP2 simulations at 6 kyr. *Climate Dynamics*, **37**, 357-375.

Rojas, M., et al., 2009: The Southern Westerlies during the last glacial maximum in PMIP2 simulations. *Climate Dynamics*, **32**, 525-548.

Romanoua, A., et al., 2012: Preindustrial Control Simulations of the Ocean-Atmosphere CO2 Exchanges in the NASA-GISSclimate model: sensitivity to physical ocean model. *OM*, **submitted**.

Rotstayn, L. D., and U. Lohmann, 2002: Simulation of the tropospheric sulfur cycle in a global model with a physically based cloud scheme. *J. Geophys. Res*, **107**.

Rotstayn, L. D., M. A. Collier, R. M. Mitchell, Y. Qin, S. K. Campbell, and S. M. Dravitzki, 2011: Simulated enhancement of ENSO-related rainfall variability due to Australian dust. *Atmospheric Chemistry and Physics*, **11**.

Rotstayn, L. D., et al., 2010: Improved simulation of Australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment. *International Journal of Climatology*, **30**, 1067-1088.

Rotstayn, L. D. J., S.J.Collier, M.A.Dravitzki, S.M.Hirst, A.C.Syktus, J.I.Wong, K.K., 2012: Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: a study using single-forcing climate simulations. *Atmos. Chem. Phys.*, **12**, 6377-6404.

Rougier, J., 2007: Probabilistic inference for future climate using an ensemble of climate model evaluations. *Climatic Change*, **81**.

Rougier, J., D. M. H. Sexton, J. M. Murphy, and D. Stainforth, 2009a: Analyzing the Climate Sensitivity of the HadSM3 Climate Model Using Ensembles from Different but Related Experiments. *Journal of Climate*, **22**.

Rougier, J., D. M. H. Sexton, J. M. Murphy, and D. Stainforth, 2009b: Analyzing the Climate Sensitivity of the HadSM3 Climate Model Using Ensembles from Different but Related Experiments. *Journal of Climate*, **22**, 3540-3557.

Roy, P., P. Gachon, and R. Laprise, 2011a: Assessment of summer extremes and climate variability over the north-east of North America as simulated by the Canadian Regional Climate Model. *International Journal of Climatology*. doi:10.1002/joc.2382, n/a-n/a.

Roy, T., et al., 2011b: Regional impacts of climate change and atmospheric CO2 on future ocean carbon uptake: A multi-model linear feedback analysis. *Journal of Climate*, **24**, 2300-2318.

Ruckstuhl, C., and J. R. Norris, 2009: How do aerosol histories affect solar "dimming" and "brightening" over Europe?: IPCC-AR4 models versus observations. *Journal of Geophysical Research-Atmospheres*, **114**.

Rummukainen, M., 2010: State-of-the-art with regional climate models. *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 82-96.

Russell, J. L., R. J. Stouffer, and K. W. Dixon, 2006: Intercomparison of the Southern Ocean circulations in IPCC coupled model control simulations. *Journal of Climate*, **19**, 4560-4575.

Ruti, P. M., et al., 2011: The West African climate system: a review of the AMMA model inter-comparison initiatives. *Atmospheric Science Letters*. doi:10.1002/asl.305, n/a-n/a.

Rutt, I. C., M. Hagdorn, N. R. J. Hulton, and A. J. Payne, 2009: The Glimmer community ice sheet model. *Journal of Geophysical Research-Earth Surface*, **114**.

Rutter, N., et al., 2009: Evaluation of forest snow processes models (SnowMIP2). *Journal of Geophysical Research-Atmospheres*, **114**.

Saha, S., et al., 2010: THE NCEP CLIMATE FORECAST SYSTEM REANALYSIS. *Bull. Amer. Meteorol. Soc.*, **91**.

Sahany, S., J. D. Neelin, K. Hales, and R. B. Neale, 2012: Temperature–Moisture Dependence of the Deep Convective Transition as a Constraint on Entrainment in Climate Models. *Journal of the Atmospheric Sciences*, **69**, 1340-1358.

Saji, N. H., S. P. Xie, and T. Yamagata, 2006: Tropical Indian Ocean variability in the IPCC twentieth-century climate simulations. *Journal of Climate*, **19**, 4397-4417.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401**, 360-363.

Sakamoto, T. T., et al., 2011: MIROC4h – a new high-resolution atmosphere-ocean coupled general circulation model. *Journal of Meteorological Society of Japan*. submitted.

——, 2012: MIROC4h – a new high-resolution atmosphere-ocean coupled general circulation model. *Journal of Meteorological Society of Japan*, **90**, 325-359.

Samuelsson, P., E. Kourzeneva, and D. Mironov, 2010: The impact of lakes on the European climate as simulated by a regional climate model. *Boreal Environment Research*, **15**, 113-129.

Sanchez-Gomez, E., S. Somot, and M. Deque, 2009: Ability of an ensemble of regional climate models to reproduce weather regimes over Europe-Atlantic during the period 1961-2000. *Climate Dynamics*, **33**, 723-736.

Sander, S. P., 2006: Chemical kinetics and photochemical data for use in atmospheric studies Evaluation 15, 523 pp. pp.

Sanderson, B., C. Piani, W. Ingram, D. Stone, and M. Allen, 2008a: Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics GCM simulations. *Climate Dynamics*, **30**, 175-190.

Sanderson, B. M., 2011a: A Multimodel Study of Parametric Uncertainty in Predictions of Climate Response to Rising Greenhouse Gas Concentrations. *Journal of Climate*, **25**, 1362–1377.

——, 2011b: On the estimation of systematic error in regression-based predictions of climate sensitivity. *Climatic Change*. submitted.

——, 2012: On the estimation of systematic error in regression-based predictions of Climate Sensitivity. *Climatic Change*. in press.

Sanderson, B. M., and R. Knutti, 2012a: Combining projections from an imperfect ensemble of climate models. *PNAS*, **submitted**.

Sanderson, B. M., and R. Knutti, 2012b: On the interpretation of constrained climate model ensembles. *Geophysical Research Letters*. in press.

Sanderson, B. M., K. M. Shell, and W. Ingram, 2010: Climate feedbacks determined using radiative kernels in a multi-thousand member ensemble of AOGCMs. *Climate Dynamics*, **35**, 1219-1236.

Sanderson, B. M., C. Piani, W. J. Ingram, D. A. Stone, and M. R. Allen, 2008b: Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics GCM simulations. *Climate Dynamics*, **30**, 175-190.

Sanderson, B. M., et al., 2008c: Constraints on model response to greenhouse gas forcing and the role of subgrid-scale processes. *Journal of Climate*, **21**, 2384-2400, doi:2310.1175/2008jcli1869.2381.

Santer, and Co-authors, 2012: Identifying human influences on atmospheric temperature: are results robust to Uncertainties. *PNAS*. submitted.

Santer, B., et al., 2009a: Incorporating model quality information in climate change detection and attribution studies. *Proc. Natl. Acad. Sci. U. S. A.* doi:DOI 10.1073/pnas.0901736106, 14778-14783.

Santer, B., et al., 2007: Identification of human-induced changes in atmospheric moisture content. *Proc. Natl. Acad. Sci. U. S. A.* doi:DOI 10.1073/pnas.0702872104, 15248-15253.

Santer, B., et al., 2008: Consistency of modelled and observed temperature trends in the tropical troposphere. *International Journal of Climatology*. doi:DOI 10.1002/joc.1756, 1703-1722.

Santer, B., et al., 2005: Amplification of surface temperature trends and variability in the tropical atmosphere. *Science*, **309**, 1551-1556.

Santer, B. D., et al., 2009b: Incorporating model quality information in climate change detection and attribution studies. *Proceedings of the National Academy of Sciences of the United States of America*, **106**.

Santer, B. D., et al., 2012: Identifying human influences on atmospheric temperature: are results robust to uncertainties? *Proceedings of the National Academy of Sciences (USA)*. submitted.

Sasaki, H., and K. Kurihara, 2008: Relationship between Precipitation and Elevation in the Present Climate Reproduced by the Non-hydrostatic Regional Climate Model. *Sola*, **4**, 109-112.

Sato, K., and T. J. Dunkerton, 1997: Estimates of momentum flux associated with equatorial Kelvin and gravity waves. *Journal of Geophysical Research-Atmospheres*, **102**, 26247-26261.

Sato, M., J. E. Hansen, M. P. McCormick, and J. B. Pollack, 1993: Stratospheric aerosol optical depth, 1850– 1990. *J. Geophys. Res.*, **98**, 22,987–922,994

Sato, T., H. Miura, M. Satoh, Y. N. Takayabu, and Y. Q. Wang, 2009: Diurnal Cycle of Precipitation in the Tropics Simulated in a Global Cloud-Resolving Model. *Journal of Climate*, **22**, 4809-4826.

Scaife, A. A., N. Butchart, C. D. Warner, and R. Swinbank, 2002: Impact of a spectral gravity wave parameterization on the stratosphere in the met office unified model. *Journal of the Atmospheric Sciences*, **59**, 1473-1489.

Scaife, A. A., J. R. Knight, G. K. Vallis, and C. K. Folland, 2005: A stratospheric influence on the winter NAO and North Atlantic surface climate. *Geophysical Research Letters*, **32**.

Scaife, A. A., T. Woollings, J. Knight, G. Martin, and T. Hinton, 2010: Atmospheric Blocking and Mean Biases in Climate Models. *Journal of Climate*, **23**, 6143-6152.

Scaife, A. A., N. Butchart, C. D. Warner, D. Stainforth, W. Norton, and J. Austin, 2000: Realistic Quasi-Biennial Oscillations in a simulation of the global climate. *Geophysical Research Letters*, **27**, 3481-3484.

Scaife, A. A., et al., 2011a: Improved Atlantic winter blocking in a climate model. *Geophysical Research Letters*, **38**.

Scaife, A. A., et al., 2011b: Climate Change and Stratosphere-Troposphere Interaction. *Climate Dynamics*. doi:10.1007/s00382-011-1080-7.

Scaife, A. A., et al., 2009: The CLIVAR C20C project: selected twentieth century climate events. *Climate Dynamics*, **33**, 603-614.

Schaller, N., I. Mahlstein, J. Cermak, and R. Knutti, 2011: Analyzing precipitation projections: A comparison of different approaches to climate model evaluation. *Journal of Geophysical Research-Atmospheres*, **116**.

Scherrer, S., 2011: Present-day interannual variability of surface climate in CMIP3 models and its relation to future warming. *International Journal of Climatology*, **31**, 1518-1529.

Schlesinger, M. E., and J. F. B. Mitchell, 1987: Climate model simulations of the equilibrium climatic response to increased carbon-dioxide. *Reviews of Geophysics*, **25**, 760-798.

Schmidli, J., C. Goodess, C. Frei, M. Haylock, Y. Hundecha, J. Ribalaygua, and T. Schmith, 2007: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. *Journal of Geophysical Research-Atmospheres*, **112**, -.

Schmidt, G. A., et al., 2012: Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1). *Geoscientific Model Development*, **5**, 185-191.

Schmidt, G. A., et al., 2006: Present day atmospheric simulations using GISS ModelE: Comparison to in-situ, satellite and reanalysis data. *J. Climate*, **19**, 153-192.

Schmith, T., 2008: Stationarity of regression relationships: Application to empirical downscaling. *Journal of Climate*, **21**, 4529-4537.

Schmittner, A., A. Oschlies, X. Giraud, M. Eby, and H. L. Simmons, 2005: A global model of the marine ecosystem for long-term simulations: Sensitivity to ocean mixing, buoyancy forcing, particle sinking, and dissolved organic matter cycling. *Global Biogeochemical Cycles*, **19**.

Schott, F. A., S.-P. Xie, and J. P. McCreary, Jr., 2009: Indian Ocean circulation and climate variability. *Reviews of Geophysics*, **47**, -.

Schramm, J. L., M. M. Holland, J. A. Curry, and E. E. Ebert, 1997: Modeling the thermodynamics of a sea ice thickness 1. Sensitivity to ice thickness resolution. *Journal of Geophysical Research*, **102**, 23,079-023,091.

Schultz, M. G., et al., 2008: Global wildland fire emissions from 1960 to 2000. *Global Biogeochemical Cycles*, **22**.

Schurgers, G., U. Mikolajewicz, M. Groger, E. Maier-Reimer, M. Vizcaino, and A. Winguth, 2008: Long-term effects of biogeophysical and biogeochemical interactions between terrestrial biosphere and climate under anthropogenic climate change. *Global and Planetary Change*, **64**, 26-37.

Scinocca, J. F., and N. A. McFarlane, 2000: The parametrization of drag induced by stratified flow over anisotropic orography. *Quarterly Journal of the Royal Meteorological Society*, **126**, 2353-2393.

Scoccimarro, E., et al., 2011: Effects of Tropical Cyclones on Ocean Heat Transport in a High Resolution Coupled General Circulation Model. *J. Clim.*, **24**, 4368-4384.

Seager, R., and R. Murtugudde, 1997: Ocean dynamics, thermocline adjustment, and regulation of tropical SST. *Journal of Climate*, **10**, 521-534.

Segui, P., A. Ribes, E. Martin, F. Habets, and J. Boe, 2010: Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins. *Journal of Hydrology*, **383**, 111-124.

Seidel, D. J., Q. Fu, W. J. Randel, and T. J. Reichler, 2008: Widening of the tropical belt in a changing climate. *Nature Geoscience*, **1**, 21-24.

Seidel, D. J., N. P. Gillett, J. R. Lanzante, K. P. Shine, and P. W. Thorne, 2011: Stratospheric temperature trends: Our evolving understanding. *Wiley Interdisciplinary Reviews: Climate Change*, **2**, 592-616.

Sellers, P. J., et al., 1996: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *Journal of Climate*, **9**, 676-705.

Selten, F. M., G. W. Branstator, H. A. Dijkstra, and M. Kliphuis, 2004: Tropical origins for recent and future Northern Hemisphere climate change. *Geophys. Res. Lett.*, **31**, L21205-.

Semenov, V. A., M. Latif, J. H. Jungclaus, and W. Park, 2008: Is the observed NAO variability during the instrumental record unusual? *Geophys. Res. Lett.*, **35**, L11701-.

Sen Gupta, A., N. C. Jourdain, J. N. Brown, and Monselesan, 2012: Climate Drift in the CMIP5 models. *Journal of Climate*. submitted.

Seneviratne, S., et al., 2012: Changes in Climate Extremes and their Impacts on the Natural Physical Environment. *IPCC WGI/WGII Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX)*, in press.

Seneviratne, S. I., D. Luethi, M. Litschi, and C. Schaer, 2006: Land-atmosphere coupling and climate change in Europe. *Nature*, **443**.

Seneviratne, S. I., et al., 2010: Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, **99**, 125-161.

Separovic, L., R. De Elia, and R. Laprise, 2008: Reproducible and Irreproducible Components in Ensemble Simulations with a Regional Climate Model. *Monthly Weather Review*. doi:DOI 10.1175/2008MWR2393.1, 4942-4961.

Separovic, L., R. Elía, and R. Laprise, 2012: Impact of spectral nudging and domain size in studies of RCM response to parameter modification. *Climate Dynamics*, **38**, 1325-1343.

Sexton, D. M. H., and J. M. Murphy, 2012: Multivariate probabilistic projections using imperfect climate models. Part II: robustness of methodological choices and consequences for climate sensitivity. *Climate Dynamics*, **38**, 2543-2558.

Sexton, D. M. H., J. M. Murphy, M. Collins, and M. J. Webb, 2012: Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. *Climate Dynamics*, **38**, 2513-2542.

Shaffer, G., and J. L. Sarmiento, 1995: BIOGEOCHEMICAL CYCLING IN THE GLOBAL OCEAN .1. A NEW, ANALYTICAL MODEL WITH CONTINUOUS VERTICAL RESOLUTION AND HIGH-LATITUDE DYNAMICS. *Journal of Geophysical Research-Oceans*, **100**, 2659-2672.

Shaffer, G., S. M. Olsen, and J. O. P. Pedersen, 2008: Presentation, calibration and validation of the low-order, DCESS Earth System Model (Version 1). *Geoscientific Model Development*, **1**, 17-51.

Shaffrey, L. C., et al., 2009: UK HiGEM: The New UK High-Resolution Global Environment Model-Model Description and Basic Evaluation. *Journal of Climate*, **22**, 1861-1896.

Sheffield, J., and E. F. Wood, 2008: Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Climate Dynamics*, **31**, 79-105.

Sheffield, J., and e. al., 2012: North American Climate in CMIP5 Experiments: Part II: Evaluation of 20th Century Intra-Seasonal to Decadal Variability. submitted ed., J. Climate.

Shell, K. M., J. T. Kiehl, and C. A. Shields, 2008: Using the radiative kernel technique to calculate climate feedbacks in NCAR's Community Atmospheric Model. *Journal of Climate*, **21**, 2269-2282.

Shevliakova, E., et al., 2009: Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink. *Global Biogeochemical Cycles*, **23**.

Shi, Y., J. Zhang, J. S. Reid, B. Holben, E. J. Hyer, and C. Curtis, 2011: An analysis of the collection 5 MODIS over-ocean aerosol optical depth product for its implication in aerosol assimilation. *Atmospheric Chemistry and Physics*, **11**, 557-565.

Shibata, K., and M. Deushi, 2005: Radiative effect of ozone on the quasi-biennial oscillation in the equatorial stratosphere. *Geophysical Research Letters*, **32**.

Shin, D., J. Kim, and H. Park, 2011: Agreement between monthly precipitation estimates from TRMM satellite, NCEP reanalysis, and merged gauge-satellite analysis. *Journal of Geophysical Research-Atmospheres*, **116**.

Shin, S. I., D. Sardeshmukh, and K. Pegion, 2010: Realism of local and remote feedbacks on tropical sea surface temperatures in climate models. *Journal of Geophysical Research-Atmospheres*, **115**.

Shindell, D. T., et al., 2012a: Interactive ozone and methane chemistry in GISS-E2 historical and future climate simulations *ACP*, **submitted**.

Shindell, D. T., et al., 2012b: Radiative forcing in the ACCMIP historical and future climate simulations. *Atmos. Chem. Phys.*

Shiogama, H., S. Emori, N. Hanasaki, M. Abe, Y. Masutomi, K. Takahashi, and T. Nozawa, 2011: Observational constraints indicate risk of drying in the Amazon basin. *Nature Communications*, **2**.

Shiogama, H., et al., 2012: Perturbed physics ensemble using the MIROC5 coupled atmosphere–ocean GCM without flux corrections: experimental design and results. *Climate Dynamics*. doi:DOI 10.1007/s00382-012-1441-x, 1-16.

Shkol'nik, I., V. Meleshko, S. Efimov, and E. Stafeeva, 2012: Changes in climate extremes on the territory of Siberia by the middle of the 21st century: An ensemble forecast based on the MGO regional climate model. *Russian Meteorology and Hydrology*, **37**, 71-84.

Shkolnik, I., V. Meleshko, and V. Kattsov, 2007: The MGO climate model for Siberia. *Russian Meteorology and Hydrology*, **32**, 351-359.

Shukla, J., T. DelSole, M. Fennessy, J. Kinter, and D. Paolino, 2006: Climate model fidelity and projections of climate change. *Geophysical Research Letters*, **33**, L07702.

Siebesma, A. P., P. M. M. Soares, and J. Teixeira, 2007: A Combined Eddy-Diffusivity Mass-Flux Approach for the Convective Boundary Layer. *Journal of the Atmospheric Sciences*, **64**, 1230-1248.

Sillmann, J., and M. Croci-Maspoli, 2009: Present and future atmospheric blocking and its impact on European mean and extreme climate. *Geophysical Research Letters*, **36**.

Sillmann, J., M. Croci-Maspoli, M. Kallache, and R. W. Katz, 2011: Extreme Cold Winter Temperatures in Europe under the Influence of North Atlantic Atmospheric Blocking. *Journal of Climate*, **24**, 5899-5913.

Sillmann, J., V. V. Kharin, X. Zhang, and F. W. Zwiers, 2012: Climate extreme indices in the CMIP5 multi-model ensemble. Part 1: Model evaluation in the present climate. *Journal of Geophysical Research*. submitted.

Simmons, H. L., S. R. Jayne, L. C. St Laurent, and A. J. Weaver, 2004: Tidally driven mixing in a numerical model of the ocean general circulation. *Ocean Modelling*, **6**, 245-263.

Simpson, I. R., M. Blackburn, and J. D. Haigh, 2009: The Role of Eddies in Driving the Tropospheric Response to Stratospheric Heating Perturbations. *Journal of the Atmospheric Sciences*, **66**, 1347-1365.

Sitch, S., et al., 2003: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, **9**, 161-185.

Sitch, S., et al., 2008: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology*, **14**, 2015-2039.

Six, K. D., and E. Maier-Reimer, 1996: Effects of Plankton Dynamics on Seasonal Carbon Fluxes in an Ocean General Circulation Model. *Global Biogeochemical Cycles*, **10**, 559-583.

Slater, A. G., and D. M. Lawrence, 2012: Diagnosing present and future permafrost from climate models. *Journal of Climate*.

Sloyan, B. M., and I. V. Kamenkovich, 2007: Simulation of Subantarctic Mode and Antarctic Intermediate Waters in climate models. *Journal of Climate*, **20**, 5061-5080.

Smirnov, D., and D. J. Vimont, 2011: Variability of the Atlantic Meridional Mode during the Atlantic Hurricane Season. *Journal of Climate*, **24**, 1409-1424.

Smith, B., P. Samuelsson, A. Wramneby, and M. Rummukainen, 2011a: A model of the coupled dynamics of climate, vegetation and terrestrial ecosystem biogeochemistry for regional applications. *Tellus Series a-Dynamic Meteorology and Oceanography*. doi:DOI 10.1111/j.1600-0870.2010.00477.x, 87-106.

Smith, P. C., P. Ciais, P. Peylin, N. De Noblet-Ducoudre, N. Viovy, Y. Meurdesoif, and A. Bondeau, 2010a: European-wide simulations of croplands using an improved terrestrial biosphere model: 2. Interannual yields and anomalous CO2 fluxes in 2003. *Journal of Geophysical Research-Biogeosciences*, **115**.

Smith, P. C., N. De Noblet-Ducoudre, P. Ciais, P. Peylin, N. Viovy, Y. Meurdesoif, and A. Bondeau, 2010b: European-wide simulations of croplands using an improved terrestrial biosphere model: Phenology and productivity. *Journal of Geophysical Research-Biogeosciences*, **115**.

Smith, R. D., M. E. Maltrud, F. O. Bryan, and M. W. Hecht, 2000: Numerical simulation of the North Atlantic Ocean at 1/10 degrees. *Journal of Physical Oceanography*, **30**, 1532-1561.

Smith, R. S., J. M. Gregory, and A. Osprey, 2008a: A description of the FAMOUS (version XDBUA) climate model and control run. *Geoscientific Model Development*, **1**, 53-68.

Smith, S. J., J. van Aardenne, Z. Klimont, R. J. Andres, A. Volke, and S. D. Arias, 2011b: Anthropogenic sulfur dioxide emissions: 1850-2005. *Atmospheric Chemistry and Physics*, **11**, 1101-1116.

Smith, S. J., J. van Aardenne, Z. Klimont, R. J. Andres, A. Volke, and S. Delgado Arias, 2011c: Anthropogenic sulfur dioxide emissions: 1850–2005. *Atmos. Chem. Phys.*, **11**, 1101-1116.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008b: Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006). *Journal of Climate*, **21**, 2283-2296.

Sobolowski, S., and T. Pavelsky, 2012: Evaluation of present and future North American Regional Climate Change Assessment Program (NARCCAP) regional climate simulations over the southeast United States. *Journal of Geophysical Research-Atmospheres*, **117**.

Soden, B. J., and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean-atmosphere models. *Journal of Climate*, **19**, 3354-3360.

Soden, B. J., I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields, 2008: Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, **21**, 3504-3520.

Sokolov, A. P., and P. H. Stone, 1998: A flexible climate model for use in integrated assessments. *Clim. Dyn.*, **14**, -303.

Sokolov, A. P., C. E. Forest, and P. H. Stone, 2010: Sensitivity of climate change projections to uncertainties in the estimates of observed changes in deep-ocean heat content. *Climate Dynamics*, **34**, 735-745.

Sokolov, A. P., et al., 2009: Probabilistic Forecast for Twenty-First-Century Climate Based on Uncertainties in Emissions (Without Policy) and Climate Parameters. *Journal of Climate*, **22**, 5175-5204.

Sokolov, A. P., et al., 2005: The MIT Integrated Global System Model (IGSM) Version 2: Model Description and Baseline Evaluation, MIT JP Report 124. MIT.

Solman, S., and N. Pessacg, 2012: Regional climate simulations over South America: sensitivity to model physics and to the treatment of lateral boundary conditions using the MM5 model. *Climate Dynamics*, **38**, 281-300.

Solomon, A., and D. Zhang, 2006: Pacific subtropical cell variability in coupled climate model simulations of the late 19th–20th century. *Ocean Modelling*, **15**, 236–249.

Solomon, S., P. J. Young, and B. Hassler, 2012a: Uncertainties in the evolution of stratospheric ozone and implications for recent temperature changes in the tropical lower stratosphere. *Geophys. Res. Lett.*, accepted.

Solomon, S., P. J. Young, and B. Hassler, 2012b: Uncertainties in the evolution of stratospheric ozone and implications for recent temperature changes in the tropical lower stratosphere *GRL*, **accepted**.

Somot, S., F. Sevault, M. Deque, and M. Crepon, 2008: 21st century climate change scenario for the Mediterranean using a coupled atmosphere-ocean regional climate model. *Global and Planetary Change*. doi:DOI 10.1016/j.gloplacha.2007.10.003, 112-126.

Son, S., et al., 2008: The impact of stratospheric ozone recovery on the Southern Hemisphere westerly jet. *Science*. doi:DOI 10.1126/science.1155939, 1486-1489.

Son, S., et al., 2010: Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel assessment. *Journal of Geophysical Research-Atmospheres*. doi:ARTN D00M07, 10.1029/2010JD014271.

Song, Z., F. Qiao, and Y. Song, 2012: Response of the equatorial basin-wide SST to non-breakingsurface wave-induced mixing in a climate model:An amendment to tropical bias. *JOURNAL OF GEOPHYSICAL RESEARCH,*, **117**.

SPARC-CCMVal, 2010: SPARC Report on the Evaluation of Chemistry-Climate Models.

Sperber, K., and D. Kim, 2012: Simplified metrics for the identification of the Madden-Julian oscillation in models. *Atmos. Sci. Let.* doi:10.1002/asl.378.

Sperber, K., et al., submitted: The Asian Summer Monsoon: An Intercomparison of CMIP-5 vs. CMIP-3 Simulations of the Late 20th Century. *Journal of Climate*.

Sperber, K. R., 2003: Propagation and the vertical structure of the Madden-Julian oscillation. *Monthly Weather Review*, **131**, 3018-3037.

Sperber, K. R., and H. Annamalai, 2008: Coupled model simulations of boreal summer intraseasonal (30-50 day) variability, Part 1: Systematic errors and caution on use of metrics. *Climate Dynamics*, **31**, 345-372.

Sperber, K. R., et al., 2010: Monsoon Fact Sheet: CLIVAR Asian-Australian Monsoon Panel.

Stainforth, D., et al., 2005a: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403-406.

Stainforth, D. A., et al., 2005b: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403-406.

Stephens, G. L., and C. D. Kummerow, 2007: The remote sensing of clouds and precipitation from space: A review. *Journal of the Atmospheric Sciences*, **64**, 3742-3765.

Stephens, G. L., et al., 2010: Dreary state of precipitation in global models. *J. Geophys. Res.*, **115**, D24211.

Stephens, G. L., et al., 2012a: An Update on the Earth's energy balance in light of new global observations. *Nature Geoscience*, **submitted**.

——, 2012b: An Update on the Earth's energy balance in light of new global observations *Nature Geoscience*, **submitted**.

Stephenson, D. B., M. Collins, J. C. Rougier, and R. E. Chandler, 2012: Statistical Problems in the Probabilistic Prediction of Climate Change, **in press**.

Stephenson, D. B., V. Pavan, M. Collins, M. M. Junge, R. Quadrelli, and C. M. G. Participating, 2006: North Atlantic Oscillation response to transient greenhouse gas forcing and the impact on European winter climate: a CMIP2 multi-model assessment. *Climate Dynamics*, **27**, 401-420.

Stevens, B., et al., 2012: The Atmospheric Component of the MPI-M Earth System Model: ECHAM6 *JOURNAL OF ADVANCES IN MODELING EARTH SYSTEMS,*, **submitted**.

Stevenson, S., 2012: Significant Changes to ENSO Strength and Impacts in the Twenty-First Century: Results from CMIP5. Submitted ed., GRL.

Stevenson, S., Fox-Kemper, B., Jochum, M., Neale, R., Deser, C., & Meehl, G. (2012). Will there be a significant change to El Nino in the 21st century? . *J. Clim., 25*, 2129–2145. doi: 10.1175/JCLI-D-11-00252.1

Stocker, B. D., K. Strassmann, and F. Joos, 2011: Sensitivity of Holocene atmospheric CO2 and the modern carbon budget to early human land use: analyses with a process-based model. *Biogeosciences*, **8**, 69-88.

Stolarski, R., and S. Frith, 2006: Search for evidence of trend slow-down in the long-term TOMS/SBUV total ozone data record: the importance of instrument drift uncertainty. *Atmospheric Chemistry and Physics*, **6**, 4057-4065.

Stoner, A. M. K., K. Hayhoe, and D. J. Wuebbles, 2009: Assessing General Circulation Model Simulations of Atmospheric Teleconnection Patterns. *Journal of Climate*, **22**, 4348-4372.

Stott, P. A., and C. E. Forest, 2007: Ensemble climate predictions using climate models and observational constraints. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*, **365**, 2029-2052.

Strassmann, K. M., F. Joos, and G. Fischer, 2008: Simulating effects of land use changes on carbon fluxes: past contributions to atmospheric CO2 increases and future commitments due to losses of terrestrial sink capacity. *Tellus Series B-Chemical and Physical Meteorology*, **60**, 583-603.

Stratton, R. A., and A. J. Stirling, 2012: Improving the diurnal cycle of convection in GCMs. *Quarterly Journal of the Royal Meteorological Society*, **138**, 1121-1134.

Stroeve, J., M. Holland, W. Meier, T. Scambos, and M. Serreze, 2007: Arctic sea ice decline: Faster than forecast. *Geophysical Research Letters*, **34**, -.

Stroeve, J. C., V. Kattsov, A. Barrett, M. Serreze, T. Pavlova, M. Holland, and W. N. Meier, 2012: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophysical Research Letters*.

Su, H., D. E. Waliser, J. H. Jiang, J. L. Li, W. G. Read, J. W. Waters, and A. M. Tompkins, 2006: Relationships of upper tropospheric water vapor, clouds and SST: MLS observations, ECMWF analyses and GCM simulations. *Geophysical Research Letters*, **33**.

Su, H., et al., 2011: Comparison of regime-sorted tropical cloud profiles observed by CloudSat with GEOS5 analyses and two general circulation model simulations. *J. Geophys. Res.*, **116**.

Suh, M., S. Oh, D. Lee, D. Cha, S. Choi, C. Jin, and S. Hong, 2012: Development of new ensemble methods based on the performance skills of regional climate models over South Korea. *Journal of Climate*, **In press**.

Sun, D.-Z., Y. Yu, and T. Zhang, 2009: Tropical Water Vapor and Cloud Feedbacks in Climate Models: A Further Assessment Using Coupled Simulations. *Journal of Climate*, **22**, 1287-1304.

Sun, Z., et al., 2012: Modifications in atmospheric physical parameterization aimed at improving SST simulation in the ACCESS coupled-model *Australian Meteorological and Oceanographic Journal,*, **submitted**.

Sutton, R. T., B. W. Dong, and J. M. Gregory, 2007: Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations. *Geophysical Research Letters*, **34**.

Svensson, G., and A. Holtslag, 2009: Analysis of Model Results for the Turning of the Wind and Related Momentum Fluxes in the Stable Boundary Layer. *Boundary-Layer Meteorology*, **132**, 261-277-277.

Svensson, G., et al., 2011: Evaluation of the Diurnal Cycle in the Atmospheric Boundary Layer Over Land as Represented by a Variety of Single-Column Models: The Second GABLS Experiment. *Boundary-Layer Meteorology*, **140**, 177-206.

Swart, N. C., and J. C. Fyfe, 2011: Ocean carbon uptake and storage influenced by wind bias in global climate models. *Nature Climate Change*.

——, 2012: Observed and simulated changes in the Southern Hemisphere surface westerly wind-stress. *Geophys Res Lett*, **accepted**.

Tachiiri, K., J. C. Hargreaves, J. D. Annan, A. Oka, A. Abe-Ouchi, and M. Kawamiya, 2010: Development of a system emulating the global carbon cycle in Earth system models. *Geoscientific Model Development*, **3**, 365-376.

Tagliabue, A., et al., 2009: Quantifying the roles of ocean circulation and biogeochemistry in governing ocean carbon-13 and atmospheric carbon dioxide at the last glacial maximum. *Climate of the Past*, **5**, 695-706.

Takahashi, M., 1999: Simulation of the stratospheric quasi-biennial oscillation in a general circulation model. *Geophysical Research Letters*, **26**, 1307-1310.

Takahashi, T., et al., 2009: Climatological mean and decadal change in surface ocean pCO(2), and net sea-air CO2 flux over the global oceans. *Deep-Sea Research Part Ii-Topical Studies in Oceanography*, **56**, 554-577.

Takata, K., S. Emori, and T. Watanabe, 2003: Development of the minimal advanced treatments of surface interaction and runoff. *Global and Planetary Change*, **38**.

Takemura, T., T. Nakajima, O. Dubovik, B. N. Holben, and S. Kinne, 2002: Single-scattering albedo and radiative forcing of various aerosol species with a global three-dimensional model. *Journal of Climate*, **15**.

Takemura, T., T. Nozawa, S. Emori, T. Y. Nakajima, and T. Nakajima, 2005: Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model. *Journal of Geophysical Research-Atmospheres*, **110**.

Takemura, T., H. Okamoto, Y. Maruyama, A. Numaguti, A. Higurashi, and T. Nakajima, 2000: Global three-dimensional simulation of aerosol optical thickness distribution of various origins. *Journal of Geophysical Research-Atmospheres*, **105**.

Takemura, T., M. Egashira, K. Matsuzawa, H. Ichijo, R. O'Ishi, and A. Abe-Ouchi, 2009: A simulation of the global distribution and radiative forcing of soil dust aerosols at the Last Glacial Maximum. *Atmospheric Chemistry and Physics*, **9**.

Takle, E. S., et al., 2007: Transferability intercomparison - An opportunity for new insight on the global water cycle and energy budget. *Bull. Amer. Meteorol. Soc.*, **88**, 375-+.

Taylor, C. M., A. Gounou, F. Guichard, P. P. Harris, R. J. Ellis, F. Couvreux, and M. De Kauwe, 2011: Frequency of Sahelian storm initiation enhanced over mesoscale soil-moisture patterns. *Nature Geoscience*, **4**, 430-433.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: AN OVERVIEW OF CMIP5 AND THE EXPERIMENT DESIGN. *Bull. Amer. Meteorol. Soc.*, **93**, 485-498.

Tebaldi, C., and R. Knutti, 2007a: The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences*, **365**, 2053-2075.

Tebaldi, C., and R. Knutti, 2007b: The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*. doi:DOI 10.1098/rsta.2007.2076, 2053-2075.

Tebaldi, C., and B. Sanso, 2009: Joint projections of temperature and precipitation change from multiple climate models: a hierarchical Bayesian approach. *Journal of the Royal Statistical Society Series a-Statistics in Society*, **172**, 83-106.

Tebaldi, C., L. O. Mearns, D. Nychka, and R. L. Smith, 2004: Regional probabilities of precipitation change: A Bayesian analysis of multimodel simulations. *Geophysical Research Letters*, **31**.

Teixeira, J., et al., 2008: Parameterization of the atmospheric boundary layer. *Bull. Amer. Meteorol. Soc.*, **89**, 453-458.

Teixeira, J., et al., 2011: Tropical and Subtropical Cloud Transitions in Weather and Climate Prediction Models: The GCSS/WGNE Pacific Cross-Section Intercomparison (GPCI). *Journal of Climate*, **24**, 5223-5256.

Terray, L., L. Corre, S. Cravatte, T. Delcroix, G. Reverdin, and R. Aurelien, 2012: Near-surface salinity as Nature's rain gauge to detect human influence on the tropical water cycle. *Journal of Climate*, (submitted).

Teutschbein, C., F. Wetterhall, and J. Seibert, 2011: Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale. *Climate Dynamics*, **37**, 2087-2105.

Textor, C., et al., 2007: The effect of harmonized emissions on aerosol properties in global models - an AeroCom experiment. *Atmospheric Chemistry and Physics*, **7**, 4489-4501.

Thomas, H., et al., 2008: Changes in the North Atlantic Oscillation influence CO2 uptake in the North Atlantic over the past 2 decades. *Global Biogeochemical Cycles*, **22**.

Thompson, D., S. Solomon, P. Kushner, M. England, K. Grise, and D. Karoly, 2011: Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. *Nature Geoscience*, **4**, 741-749.

Thompson, D. W. J., M. P. Baldwin, and J. M. Wallace, 2002: Stratospheric connection to Northern Hemisphere wintertime weather: Implications for prediction. *Journal of Climate*, **15**, 1421-1428.

Thorndike, A. S., D. A. Rothrock, G. A. Maykut, and R. Colony, 1975: THICKNESS DISTRIBUTION OF SEA ICE. *Journal of Geophysical Research-Oceans and Atmospheres*, **80**.

Thorne, P. W., et al., 2011: A quantification of uncertainties in historical tropical tropospheric temperature trends from radiosondes. *J. Geophys. Res.*, **116**, D12116.

Thornton, P. E., J. F. Lamarque, N. A. Rosenbloom, and N. M. Mahowald, 2007: Influence of carbon-nitrogen cycle coupling on land model response to CO2 fertilization and climate variability. *Global Biogeochemical Cycles*, **21**.

Till Kuhlbrodt, and J. Gregory, 2012: Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. *Geophysical Research Letters*, ***submitted***.

Timbal, B., and D. Jones, 2008: Future projections of winter rainfall in southeast Australia using a statistical downscaling technique. *Climatic Change*, **86**, 165-187.

Timmermann, A., S. Lorenz, S.-I. An, A. Clement, and S.-P. Xie, 2007: The Effect of Orbital Forcing on the Mean Climate and Variability of the Tropical Pacific. *J. Climate*, **20**, 4147-4159.

Timmermann, R., H. Goosse, G. Madec, T. Fichefet, C. Ethe, and V. Duliere, 2005: On the representation of high latitude processes in the ORCA-LIM global coupled sea ice-ocean model. *Ocean Modelling*, **8**.

Ting, M., Y. Kushnir, R. Seager, and C. Li, 2009: Forced and Internal Twentieth-Century SST Trends in the North Atlantic. *Journal of Climate*, **22**, 1469-1481.

Tjernstrom, M., J. Sedlar, and M. Shupe, 2008: How well do regional climate models reproduce radiation and clouds in the Arctic? An evaluation of ARCMIP simulations. *Journal of Applied Meteorology and Climatology*, **47**, 2405-2422.

Tjiputra, J. F., K. Assmann, M. Bentsen, I. Bethke, O. H. Ottera, C. Sturm, and C. Heinze, 2010: Bergen Earth system model (BCM-C): model description and regional climate-carbon cycle feedbacks assessment. *Geoscientific Model Development*, **3**, 123-141.

Tjiputra, J. F., et al., 2012: Evaluation of the carbon cycle components inthe Norwegian Earth System Model (NorESM). *Geosci. Model Dev.*, **submitted**.

Trenberth, K. E., and J. M. Caron, 2001: Estimates of meridional atmosphere and ocean heat transports. *Journal of Climate*, **14**, 3433-3443.

Trenberth, K. E., and J. T. Fasullo, 2008: An observational estimate of inferred ocean energy divergence. *Journal of Physical Oceanography*, **38**, 984-999.

——, 2009: Global warming due to increasing absorbed solar radiation. *Geophysical Research Letters*, **36**.

——, 2010a: CLIMATE CHANGE Tracking Earth's Energy. *Science*, **328**, 316-317.

——, 2010b: Simulation of Present-Day and Twenty-First-Century Energy Budgets of the Southern Oceans. *Journal of Climate*, **23**, 440-454.

Trenberth, K. E., D. P. Stepaniak, and J. M. Caron, 2000: The global monsoon as seen through the divergent atmospheric circulation. *Journal of Climate*, **13**, 3969-3993.

Trenberth, K. E., J. T. Fasullo, and J. Kiehl, 2009: Earth's Global Energy Budget. *Bull. Amer. Meteorol. Soc.*, **90**, 311-323.

Tryhorn, L., and A. DeGaetano, 2011: A comparison of techniques for downscaling extreme precipitation over the Northeastern United States. *International Journal of Climatology*, **31**, 1975-1989.

Tschumi, T., F. Joos, and P. Parekh, 2008: How important are Southern Hemisphere wind changes for low glacial carbon dioxide? A model study. *Paleoceanography*, **23**.

Tschumi, T., F. Joos, M. Gehlen, and C. Heinze, 2011: Deep ocean ventilation, carbon isotopes, marine sedimentation and the deglacial CO(2) rise. *Climate of the Past*, **7**, 771-800.

Tsigaridis, K., and M. Kanakidou, 2007: Secondary organic aerosol importance in the future atmosphere. *Atmos. Environ.*, **41**.

Tsujino, H., M. Hirabara, H. Nakano, T. Yasuda, T. Motoi, and G. Yamanaka, 2011: Simulating present climate of the global ocean–ice system using the Meteorological Research Institute Community Ocean Model (MRI. COM): simulation characteristics and variability in the Pacific sector. *Journal of oceanography*. 1-31.

Tsushima, Y., et al., 2006: Importance of the mixed-phase cloud distribution in the control climate for assessing the response of clouds to carbon dioxide increase: a multi-model study. *Clim. Dyn.*, **27**, 113-126.

Turner, A. G., and H. annamalai, 2012: Climate change and the south Asian summer monsoon. *Nature Climate Change*, **2**, 1-9.

Ulbrich, U., G. C. Leckebusch, and J. G. Pinto, 2009: Extra-tropical cyclones in the present and future climate: a review. *Theoretical and Applied Climatology*, **96**, 117-131.

Ulbrich, U., J. G. Pinto, H. Kupfer, G. C. Leckebusch, T. Spangehl, and M. Reyers, 2008: Changing northern hemisphere storm tracks in an ensemble of IPCC climate change simulations. *Journal of Climate*, **21**, 1669-1679.

UNESCO, 1981: Tenth report of the joint panel on oceanographic tables and standards UNESCO.

Uotila, P., A. Lynch, J. Cassano, and R. Cullather, 2007: Changes in Antarctic net precipitation in the 21st century based on Intergovernmental Panel on Climate Change (IPCC) model scenarios. *Journal of Geophysical Research-Atmospheres*, **112**, -.

Uotila, P., S. O’Farrell, S. J. Marsland, and D. Bi, 2012a: The sea-ice performance of the Australian climate models participating in the CMIP5 *Australian Meteorological and Oceanographic Journal*, **submitted**.

Uotila, P., S. O'Farrell, S. J. Marsland, and D. Bi, 2012b: A sea-ice sensitivity study with a global ocean-ice model. *Ocean Modelling*, **51**.

Uppala, S. M., et al., 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131**, 2961-3012.

van den Hurk, B., and E. van Meijgaard, 2010: Diagnosing Land-Atmosphere Interaction from a Regional Climate Model Simulation over West Africa. *Journal of Hydrometeorology*. doi:DOI 10.1175/2009JHM1173.1, 467-481.

van Haren, R., G. J. van Oldenborgh, G. Lenderink, M. Collins, and W. Hazeleger, 2012: SST and circulation trend biases cause an underestimation of European precipitation trends. *Climate Dynamics*. doi:10.1007/s00382-012-1401-5.

van Oldenborgh, G., et al., 2009: Western Europe is warming much faster than expected. *Climate of the Past*, **5**, 1-12.

van Oldenborgh, G. J., S. Y. Philip, and M. Collins, 2005: El Niño in a changing climate: a multi-model study. *Ocean Science*, **1**, 81-–95.

van Roosmalen, L., J. H. Christensen, M. B. Butts, K. H. Jensen, and J. C. Refsgaard, 2010: An intercomparison of regional climate model data for hydrological impact studies in Denmark. *Journal of Hydrology*, **380**, 406-419.

van Vliet, M., S. Blenkinsop, A. Burton, C. Harpham, H. Broers, and H. Fowler: A multi-model ensemble of downscaled spatial climate change scenarios for the Dommel catchment, Western Europe. *Climatic Change*. doi:10.1007/s10584-011-0131-8, 1-29.

——, 2011: A multi-model ensemble of downscaled spatial climate change scenarios for the Dommel catchment, Western Europe. *Climatic Change*. doi:10.1007/s10584-011-0131-8, 1-29.

van Vuuren, D., et al., 2011: The representative concentration pathways: an overview. *Climatic Change*, **109**, 5-31.

Vancoppenolle, M., T. Fichefet, and H. Goosse, 2009a: Simulating the mass balance and salinity of arctic andantarctic sea ice. 2. importance of sea ice salinity variations. *Ocean Modelling*, **27**, 54-69.

Vancoppenolle, M., T. Fichefet, H. Goosse, S. Bouillon, G. Madec, and M. A. M. Maqueda, 2009b: Simulating the mass balance and salinity of arctic and antarctic sea ice. 1. model description and validation. *Ocean Modelling*, **27**, 33-53.

Vancoppenolle, M., H. Goosse, A. de Montety, T. Fichefet, B. Tremblay, and J. L. Tison, 2010: Interactions between brine motion, nutrients and primary production in sea ice. *Journal of Geophysical Research*, **115, C02005**.

Vannière, Benoit, Guilyardi, Eric, Madec, Gurvan, Doblas-Reyes, Francisco J., & Woolnough, Steve. (2012). Using seasonal hindcasts to understand the origin of the equatorial cold tongue bias in CGCMs and its impact on ENSO: *Climate Dynamics*. DOI: 10.1007/s00382-012-1429-6

Vautard, R., et al., 2012: The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. *Climate Dynamics*, **Submitted**.

Vecchi, G. A., and B. J. Soden, 2007: Global warming and the weakening of the tropical circulation. *Journal of Climate*, **20**, 4316-4340.

Vecchi, G. A., B. J. Soden, A. T. Wittenberg, I. M. Held, A. Leetmaa, and M. J. Harrison, 2006a: Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature*, **327**, 216-–219.

Vecchi, G. A., B. J. Soden, A. T. Wittenberg, I. M. Held1, A. Leetmaa, and M. J. Harrison, 2006b: Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature*, **327**, 216-–219.

Veljovic, K., B. Rajkovic, M. J. Fennessy, E. L. Altshuler, and F. Mesinger, 2010: Regional climate modeling: Should one attempt improving on the large scales? Lateral boundary condition scheme: Any impact? *Meteorologische Zeitschrift*, **19**, 237-246.

Verlinde, J., et al., 2007: The Mixed-Phase Arctic Cloud Experiment. *Bull. Amer. Meteorol. Soc.*, **88**, 205-221.

Verseghy, D. L., 2000: The Canadian Land Surface Scheme (CLASS): Its history and future. *Atmosphere-Ocean*, **38**.

Vial, J., and T. J. Osborn, 2012: Assessment of atmosphere-ocean general circulation model simulations of winter northern hemisphere atmospheric blocking. *Climate Dynamics*, **39**, 95-112.

Vichi, M., et al., 2011: Global and regional ocean carbon uptake and climate change: sensitivity to a substantial mitigation scenario. *Climate Dynamics*. 1-19.

Vizcaino, M., U. Mikolajewicz, J. Jungclaus, and G. Schurgers, 2010: Climate modification by future ice sheet changes and consequences for ice sheet mass balance. *Climate Dynamics*, **34**, 301-324.

Vizcaino, M., U. Mikolajewicz, M. Groger, E. Maier-Reimer, G. Schurgers, and A. M. E. Winguth, 2008: Long-term ice sheet-climate interactions under anthropogenic greenhouse forcing simulated with a complex Earth System Model. *Climate Dynamics*, **31**, 665-690.

Voldoire, A., et al., 2012: The CNRM-CM5.1 global climate model : description and basic evaluation. *Climate Dynamics*. doi:doi:10.1007/s00382-011-1259-y

Volodin, E., 2008a: Relation between temperature sensitivity to doubled carbon dioxide and the distribution of clouds in current climate models. *Izvestiya Atmospheric and Oceanic Physics*, **44**, 288-299.

Volodin, E. M., 2007: Atmosphere-ocean general circulation model with the carbon cycle. *Izvestiya Atmospheric and Oceanic Physics*, **43**.

——, 2008b: Methane cycle in the INM RAS climate model. *Izvestiya Atmospheric and Oceanic Physics*, **44**, 153-159.

——, 2008c: Relation between temperature sensitivity to doubled carbon dioxide and the distribution of clouds in current climate models. *Izvestiya Atmospheric and Oceanic Physics*, **44**, 288-299.

Volodin, E. M., and V. N. Lykosov, 1998: Parametrization of heat and moisture transfer in the soil-vegetation system for use in atmospheric general circulation models: 1. Formulation and simulations based on local observational data. *Izvestiya Akademii Nauk Fizika Atmosfery I Okeana*, **34**.

Volodin, E. M., N. A. Dianskii, and A. V. Gusev, 2010: Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. *Izvestiya Atmospheric and Oceanic Physics*, **46**, 414-431.

von Salzen, K., et al., 2012: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4): Simulation of Clouds and Precipitation and Their Responses to Short-Term Climate Variability. *Atmosphere Ocean*, **submitted**.

Vrac, M., and P. Naveau, 2008: Stochastic downscaling of precipitation: From dry events to heavy rainfall (vol 43, art no W07402, 2007). *Water Resources Research*, **44**, -.

Waelbroeck, C., et al., 2009: Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum. *Nature Geoscience*, **2**, 127-132.

Wahl, S., M. Latif, W. Park, and N. Keenlyside, 2011: On the Tropical Atlantic SST warm bias in the Kiel Climate Model. *Climate Dynamics*, **36**, 891-906.

Waliser, D., K. W. Seo, S. Schubert, and E. Njoku, 2007: Global water cycle agreement in the climate models assessed in the IPCC AR4. *Geophysical Research Letters*, **34**.

Waliser, D., Y. Chao, J. Ci, H. Zhang, and K. Seo, 2011a: How Well Do Global Climate Models Simulate Mean Sea Surface Salinity? . *Journal of Climate, submitted*.

Waliser, D., et al., 2009a: MJO Simulation Diagnostics. *Journal of Climate*, **22**, 3006-3030.

Waliser, D. E., J. L. F. Li, T. S. L'Ecuyer, and W. T. Chen, 2011b: The impact of precipitating ice and snow on the radiation balance in global climate models. *Geophysical Research Letters*, **38**.

Waliser, D. E., et al., 2003: AGCM simulations of intraseasonal variability associated with the Asian summer monsoon. *Climate Dynamics*, **21**, 423-446.

Waliser, D. E., et al., 2009b: Cloud ice: A climate model challenge with signs and expectations of progress. *J. Geophys. Res.*, **114**, D00A21, doi 10.1029/2008jd010015.

Walker, G. T., 1924: Correlation in seasonal variations of wether. IX. A further study of world weather, 275-332 pp.

Walther, A., J.-H. Jeong, G. Nikulin, C. Jones, and D. Chen, 2011: Evaluation of the warm season diurnal cycle of precipitation over Sweden simulated by the Rossby Centre regional climate model RCA3. *Atmospheric Research*, **Accepted**.

Wang, B., and Q. H. Ding, 2008: Global monsoon: Dominant mode of annual variation in the tropics. *Dynamics of Atmospheres and Oceans*, **44**, 165-183.

Wang, B., H. J. Kim, K. Kikuchi, and A. Kitoh, 2011a: Diagnostic metrics for evaluation of annual and diurnal cycles. *Climate Dynamics*, **37**, 941-955.

Wang, B., H. Wan, Z. Z. Ji, X. Zhang, R. C. Yu, Y. Q. Yu, and H. T. Liu, 2004: Design of a new dynamical core for global atmospheric models based on some efficient numerical methods. *Science in China Series a-Mathematics*, **47**.

Wang, C., and J. Picaut, 2004: Understanding ENSO Physics - A Review. *Earth's climate: the ocean-atmosphere interaction*, AGU, Washington D.C., C. Wang, S.-P. Xie and J.A. Carton Eds., 21-48.

Wang, J. F., and X. B. Zhang, 2008: Downscaling and projection of winter extreme daily precipitation over North America. *Journal of Climate*, **21**, 923-937.

Wang, M., and J. E. Overland, 2012: Summer Arctic sea ice will be gone sooner or later - an update from CMIP5 models. *Geophysical Research Letters*, **submitted**.

Wang, M., J. E. Overland, and N. A. Bond, 2010: Climate projections for selected large marine ecosystems. *Journal of Marine Systems*, **79**, 258-266.

Wang, S. S., A. P. Trishchenko, K. V. Khlopenkov, and A. Davidson, 2006: Comparison of International Panel on Climate Change Fourth Assessment Report climate model simulations of surface albedo with satellite products over northern latitudes. *Journal of Geophysical Research-Atmospheres*, **111**.

Wang, Y., M. Notaro, Z. Liu, R. Gallimore, S. Levis, and J. E. Kutzbach, 2008: Detecting vegetation-precipitation feedbacks in mid-Holocene North Africa from two climate models. *Climate of the Past*, **4**, 59-67.

Wang, Y. P., and B. Z. Houlton, 2009: Nitrogen constraints on terrestrial carbon uptake: Implications for the global carbon-climate feedback. *Geophysical Research Letters*, **36**.

Wang, Y. P., et al., 2011b: Diagnosing errors in a land surface model (CABLE) in the time and frequency domains. *Journal of Geophysical Research-Biogeosciences*, **116**.

Wania, R., I. Ross, and I. C. Prentice, 2009: Integrating peatlands and permafrost into a dynamic global vegetation model: 1. Evaluation and sensitivity of physical land surface processes. *Global Biogeochemical Cycles*, **23**.

Watanabe, M., 2008: Two regimes of the equatorial warm pool. Part I: A simple tropical climate model. *Journal of Climate*, **21**, 3533-3544.

Watanabe, M., Chikira, M., Imada, Y., & Kimoto, M. (2011). Convective control of ENSO simulated in MIROC. *J. Clim., 24*, 543–562.

Watanabe, M., et al., 2010: Improved Climate Simulation by MIROC5. Mean States, Variability, and Climate Sensitivity. *Journal of Climate*, **23**, 6312-6335.

Watterson, I., and P. Whetton, 2011: Distributions of decadal means of temperature and precipitation change under global warming. *Journal of Geophysical Research-Atmospheres*, **116**.

Waugh, D., and V. Eyring, 2008: Quantitative performance metrics for stratospheric-resolving chemistry-climate models. *Atmospheric Chemistry and Physics*. 5699-5713.

Waugh, D., L. Oman, P. Newman, R. Stolarski, S. Pawson, J. Nielsen, and J. Perlwitz, 2009: Effect of zonal asymmetries in stratospheric ozone on simulated Southern Hemisphere climate trends. *Geophysical Research Letters*. doi:ARTN L18701, 10.1029/2009GL040419, -.

Weaver, A. J., et al., 2001: The UVic Earth System Climate Model: Model description, climatology, and applications to past, present and future climates. *Atmosphere-Ocean*, **39**, 361-428.

Webb, M., C. Senior, S. Bony, and J. J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dynamics*, **17**, 905-922.

Webb, M., et al., 2006: On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dynamics*, **27**, 17-38.

Weber, S. L., et al., 2007: The modern and glacial overturning circulation in the Atlantic Ocean in PMIP coupled model simulations. *Climate of the Past*, **3**, 51-64.

Webster, P. J., A. M. Moore, J. P. Loschnigg, and R. R. Leben, 1999: Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997-98. *Nature*, **401**, 356-360.

Wehner, M., 2012: Very extreme seasonal precipitation in the NARCCAP ensemble: model performance and projections. *Climate Dynamics*. doi:10.1007/s00382-012-1393-1, 1-22.

Wehner, M. F., R. L. Smith, G. Bala, and P. Duffy, 2010: The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model. *Climate Dynamics*, **34**, 241-247.

Wentz, F. J., 1997: A well-calibrated ocean algorithm for special sensor microwave / imager, **102**, 8703-8718.

Wetzel, P., E. Maier-Reimer, M. Botzet, J. H. Jungclaus, N. Keenlyside, and M. Latif, 2006: Effects of ocean Biology on the Penetrative Radiation in a Coupled Climate Model. *Journal of Climate*, **19**, 3973-3987.

Whetton, P., I. Macadam, J. Bathols, and J. O'Grady, 2007a: Assessment of the use of current climate patterns to evaluate regional enhanced greenhouse response patterns of climate models. *Geophysical Research Letters*, **34**, L14701.

——, 2007b: Assessment of the use of current climate patterns to evaluate regional enhanced greenhouse response patterns of climate models. *Geophysical Research Letters*. doi:ARTN L14701, 10.1029/2007GL030025, -.

Wilcox, L. J., A. J. Charlton-Perez, and L. J. Gray, 2012: Trends in Austral jet position in ensembles of high- and low-top CMIP5 models. *J. Geophys. Res.,*, **1*17(*D13), D13115**.

Wild, M., C. N. Long, and A. Ohmura, 2006: Evaluation of clear-sky solar fluxes in GCMs participating in AMIP and IPCC-AR4 from a surface perspective. *Journal of Geophysical Research-Atmospheres*, **111**.

Wilks, D. S., 1995: *Statistical methods in the atmospheric sciences.* Academic Press, 467 pp.

Willett, K., P. Jones, P. Thorne, and N. Gillett, 2010: A comparison of large scale changes in surface humidity over land in observations and CMIP3 general circulation models. *Environmental Research Letters*, **5**, -.

Williams, C. J. R., R. P. Allan, and D. R. Kniveton, 2012: Diagnosing atmosphere-land feedbacks in CMIP5 climate models. submitted.

Williams, K., and M. Webb, 2009: A quantitative performance assessment of cloud regimes in climate models. *Climate Dynamics*. doi:DOI 10.1007/s00382-008-0443-1, 141-157.

Williams, K. D., and G. Tselioudis, 2007: GCM intercomparison of global cloud regimes: present-day evaluation and climate change response. *Climate Dynamics*, **29**, 231-250.

Williams, K. D., and M. E. Brooks, 2008: Initial tendencies of cloud regimes in the Met Office unified model. *Journal of Climate*, **21**, 833-840.

Williams, K. D., W. J. Ingram, and J. M. Gregory, 2008: Time variation of effective climate sensitivity in GCMs. *Journal of Climate*, **21**, 5076-5090.

Williamson, D. G., M., L. Allison, A. Blaker, P. Challenor, and L. Jackson, 2012: History matching for the quantification and reduction of parametric uncertainty in climate model projections. submitted.

Williamson, D. L., and J. G. Olson, 2007: A comparison of forecast errors in CAM2 and CAM3 at the ARM Southern Great Plains site. *Journal of Climate*, **20**, 4572-4585.

Williamson, M. S., T. M. Lenton, J. G. Shepherd, and N. R. Edwards, 2006: An efficient numerical terrestrial scheme (ENTS) for Earth system modelling. *Ecol. Model.*, **198**, 362-374.

Willis, J. K., 2010: Can in situ floats and satellite altimeters detect long-term changes in Atlantic Ocean overturning? *Geophys. Res. Lett.*, **37**, L06602-.

Winterfeldt, J., and R. Weisse, 2009: Assessment of Value Added for Surface Marine Wind Speed Obtained from Two Regional Climate Models. *Monthly Weather Review*. doi:DOI 10.1175/2009MWR2704.1, 2955-2965.

Winterfeldt, J., B. Geyer, and R. Weisse, 2011: Using QuikSCAT in the added value assessment of dynamically downscaled wind speed. *International Journal of Climatology*, **31**, 1028-1039.

Winton, M., 2000: A reformulated three-layer sea ice model. *Journal of Atmospheric and Oceanic Technology*, **17**.

——, 2011: Do Climate Models Underestimate the Sensitivity of Northern Hemisphere Sea Ice Cover? *Journal of Climate*, **24**, 3924-3934.

Wittenberg, A. T., 2009: Are historical records sufficient to constrain ENSO simulations? *Geophys. Res. Lett.*, **36**, -.

Wittenberg, A. T., A. Rosati, N. C. Lau, and J. J. Ploshay, 2006: GFDL's CM2 Global Coupled Climate Models. Part III: Tropical Pacific Climate and ENSO. *Jounal of Climate*, **19**, 698-722.

WMO, 2011: *Scientific Assessment of Ozone Depletion: 2010.*

Wohlfahrt, J., et al., 2008: Evaluation of coupled ocean-atmosphere simulations of the mid-Holocene using palaeovegetation data from the northern hemisphere extratropics. *Climate Dynamics*, **31**, 871-890.

Wood, R., et al., 2011: The VAMOS Ocean-Cloud-Atmosphere-Land Study Regional Experiment (VOCALS-REx): goals, platforms, and field operations. *Atmos. Chem. Phys.*, **11**, 627-654.

Woollings, T., 2008: Vertical structure of anthropogenic zonal-mean atmospheric circulation change. *Geophysical Research Letters*, **35**.

Woollings, T., A. Charlton-Perez, S. Ineson, A. G. Marshall, and G. Masato, 2010a: Associations between stratospheric variability and tropospheric blocking. *Journal of Geophysical Research-Atmospheres*, **115**.

Woollings, T., B. Hoskins, M. Blackburn, D. Hassell, and K. Hodges, 2010b: Storm track sensitivity to sea surface temperature resolution in a regional atmosphere model. *Climate Dynamics*, **35**, 341-353.

Wright, D. G., and T. F. Stocker, 1992: SENSITIVITIES OF A ZONALLY AVERAGED GLOBAL OCEAN CIRCULATION MODEL. *Journal of Geophysical Research-Oceans*, **97**, 12707-12730.

Wu, Q., and G. North, 2003: Statistics of calendar month averages of surface temperature: A possible relationship to climate sensitivity. *Journal of Geophysical Research-Atmospheres*, **108**, 4071.

Wu, Q., D. Karoly, and G. North, 2008a: Role of water vapor feedback on the amplitude of season cycle in the global mean surface air temperature. *Geophysical Research Letters*, **35**, L08711.

Wu, Q. G., D. J. Karoly, and G. R. North, 2008b: Role of water vapor feedback on the amplitude of season cycle in the global mean surface air temperature. *Geophysical Research Letters*, **35**.

Wu, T., 2012: A mass-flux cumulus parameterization scheme for large-scale models: description and test with observations. *Climate Dynamics*, **38**.

Wu, T., R. Yu, and F. Zhang, 2008c: A modified dynamic framework for the atmospheric spectral model and its application. *Journal of the Atmospheric Sciences*, **65**.

Wu, T., et al., 2010a: Erratum - The Beijing Climate Center atmospheric general circulation model: description and its performance for the present-day climate. *Climate Dynamics*, **34**, 149-150.

——, 2010b: The Beijing Climate Center atmospheric general circulation model: description and its performance for the present-day climate. *Climate Dynamics*, **34**, 123-147.

Wyant, M. C., C. S. Bretherton, and P. N. Blossey, 2009: Subtropical Low Cloud Response to a Warmer Climate in a Superparameterized Climate Model. Part I: Regime Sorting and Physical Mechanisms. *Journal of Advances in Modeling Earth Systems*, **1**, doi:10.3894/JAMES.2009.3891.3897.

Wyser, K., et al., 2008: An evaluation of Arctic cloud and radiation processes during the SHEBA year: simulation results from eight Arctic regional climate models. *Climate Dynamics*, **30**, 203-223.

Xavier, P. K., 2012: Intraseasonal convective moistening in CMIP3 models. *Journal of climate*, **25**, 2569-2577.

Xavier, P. K., J. P. Duvel, P. Braconnot, and F. J. Doblas-Reyes, 2010: An Evaluation Metric for Intraseasonal Variability and its Application to CMIP3 Twentieth-Century Simulations. *Journal of Climate*, **23**, 3497-3508.

Xiao, X., et al., 1998: Transient climate change and net ecosystem production of the terrestrial biosphere. *Global Biogeochemical Cycles*, **12**.

Xie, L., T. Z. Yan, L. J. Pietrafesa, J. M. Morrison, and T. Karl, 2005: Climatology and interannual variability of North Atlantic hurricane tracks. *Journal of Climate*, **18**, 5370-5381.

Xie, P., and P. A. Arkin, 1997: Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs. *Bull. Amer. Meteorol. Soc.*, **78**, 2539-2558.

Xie, S., J. Boyle, S. A. Klein, X. Liu, and S. Ghan, 2008: Simulations of Arctic mixed-phase clouds in forecasts with CAM3 and AM2 for M-PACE. *J. Geophys. Res.*, **113**, D04211.

Xie, S., H.-Y. Ma, J. S. Boyle, S. A. Klein, and Y. Zhang, 2012: On the Correspondence between Short- and Long- Timescale Systematic Errors in CAM4/CAM5 for the Years of Tropical Convection. *Journal of Climate*. doi:10.1175/jcli-d-12-00134.1.

Xie, S. P., H. Annamalai, F. A. Schott, and J. P. McCreary, 2002: Structure and mechanisms of South Indian Ocean climate variability. *Journal of Climate*, **15**, 864-878.

Xin, X.-G., T.-J. Zhou, and R.-C. Yu, 2008: The Arctic Oscillation in coupled climate models. *Chinese Journal of Geophysics-Chinese Edition*, **51**, 337-351.

Xin, X., L. Zhang, J. Zhang, T. Wu, and Y. Fang, 2012a: Climate change projections over East Asia with BCC_CSM1.1 under RCP scenarios *JMSJ*, **submitted**.

Xin, X., T. WU, J. LI, Z. WANG, W. LI, and F. WU, 2012b: How well does BCC_CSM1.1 reproduce the 20th century climate change over China? . *AOSL*, **submitted**.

Xue, Y. K., R. Vasic, Z. Janjic, F. Mesinger, and K. E. Mitchell, 2007: Assessment of dynamic downscaling of the continental US regional climate using the Eta/SSiB regional climate model. *Journal of Climate*, **20**, 4172-4193.

Yakovlev, N. G., 2009: Reproduction of the large-scale state of water and sea ice in the Arctic Ocean in 1948-2002: Part I. Numerical model. *Izvestiya Atmospheric and Oceanic Physics*, **45**.

Yang, D., and T. Ohata, 2001: A Bias-Corrected Siberian Regional Precipitation Climatology. *Journal of Hydrometeorology*, **2**, 122-139.

Yang, W., J. Andreasson, L. Graham, J. Olsson, J. Rosberg, and F. Wetterhall, 2010: Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrology Research*, **41**, 211-229.

Yasutaka, W., N. Masaomi, K. Sachie, and M. Chiashi, 2008: Climatological Reproducibility Evaluation and Future Climate Projection of Extreme Precipitation Events in the Baiu Season Using a High-Resolution Non-Hydrostatic RCM in Comparison with an AGCM. *Journal of the Meteorological Society of Japan. Ser. II*, **86**, 951-967.

Yhang, Y. B., and S. Y. Hong, 2008: Improved physical processes in a regional climate model and their impact on the simulated summer monsoon circulations over East Asia. *Journal of Climate*, **21**, 963-979.

Yin, J., S. Griffies, and R. Stouffer, 2010: Spatial Variability of Sea Level Rise in Twenty-First Century Projections. *Journal of Climate*. doi:DOI 10.1175/2010JCLI3533.1, 4585-4607.

Yin, X., A. Gruber, and P. Arkin, 2004: Comparison of the GPCP and CMAP merged gauge-satellite monthly precipitation products for the period 1979-2001. *Journal of Hydrometeorology*, **5**, 1207-1222.

Yokohata, T., M. J. Webb, M. Collins, K. D. Williams, M. Yoshimori, J. C. Hargreaves, and J. D. Annan, 2010a: Structural Similarities and Differences in Climate Responses to CO2 Increase between Two Perturbed Physics Ensembles. *Journal of Climate*, **23**, 1392-1410.

Yokohata, T., M. Webb, M. Collins, K. Williams, M. Yoshimori, J. Hargreaves, and J. Annan, 2010b: Structural Similarities and Differences in Climate Responses to CO2 Increase between Two Perturbed Physics Ensembles. *Journal of Climate*, **23**, 1392-1410.

Yokohata, T., J. Annan, M. Collins, C. Jackson, M. Tobis, M. Webb, and J. Hargreaves, 2012a: Reliability of multi-model and structurally different single-model ensembles. *Climate Dynamics*, **39**, 599-616.

Yokohata, T., et al., 2012b: Reliability and importance of structural diversity of climate model ensembles. *Climate Dynamics*. submitted.

Yokohata, T., et al., 2008: Comparison of equilibrium and transient responses to CO2 increase in eight state-of-the-art climate models. *Tellus Series a-Dynamic Meteorology and Oceanography*, **60**, 946-961.

Yokoi, S., et al., 2011: Application of Cluster Analysis to Climate Model Performance Metrics. *Journal of Applied Meteorology and Climatology*, **50**, 1666-1675.

Yokoi, T., T. Tozuka, and T. Yamagata, 2009: Seasonal Variations of the Seychelles Dome Simulated in the CMIP3 Models. *Journal of Physical Oceanography*, **39**, 449-457.

Yoshimori, M., T. Yokohata, and A. Abe-Ouchi, 2009: A Comparison of Climate Feedback Strength between CO2 Doubling and LGM Experiments. *Journal of Climate*, **22**, 3374-3395.

Yoshimori, M., J. C. Hargreaves, J. D. Annan, T. Yokohata, and A. Abe-Ouchi, 2011: Dependency of Feedbacks on Forcing and Climate State in Physics Parameter Ensembles. *Journal of Climate*, **24**, 6440-6455.

Young, P. J., et al., 2012a: Late twentieth century Southern Hemisphere stratospheric temperature trends in observations and CCMVal-2, CMIP3 and CMIP5 models. *J. Geophys. Res.*, submitted.

Young, P. J., et al., 2012b: Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). *Atmos. Chem. Phys.*, **submitted**.

Yu, B., and F. W. Zwiers, 2010: Changes in equatorial atmospheric zonal circulations in recent decades. *Geophys. Res. Lett.*, **37**, L05701.

Yu, J.-Y., and S. T. Kim, 2011: Reversed Spatial Asymmetries between El Nino and La Nina and Their Linkage to Decadal ENSO Modulation in CMIP3 Models. *Journal of Climate*, **24**, 5423-5434.

Yu, W., M. Doutriaux, G. Seze, H. LeTreut, and M. Desbois, 1996: A methodology study of the validation of clouds in GCMs using ISCCP satellite observations. *Climate Dynamics*, **12**, 389-401.

Yukimoto, S., et al., 2011: Meteorological Research Institute-Earth System Model v1 (MRI-ESM1) - Model Description.

Yukimoto, S., et al., 2012: A New Global Climate Model of the Meteorological Research Institute: MRI-CGCM3-Model Description and Basic Performance. *Journal of the Meteorological Society of Japan*, **90A**.

Zaehle, S., P. Friedlingstein, and A. D. Friend, 2010a: Terrestrial nitrogen feedbacks may accelerate future climate change. *Geophysical Research Letters*, **37**.

Zaehle, S., A. D. Friend, P. Friedlingstein, F. Dentener, P. Peylin, and M. Schulz, 2010b: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 2. Role of the nitrogen cycle in the historical terrestrial carbon balance. *Global Biogeochemical Cycles*, **24**.

Zahn, M., and H. von Storch, 2008: A long-term climatology of North Atlantic polar lows. *Geophysical Research Letters*, **35**.

Zalesny, V. B., et al., 2010: Numerical simulation of large-scale ocean circulation based on the multicomponent splitting method. *Russian Journal of Numerical Analysis and Mathematical Modelling*, **25**.

Zaliapin, I., and M. Ghil, 2010: Another look at climate sensitivity. *Nonlinear Processes in Geophysics*, **17**, 113-122.

Zappa, G. S. L. C. H. K. I., 2012: The ability of CMIP5 models to simulate North Atlantic extratropical cyclones. *Journal of Climate*, **Submitted**.

Zelinka, M. D., S. A. Klein, and D. L. Hartmann, 2012: Computing and Partitioning Cloud Feedbacks Using Cloud Property Histograms. Part II: Attribution to Changes in Cloud Amount, Altitude, and Optical Depth. *Journal of Climate*, **25**, 3736-3754.

Zeng, N., 2003: Glacial-interglacial atmospheric CO2 change - The glacial burial hypothesis. *Advances in Atmospheric Sciences*, **20**, 677-693.

——, 2006: Quasi-100ky glacial-interglacial cycles triggered by subglacial burial carbon release. *Climate of the Past*, **2**, 371-397.

Zeng, N., J. D. Neelin, and C. Chou, 2000: A quasi-equilibrium tropical circulation model - Implementation and simulation. *Journal of the Atmospheric Sciences*, **57**, 1767-1796.

Zeng, N., A. Mariotti, and P. Wetzel, 2005: Terrestrial mechanisms of interannual CO(2) variability. *Global Biogeochemical Cycles*, **19**.

Zeng, N., H. F. Qian, E. Munoz, and R. Iacono, 2004: How strong is carbon cycle-climate feedback under global warming? *Geophysical Research Letters*, **31**.

Zhang, D., and M. J. McPhaden, 2006: Decadal variability of the shallow Pacific meridional overturning circulation: Relation to tropical sea surface temperatures in observations and climate change models. *Ocean Modelling*. 250-273.

Zhang, F., and A. Georgakakos, 2011: Joint variable spatial downscaling. *Climatic Change*. doi:10.1007/s10584-011-0167-9, 1-28.

Zhang, J., J. S. Reid, D. L. Westphal, N. L. Baker, and E. J. Hyer, 2008a: A system for operational aerosol optical depth data assimilation over global oceans. *Journal of Geophysical Research-Atmospheres*, **113**.

Zhang, Q., H. S. Sundqvist, A. Moberg, H. Kornich, J. Nilsson, and K. Holmgren, 2010a: Climate change between the mid and late Holocene in northern high latitudes - Part 2: Model-data comparisons. *Climate of the Past*, **6**, 609-626.

Zhang, W., and F.-F. Jin, 2012: Do Current Coupled Models Better Simulate ENSO-SSTA Meridional Width? , Submitted ed., GRL.

Zhang, X., 2007: A comparison of explicit and implicit spatial downscaling of GCM output for soil erosion and crop production assessments. *Climatic Change*, **84**, 337-363.

Zhang, X., et al., 2007: Detection of human influence on twentieth-century precipitation trends. *Nature*. doi:DOI 10.1038/nature06025, 461-U464.

Zhang, Y., S. A. Klein, J. Boyle, and G. G. Mace, 2010b: Evaluation of tropical cloud and precipitation statistics of Community Atmosphere Model version 3 using CloudSat and CALIPSO data. *J. Geophys. Res.*, **115**, D12205.

Zhang, Y., D. J. Seidel, J.-C. Golaz, C. Deser, and R. A. Tomas, 2011: Climatological Characteristics of Arctic and Antarctic Surface-Based Inversions*. *Journal of Climate*, **24**, 5167-5186.

Zhang, Y., et al., 2008b: On the diurnal cycle of deep convection, high-level cloud, and upper troposphere water vapor in the Multiscale Modeling Framework. *J. Geophys. Res.*, **113**, D16105.

Zhao, M., I. M. Held, and S.-J. Lin, 2012: Some Counterintuitive Dependencies of Tropical Cyclone Frequency on Parameters in a GCM. *Journal of the Atmospheric Sciences*, **69**, 2272-2283.

Zhao, M., I. M. Held, S. J. Lin, and G. A. Vecchi, 2009: Simulations of Global Hurricane Climatology, Interannual Variability, and Response to Global Warming Using a 50-km Resolution GCM. *Journal of Climate*, **22**, 6653-6678.

Zhao, T. L., et al., 2008: A three-dimensional model study on the production of BrO and Arctic boundary layer ozone depletion. *Journal of Geophysical Research*, **113**, D24304.

Zheng, X.-T., S.-P. Xie, and Q. Liu, 2011: Response of the Indian Ocean basin mode and its capacitor effect to global warming. *Jounal of Climate*, **in press**.

Zheng, Y., J.-L. Lin, and T. Shinoda, 2012: The equatorial Pacific cold tongue simulated by IPCC AR4 coupled GCMs: Upper ocean heat budget and feedback analysis. *Journal of Geophysical Research-Oceans*, **117**.

Zhou, T., B. Wu, and B. Wang, 2009: How Well Do Atmospheric General Circulation Models Capture the Leading Modes of the Interannual Variability of the Asian-Australian Monsoon? *Journal of Climate*, **22**, 1159-1173.

Zhu, Y., and H. Wang, 2010: The Arctic and Antarctic Oscillations in the IPCC AR4 Coupled Models. *Acta Meteorologica Sinica*, **24**, 176-188.

Zickfeld, K., et al., submitted: Long-term climate change commitment and reversibility. *J. Climate*.

Zunz, V., H. Goosse, and F. Massonnet, 2012: How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent? *The Cryosphere*.

Zwiers, F. W., X. Zhang, and Y. Feng, 2011: Anthropogenic Influence on Long Return Period Daily Temperature Extremes at Regional Scales. *Journal of Climate*, **24**, 881-892.

Tables

Table 9.1: Salient features of the AOGCMs and ESMs participating in CMIP5. Column 1: identification (Model ID) along with the calendar year ('vintage') of the first publication for each model; Column 2: sponsoring institution(s), main reference(s) and flux correction implementation (*not yet described*); Subsequent Columns for each of the 8 CMIP5 realms: component name, code independence and main component reference(s). Additionally, there are standard entries for the Atmosphere realm: horizontal grid resolution, number of vertical levels, grid top (low or high top); and for the Ocean realm: horizontal grid resolution, number of vertical levels, top level, vertical coordinate type, ocean free surface type ("Top BC"). This table information was automatically extracted from the CMIP5 online questionnaire (http://q.cmip5.ceda.ac.uk/) as of 12 November 2011.

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
(1) ACCESS1.0 (2) 2011	(1) Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia (2) (Bi et al., 2012b) (Dix et al., 2012)	(1) Aerosols (CLASSIC) (2) (Bellouin et al., 2011; Dix et al., 2012)	(1) Atmosphere (as in HadGEM2 (r1.1)) (2) 192x145 N96 (3) 38 (4) 39,255m (5) (Bi et al., 2012b; Rashid et al., 2012) (Martin et al., 2011)	Not implemented	(1) Land Ice (2) (Kowalczyk et al., 2012)	(1) Land Surface (MOSES2.2) (2) (Kowalczyk et al., 2012), (Cox et al., 1999a), (Essery et al., 2003)	Not implemented	(1) ACCESS-OM (MOM4p1) (2) primarily 1 degree latitude/longitude tripolar with enhanced resolution near equator and at high latitudes (3) 50 (4) 0-10 m (5) z* (6) non-linear split-explicit (7) (Bi et al., 2012a); (Marsland et al., 2012)	(1) Sea Ice (CICE4.1) (2) (Bi et al., 2012a); (Uotila et al., 2012a; Uotila et al., 2012b)
(1) ACCESS1.3 (2) 2011	(1) Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia (2)	(1) Aerosols (CLASSIC) (2) (Bellouin et al., 2011; Dix et al., 2012)	(1) Atmosphere (similar to UK Met Office Global Atmosphere 1.0) (2) 192x145 N96 (3) 38 (4) 39,255m (5)	Not implemented	(1) Land Ice (2) (Kowalczyk et al., 2012)	(1) CABLE (2) (Kowalczyk et al., 2012) (Kowalczyk et al., 2006) (Wang et al., 2011b)	Not implemented	(1) ACCESS-OM (MOM4p1) (2) primarily 1 degree latitude/longitude tripolar with enhanced resolution near equator and at	(1) Sea Ice (CICE4.1) (2) (Bi et al., 2012a); (Uotila et al., 2012a; Uotila et al., 2012b)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
	(Bi et al., 2012b) (Dix et al., 2012)		(Bi et al., 2012b) (Sun et al., 2012), (Rashid et al., 2012) (Hewitt et al., 2011)					high latitudes (3) 50 (4) 0-10 m (5) z* (6) non-linear split-explicit (7) (Bi et al., 2012a); (Marsland et al., 2012)	
(1) BCC-CSM1.1 (2) 2011	(1) Beijing Climate Center, China Meteorological Administration (2) (Wu, 2012; Xin et al., 2012a; Xin et al., 2012b)	Not implemented	(1) BCC_AGCM2.1 (2) T42 T42L26 (3) 26 (4) 2.917hPa (5) (Wu et al., 2008c) , (Wu et al., 2010b), (Wu et al., 2010a), (Wu, 2012)	Not implemented	Not implemented	(1) BCC-AVIM1.0 (2) (Wu, 2012)	(1) OceanBiogeoChemistry (2) Based on the protocols from the Ocean Carbon Cycle Model Intercomparison Project–Phase 2 (OCMIP2, http://www.ipsl.jussieu.fr/OCMIP/phase2/)	(1) MOM4-L40 (2) 1° with enhanced resolution in the meridional direction in the tropics (1/3° meridional resolution at the equator) tripolar (3) 40 (4) 25 m (5) z (6) linear split-explicit (7) (Griffies et al., 2005)	(1) GFDL Sea Ice Simulator SIS (2) (Winton, 2000)
(1) BCC-CSM1.1(m) (2) 2011	(1) Beijing Climate Center, China Meteorological Administration (2) (Wu, 2012; Xin et al., 2012a; Xin et al., 2012b)	Not implemented	(1) BCC_AGCM2.1 (2) T106 (3) 26 (4) 2.917hPa (5) (Wu et al., 2008c) , (Wu et al., 2010b), (Wu et al., 2010a), (Wu, 2012)	Not implemented	Not implemented	(1) BCC-AVIM1.0 (2) (Wu, 2012)	(1) Ocean BiogeoChemistry (2) Based on the protocols from the Ocean Carbon Cycle Model Intercomparison Project–Phase 2 (OCMIP2, http://www.ipsl.jussieu.fr/OCMIP/	(1) MOM4-L40 (2) Tri-polar: 1° with enhanced resolution in the meridional direction in the tropics (1/3° meridional resolution at the equator) (3) 40 (4) 25m (5) z	(1) GFDL Sea Ice Simulator (SIS) (2) (Winton, 2000)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
							phase2/)	(6) implicit (7) (Griffies et al., 2005)	
(1) BNU-ESM (2) 2011	(1) Beijing Normal University (2)	(1) Aerosols (2)	(1) CAM3.5 (2) T42 (3) 26 (4) 2.194hPa	Not implemented	Not implemented	(1)CoLM+BNUDGVM(C/N) (2) (Dai and Trenberth, 2004; Dai et al., 2004b; Dai et al., 2003)	(1) IBGC (2)	(1) MOM4p1 (2) 200(lat)*360(lon) (3) 50	(1) CICE4.1 (2)
(1) CanESM2 (2) 2010	(1) Canadian Center for Climate Modelling and Analysis (2) (Arora et al., 2011b), (von Salzen et al., 2012)	(1) Aerosols (2) (Lohmann et al., 1999) (Croft et al., 2005) (von Salzen et al., 2012)	(1) Atmosphere (2) Spectral T63 (3) 35 levels (4) 0.5hPa (5) (von Salzen et al., 2012)	(1) Atmospheric Chemistry (2) (von Salzen et al., 2012)	Not implemented	(1) CLASS 2.7; CTEM (2) (Verseghy, 2000) (Arora et al., 2009; von Salzen et al., 2012)	(1) CMOC (2) (Arora et al., 2009), (Christian et al., 2010)	(1) Ocean (2) 256x192 (3) 40 (4) 0 m (5) depth (6) rigid lid (7) (Merryfield et al., 2012)	(1) sea ice (2) (Merryfield et al., 2012)
(1) CanCM4 (2) 2010	(1) Canadian Center for Climate Modelling and Analysis (2) (von Salzen et al., 2012)	(1) Aerosols (2) (Lohmann et al., 1999) (Croft et al., 2005) (von Salzen et al., 2012)	(1) Atmosphere (2) Spectral T63 (3) 35 levels (4) 0.5hPa (5) (von Salzen et al., 2012)	(1) Atmospheric Chemistry (2) (von Salzen et al., 2012)	Not implemented	(1) CLASS 2.7 (2) (Verseghy, 2000), (von Salzen et al., 2012)	Not implemented	(1) Ocean (2) 256x192 (3) 40 (4) 0 m (5) depth (6) rigid lid (7) (Merryfield et al., 2012)	(1) sea ice (2) (Merryfield et al., 2012)
(1) CanAM4 (2) 2010	(1) Canadian Center for Climate Modelling and Analysis (2) (von Salzen et al., 2012)	(1) Aerosols (2) (Lohmann et al., 1999) (Croft et al., 2005) (von Salzen et al., 2012)	(1) Atmosphere (2) Spectral T63 (3) 35 levels (4) 0.5hPa (5) (von Salzen et al., 2012)	(1) Atmospheric Chemistry (2) (von Salzen et al., 2012)	Not implemented	(1) CLASS 2.7 (2) (Verseghy, 2000), (von Salzen et al., 2012)	Not implemented	Not implemented	Not implemented
(1) CCSM4 (2) 2010	(1) US National Centre for Atmospheric	(1) Aerosols (2) (Neale et al., 2010)	(1) CAM4 (2) 0.9x1.25 f09 (3) 27	Not implemented	Not implemented	(1) Community Land Model 4 (CLM4) (2) (Oleson et al.,	Not implemented	(1) POP2 with modifications (2) Nominal 1 degree	(1) CICE4 with modifications (2)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
	Research (2) (Gent et al., 2011)	(Oleson et al., 2010) (Holland et al., 2012b)	(4) 2.194067 hPa (5) (Neale et al., 2012; Neale et al., 2010)			2010) (Lawrence et al., 2011), (Lawrence et al., 2012)		(1.125 degree in longitude, 0.27-0.64 degree variable in latitude) (3) 60 (4) 10 m thick with surface variables at 5 m (5) depth (level) (6) linearized, implicit free surface with constant-volume ocean (7) (Danabasoglu et al., 2012)	(Hunke and Lipscomb, 2011) (Holland et al., 2012b)
(1) CESM1(CAM5) (2) 2010	(1) NSF-DOE-NCAR (2) (Hurrell et al., 2012)	(1) Aerosols (2) (Neale et al., 2010) (Oleson et al., 2010) (Holland et al., 2012b)	(1) Community Atmosphere Model 5 (CAM5) (2) 0.9x1.25 f09 (3) 27 (4) 2.194067 hPa (5) (Neale et al., 2012; Neale et al., 2010)	Not implemented	Not implemented	(1) CLM4 (2) (Oleson et al., 2010) (Lawrence et al., 2011), (Lawrence et al., 2012)	Not implemented	(1) POP2 with modifications (2) Nominal 1 degree (1.125 degree in longitude, 0.27-0.64 degree variable in latitude) (3) 60 (4) 10 m thick with surface variables at 5 m (5) depth (level) (6) linearized, implicit free surface with constant-volume ocean (7) (Danabasoglu et al., 2012)	(1) CICE4 with modifications (2) (Hunke and Lipscomb, 2011) (Holland et al., 2012b)
(1) CESM1(BGC)	(1) NSF-DOE-	(1) Aerosols	(1) CAM4	Not implemented	Not implemented	(1) CLM4	(1) Biogeochemical	(1) POP2 with	(1) CICE4 with

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
(2) 2010	NCAR (2) (Lindsay et al., 2012; Long et al., 2012) (Hurrell et al., 2012)	(2) (Neale et al., 2010) (Oleson et al., 2010) (Holland et al., 2012b)	(2) 0.9x1.25 f09 (3) 27 (4) 2.194067 hPa (5) (Neale et al., 2012; Neale et al., 2010)			(2) (Oleson et al., 2010) (Lawrence et al., 2011), (Lawrence et al., 2012)	Elemental Cycling (BEC) (2) (Moore et al., 2012)	modifications (2) Nominal 1 degree (1.125 degree in longitude, 0.27-0.64 degree variable in latitude) (3) 60 (4) 10 m thick with surface variables at 5 m (5) depth (level) (6) linearized, implicit free surface with constant-volume ocean (7) (Danabasoglu et al., 2012)	modifications (2) (Hunke and Lipscomb, 2011) (Holland et al., 2012b)
(1) CESM1(WACCM) (2) 2010	(1) NSF-DOE-NCAR (2) (Hurrell et al., 2012; Marsh et al., 2012)	(1) Aerosols (2) (Marsh et al., 2012)	(1) WACCM4 (2) 1.9° x 2.5° (3) 66 (4) 5.1x10^{-6} hPa (5) (Marsh et al., 2012)	(1) Atmos Chemistry (2) (Marsh et al., 2012)	Not implemented	(1) CLM4 (2) (Oleson et al., 2010) (Lawrence et al., 2011), (Lawrence et al., 2012)	Not implemented	(1) POP2 with modifications (2) Nominal 1 degree (1.125 degree in longitude, 0.27-0.64 degree variable in latitude) (3) 60 (4) 10 m thick with surface variables at 5 m (5) depth (level) (6) linearized, implicit free surface with constant-volume ocean (7) (Danabasoglu et al., 2012)	(1) CICE4 with modifications (2) (Hunke and Lipscomb, 2011) (Holland et al., 2012b)
(1) CESM1(CHEM)	(1) NSF-DOE-	(1) Aerosols	(1) Atmosphere,	(1) Atmos	Not implemented	(1) Community Land	Not implemented	(1) POP2 with	(1) CICE4 with

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
(2) 2010	NCAR (2) (Eyring et al., 2012a) (Cameron-Smith et al., 2006; Hurrell et al., 2012)	(2) (Lamarque et al., 2012), (Neale et al., 2010) (Oleson et al., 2010) (Holland et al., 2012b)	CAM4-CHEM (2) 0.9x1.25 f09 (3) 27 (4) 2.194067 hPa (5) Lamarque et al., 2012, (Neale et al., 2012; Neale et al., 2010)	Chemistry, CAM-CHEM (2) (Lamarque et al., 2012)		Model 4 (CLM4) (2) (Oleson et al., 2010) (Lawrence et al., 2011), (Lawrence et al., 2012)		modifications (2) Nominal 1 degree (1.125 degree in longitude, 0.27-0.64 degree variable in latitude) (3) 60 (4) 10 m thick with surface variables at 5 m (5) depth (level) (6) linearized, implicit free surface with constant-volume ocean (7) (Danabasoglu et al., 2012)	modifications (2) (Hunke and Lipscomb, 2011) (Holland et al., 2012b)
(1)CESM1(CISM) (2) 2010	(1) NSF-DOE-NCAR (2) (Hurrell et al., 2012; Lipscomb et al., 2012)	(1) Aerosols (2) (Neale et al., 2010) (Oleson et al., 2010) (Holland et al., 2012b)	(1) CAM4 (2) 0.9x1.25 f09 (3) 27 (4) 2.194067 hPa (5) (Neale et al., 2012; Neale et al., 2010)	Not implemented	(1)Glimmer-CISM (2) (Lipscomb et al., 2012; Rutt et al., 2009)	(1) Community Land Model 4 (CLM4) (2) (Oleson et al., 2010) (Lawrence et al., 2011), (Lawrence et al., 2012)	Not implemented	(1) POP2 with modifications (2) Nominal 1 degree (1.125 degree in longitude, 0.27-0.64 degree variable in latitude) (3) 60 (4) 10 m thick with surface variables at 5 m (5) depth (level) (6) linearized, implicit free surface with constant-volume ocean (7) (Danabasoglu et al., 2012)	(1) CICE4 with modifications (2) (Hunke and Lipscomb, 2011) (Holland et al., 2012b)
(1) CMCC-CM	(1) Centro Euro-	Not implemented	(1) ECHAM5	Not implemented	Not implemented	Not implemented	Not implemented	(1) OPA8.2	(1) LIM2

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
(2) 2009	Mediterraneo per I Cambiamenti Climatici (2) (Scoccimarro et al., 2011) (Fogli et al., 2009)		(2) 0.75°x0.75° (T159) (3) 31 (4) 10 hPa (5) (Roeckner et al., 2006)					(2) 2° average, 0.5° at the equator (ORCA2) (3) 31 (4) 5 m (5) depth (Z-level) (6) linear implicit (7) (Madec et al., 1998)	(2) (Timmermann et al., 2005)
(1) CMCC-CMS (2) 2009	(1) Centro Euro-Mediterraneo per I Cambiamenti Climatici (2) (Fogli et al., 2009; Manzini and al., 2012)	Not implemented	(1) ECHAM5 (2) 1.875°x1.875° (T63) (3) 95 (4) 0.01 hPa (5) (Manzini et al., 2006; Roeckner et al., 2006)	Not implemented	Not implemented	Not implemented	Not implemented	(1) OPA8.2 (2) (2) 2° average, 0.5° at the equator (ORCA2) (3) 31 (4) 5 m (5) depth (Z-level) (6) linear implicit (7) (Madec et al., 1998)	(1) LIM2 (2) (Timmermann et al., 2005)
(1) CMCC-CESM (2) 2009	(1) Centro Euro-Mediterraneo per I Cambiamenti Climatici (2) (Fogli et al., 2009; Vichi et al., 2011)	Not implemented	(1) ECHAM5 (2) 3.75°x3.75° (T31) (3) 39 (4) 0.01 hPa (5) (Manzini et al., 2006; Roeckner et al., 2006)	Not implemented	Not implemented	(1) SILVA (2) (Alessandri et al., 2012)	(1) PELAGOS (2) Vichi et al. 2007	(1) OPA8.2 (2) 2° average, 0.5° at the equator (ORCA2) (3) 31 (4) 5 m (5) depth (Z-level) (6) linear implicit (7) (Madec et al., 1998)	(1) LIM2 (2) (Timmermann et al., 2005)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
(1) CNRM-CM5 (2) 2010	(1) Centre National de Recherches Meteorologiques and Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique. (2) (Voldoire et al., 2012)	Not implemented	(1) ARPEGE (Atmosphere) (2) tl127r (3) 31 (4) 10 hPa (5) (Météo-France, 2011), (Deque et al., 1994), (Voldoire et al., 2012)	Not implemented	Not implemented	(1) SURFEX (Land and Ocean Surface) (2) (Voldoire et al., 2012)	Not implemented	(1) NEMO (2) 0.7 degree on average ORCA1 (3) 42 (4) 5 m (5) Z-coordinate (6) linear filtered (7) (Madec, 2008)	(1) Gelato5 (Sea Ice) (2) (Melia, 2002)
(1) CSIRO-Mk3.6.0 (2) 2009	(1) Queensland Climate Change Centre of Excellence and Commonwealth Scientific and Industrial Research Organisation (2) (Rotstayn, 2012)	(1) Aerosols (2) (Rotstayn et al., 2011) (Rotstayn, 2012), (Rotstayn and Lohmann, 2002))	(1) Atmosphere (2) ~1.875x1.875 (spectral T63) (3) 18 (4) ~4.5 hPa (5) (Rotstayn, 2012), (Gordon et al., 2010) (Gordon et al., 2002)	Not implemented	Not implemented	(1) Land Surface (2) (Gordon et al., 2010) (Gordon et al., 2002)	Not implemented	(1) Modified MOM2.2 (2) ~0.9x1.875 (3) 31 (4) 5 m (5) depth (6) rigid lid (7) (Gordon et al., 2010) (Gordon et al., 2002)	(1) Sea ice (2) (O'Farrell, 1998), (Gordon et al., 2010)
(1) EC-EARTH (2) 2010	(1) Europe (2)	Not implemented	(1) IFS c31r1 (2) 1.125 longitudinal spacing, Gaussian grid T159L62 (3) 62 (4) 1 hPa (5) Hazeleger et al. 2011, Hazeleger et al. 2012	Not implemented	Not implemented	(1) HTESSEL (2)	Not implemented	(1) NEMO_ecmwf (2) The grid is a tripolar curvilinear grid with a 1 degree resolution. ORCA1 (3) 31 (4) 1 m (5) (6) linear filtered (7)	(1) LIM2 (2)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
(1) FGOALS-g2 (2) 2011	1) LASG(Institute of Atmospheric Physics)-CESS(Tsinghua University) (2)	Not implemented	(1) GAMIL2 (2) 2.8125x2.8125degree (3) 26 layers (4) 2.194hPa (5) (Wang et al., 2004)	Not implemented	Not implemented	(1) CLM3 (2) Oleson et al., 2010	Not implemented	(1) LICOM2 (2) 1x1degree with 0.5 meridional degree in the tropical region (3) 30 (4) 10 m (5) ç-coor (6) (7)	(1) CICE4-LASG (2) (Wang and Houlton, 2009), (Liu, 2010)
(1) FGOALS-s2 (2) 2011	(1) The State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, The Institute of Atmospheric Physics (2) (Bao et al., 2010; Bao et al., 2012)	Not implemented	(1) SAMIL2.4.7 (2)R42 (2.81°x1.66°) (3) 26 (4) 2.19hPa (5)	Not implemented	Not implemented	(1) CLM3.0 (2) (Oleson, 2004)	(1) IAP-OBM (2) Xu et al., 2012;	(1) LICOM (2) The zonal resolution is 1°. The meridional resolution is 0.5° between 10°S and 10°N and increases from 0.5° to 1° from 10° (3) 30 layers (4) 10 meter (for vertical velocity and pressure) and 5 meter (for Temperature and salinity, zonal and meridional velocity) (5) depth (6) linear split-explicit (7) (Lin et al., 2012a)	(1) CSIM5 (2) (Briegleb et al., 2004)
(1) FIO-ESM v1.0 (2) 2011	(1) The First Institute of Oceanography, State Oceanic Administration, China (2) (Song et al., 2012)	Not implemented	(1) CAM3.0 (2) T42 (3) 26 (4) 3.545hPa (5) (Collins et al., 2006b)	Not implemented	Not implemented	(1) CLM3.5 (2) (Oleson et al., 2008c)	(1) Improved OCMIP-2 biogeochemical model (2)	(1) POP2.0 (2) 320x384 (3) 40 (4) 5 m (5) depth (6) linear implicit (7)	(1) CICE4.0 (2) (Hunke and Lipscomb, 2011)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
	(Fangli et al., 2004)								
(1) GFDL-ESM2M (2) 2011	(1) NOAA Geophysical Fluid Dynamics Laboratory (2) (Delworth et al., 2006)	Aerosols	Atmosphere	Not implemented	Not implemented	(1) Land Surface (2)	(1) Ocean Biogeo Chemistry (2)	(1) Ocean (2) 1 degree tripolar 360X200L50 (3) 50 (4) 0 m (5) other (6) non-linear split-explicit (7)	(1) SIS (2)
(1) GFDL-ESM2G (2) 2011	(1) NOAA Geophysical Fluid Dynamics Laboratory (2) (Delworth et al., 2006)	Aerosols	Atmosphere	Not implemented	Not implemented	(1) Land Surface (2)	(1) Ocean Biogeo Chemistry (2)	(1) Ocean (2) 1 degree tripolar360x210L63 (3) 63 (4) 0 m (5) other (6) non-linear split-explicit (7)	(1) SIS (2)
(1) gfdl-cm2p1 (2) 2011	(1) NOAA Geophysical Fluid Dynamics Laboratory (2) (Delworth et al., 2006)	(1) Aerosols (2)	(1) Atmosphere (2) 2.5 degree longitude, 2 degree latitude M45L24 (3) 24 (4) midpoint of top box is 3.65 hPa	Not implemented	(1) Land Ice (2)	(1) Land Surface (2)	Not implemented	(1) Ocean (2) 1 degree tripolar360x200L50 (3) 50 (4) 0 m (5) depth (6) non-linear split-explicit (7)	(1) SIS (2)
(1) GFDL-CM3 (2) 2011	(1) NOAA Geophysical Fluid Dynamics Laboratory (2) (Delworth et al., 2006; Donner et al., 2011)	(1) Aerosols (2) (Levy et al., 2012)	(1) Atmosphere (2) ~200km C48L48 (3) 48 (4) 0.01 hPa (5) (Donner et al., 2011)	(1) Atmospheric Chemistry (2) (Austin and Wilson, 2006) (Horowitz et al., 2003) (Sander, 2006)	Not implemented	(1) Land Surface (2) (Milly and Shmakin, 2002) (Shevliakova et al., 2009)	Not implemented	(1) Ocean (2) 1 degree tripolar360x200L50 (3) 50 (4) 0 m (5) depth (6) non-linear split-	(1) SIS (2) (Griffies, 2012)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
								explicit (7) (Griffies, 2012))	
(1) GFDL-HIRAM-C180 (2) 2011	(1) NOAA Geophysical Fluid Dynamics Laboratory (2) (Delworth et al., 2006; Donner et al., 2011)	Not implemented	(1) Atmosphere (2) Averaged cell size: approximately 50x50 km. C180L32 (3) 32 (4) 2.164 hPa (5) (Donner et al., 2011)	Not implemented	Not implemented	(1) Land Surface (2) (Milly and Shmakin, 2002) (Shevliakova et al., 2009)	Not implemented	Not implemented	Not implemented
(1) GFDL-HIRAM-C360 (2)	(1) NOAA Geophysical Fluid Dynamics Laboratory (2) (Delworth et al., 2006; Donner et al., 2011)	Not implemented	(1) Atmosphere (2) Averaged cell size: approximately 25x25 km. C360L32 (3) 32 (4) 2.164 hPa (5) (Donner et al., 2011);	Not implemented	Not implemented	(1) Land Surface (2) (Milly and Shmakin, 2002) (Shevliakova et al., 2009)	Not implemented	Not implemented	Not implemented
(1) GISS-E2CS-H (2) 2011	(1) NASA Goddard Institute for Space Studies USA (2) (Schmidt et al., 2006);	(1) Aerosols (2) (Bauer et al., 2007; Koch et al., 2011; Menon et al., 2010; Tsigaridis and Kanakidou, 2007)	(1) Atmosphere (2) Nominally 1 deg (3) 40 (4) 0.1 hPa	(1) G-PUCCINI (2) (Shindell et al., 2012a)	(1) Land Ice (2)	(1) Land Surface (2)	Not implemented (2) (Romanoua et al., 2012)	(1) HYCOM Ocean (2) 0.2 to 1 deg latitude x 1 deg longitude HYCOM (3) 26 (4) 0 m (5) hybrid Z-isopycnic (6) non-linear split-explicit (7)	(1) Sea Ice (2)
(1) GISS-E2-R (2) 2011	(1) NASA Goddard Institute for Space Studies USA (2) (Schmidt et al., 2006)	(1) Aerosols (2) (Bauer et al., 2007; Koch et al., 2011; Menon et al., 2010; Tsigaridis and Kanakidou, 2007)	(1) Atmosphere (2) 2 deg latitude x 2.5 deg longitude F (3) 40 (4) 0.1 hPa	(1) G-PUCCINI (2) (Shindell et al., 2012a)	(1) Land Ice (2)	(1) Land Surface (2)	Not implemented (2) (Romanoua et al., 2012)	(1) Russell Ocean (2) 1 deg latitude x 1.25 deg longitude Russell 1x1Q (3) 32 (4) 0 m	(1) Sea Ice (2)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
								(5) Z*-coordinate (6) other (7)	
(1) GISS-E2CS-R (2) 2011	(1) NASA Goddard Institute for Space Studies USA (2) (Schmidt et al., 2006)	(1) Aerosols (2) (Bauer et al., 2007; Koch et al., 2011; Menon et al., 2010; Tsigaridis and Kanakidou, 2007)	(1) Atmosphere (2) Nominally 1 deg (3) 40 (4) 0.1 hPa	(1) G-PUCCINI (2) (Shindell et al., 2012a)	(1) Land Ice (2)	(1) Land Surface (2)	Not implemented (2) (Romanoua et al., 2012)	(1) Russell Ocean (2) 1 deg latitude x 1.25 deg longitude Russell 1x1Q (3) 32 (4) 0 m (5) Z*-coordinate (6) other (7)	(1) Sea Ice (2)
(1) GISS-E2-H (2) 2004	(1) NASA Goddard Institute for Space Studies USA (2) (Schmidt et al., 2006)	(1) Aerosols (2) (Bauer et al., 2007; Koch et al., 2011; Menon et al., 2010; Tsigaridis and Kanakidou, 2007)	(1) Atmosphere (2) 2 deg latitude x 2.5 deg longitude F (3) 40 (4) 0.1 hPa	(1) G-PUCCINI (2) (Shindell et al., 2012a)	(1) Land Ice (2)	(1) Land Surface (2)	Not implemented (2) (Romanoua et al., 2012)	(1) HYCOM Ocean (2) 0.2 to 1 deg latitude x 1 deg longitude HYCOM (3) 26 (4) 0 m (5) hybrid Z-isopycnic (6) non-linear split-explicit (7)	(1) Sea Ice (2)
(1) HadGEM2-ES (2) 2009	(1) UK Met Office Hadley Centre (2) (Bellouin et al., 2007; Collins et al., 2008; Martin et al., 2011)	(1) Aerosols (2) (Bellouin et al., 2011)	(1) Atmosphere (2) 1.875 degrees in longitude by 1.25 degrees in latitude N96 (3) 38 (4) 39254.8 m (5) (Davies et al., 2005)	(1) Atmospheric Chemistry (2) (O'Connor et al., 2009)	(1) Land Ice (2)(Johns et al., 2006)	(1) Land Surface (2) (Cox et al., 1999a) (Essery et al., 2003)	(1) Ocean Biogeo Chemistry (2) (Halloran, 2012; Palmer and Totterdell, 2001)	(1) Ocean (2) 1 deg by 1 deg between 30 N/S and the poles; meridional resolution increases to 1/3 deg at the equator N180 (3) 40 (4) 5.0 m (5) depth (6) linear implicit (7) (Bryan and	(1) Sea Ice (2) (McLaren et al., 2006; Thorndike et al., 1975)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
								Lewis, 1979) (Johns et al., 2006);	
(1) HadGEM2-CC (2) 2010	(1) UK Met Office Hadley Centre (2) (Bellouin et al., 2007; Collins et al., 2008; Martin et al., 2011)	(1) Aerosols (2) (Bellouin et al., 2011)	(1) Atmosphere (2) 1.875 deg in longitude by 1.25 deg in latitude N96 (3) 60 (4) 84132.439 m (5) (Davies et al., 2005)	(1) Atmospheric Chemistry (2) (Martin et al., 2011) (Jones et al., 2001)	(1) Land Ice (2) Johns_2006;	(1) Land Surface (2) (Cox et al., 1999a; Essery et al., 2003)	(1) Ocean Biogeo Chemistry (2) (Halloran, 2012; Palmer and Totterdell, 2001)	(1) Ocean (2) 1.875 deg in longitude by 1.25 deg in latitude N96 (3) (4) (5) hybrid height (6) linear implicit (7) (Bryan and Lewis, 1979) (Johns et al., 2006);	(1) Sea Ice (2) (McLaren et al., 2006; Thorndike et al., 1975)
(1) HadGEM2-A (2) 2009	(1) UK Met Office Hadley Centre (2) (Bellouin et al., 2007; Collins et al., 2008; Martin et al., 2011)	(1) Aerosols (2) (Bellouin et al., 2011)	(1) Atmosphere (2) 1.875 degrees in longitude by 1.25 degrees in latitude N96 (3) 38 (4) 39254.8 m (5) (Davies et al., 2005)	Not implemented	(1) Land Ice (2) (Johns et al., 2006);	(1) Land Surface (2) (Cox et al., 1999a) (Essery et al., 2003)	Not implemented	Not implemented	Not implemented
(1) HadGEM2-AO (2) 2009	(1) UK Met Office Hadley Centre (2) (Bellouin et al., 2007; Collins et al., 2008; Martin et al., 2011)	(1) Aerosols (2) Bellouin et al. 2007;	(1) Atmosphere (2) 1.875 deg in longitude by 1.25 deg in latitude N96 (3) 60 (4) 84132.439 m (5) (Davies et al., 2005)	Not implemented	Not implemented	(1) Land Surface (2) (Cox et al., 1999a; Essery et al., 2003)	Not implemented	(1) Ocean (2) 1.875 deg in longitude by 1.25 deg in latitude N96 (3) (4) (5) hybrid height (6) linear implicit (7) (Bryan and Lewis, 1979) (Johns et al., 2006);	(1) Sea Ice (2) (McLaren et al., 2006; Thorndike et al., 1975)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
(1) HadCM3 (2) 1998	(1) UK Met Office Hadley Centre (2) (Collins et al., 2001; Gordon et al., 2000; Johns et al., 2003; Pope et al., 2000)	(1) Aerosols (2) (Jones et al., 2001)	(1) Atmosphere (2) HadAM3 (N48L19) 3.75 degrees in longitude by 2.5 degrees in latitude. N48 (3) 19 (4) 0.005 hPa (5) (Pope et al., 2000)	Not implemented	Not implemented	(1) Land Surface (2) (Collatz et al., 1992; Collatz et al., 1991; Cox, 2001b; Cox et al., 1999a; Mercado et al., 2007)	Not implemented	(1) Ocean HadOM (lat: 1.25 lon: 1.25 L20) (2) 1.25 deg in longitude by 1.25 deg in latitude N144 (3) 20 (4) 5.0 m (5) depth (6) linear implicit (7) (UNESCO, 1981)	(1) Sea Ice (2)
(1) INM-CM4 (2) 2009	(1) Russian Institute for Numerical Mathematics (2) (Volodin et al., 2010)	Not implemented	(1) Atmosphere (2) 2x1.5 degrees in longitude and latitude latitude-longitude (3) 21 (4) sigma=0.01	Not implemented	(1) Land Ice (2)	(1) Land Surface (2) (Alekseev et al., 1998; Volodin and Lykosov, 1998)	(1) Ocean Biogeo Chemistry (2) (Volodin, 2007)	(1) Ocean (2) 1x0.5 degrees in longitude and latitude generalized spherical coordinates with poles displaced outside ocean (3) 40 (4) sigma=0.0010426 (5) sigma (6) linear implicit (7) (Volodin et al., 2010; Zalesny et al., 2010)	(1) Sea Ice (2) (Yakovlev, 2009)
(1) IPSL-CM5A-LR (2) 2010	(1) Institut Pierre Simon Laplace (2) (Dufresne et al., 2012)	Not implemented	(1) Atmosphere (2) 96x95 equivalent to 1,9° x 3,75° LMDZ96x95 (3) 39 (4) 0.04 hPa (5)(Hourdin F. et al., 2012)	Not implemented	Not implemented	(1) Land Surface (2) (Krinner et al., 2005)	(1) Ocean Biogeo Chemistry (2) (Aumont and Bopp, 2006; Aumont et al., 2003)	(1) Ocean (2) 2° ORCA2 (3) 31 (4) 0m (5) depth (6) linear filtered (7) (Madec, 2008)	(1) Sea ice (2) (Fichefet and Maqueda, 1999)
(1) IPSL-CM5A-MR (2) 2009	(1) Institut Pierre Simon Laplace	Not implemented	(1) Atmosphere (2) 144x143	Not implemented	Not implemented	(1) Land Surface (2) (Krinner et al.,	(1) Ocean Biogeo Chemistry	(1) Ocean (2) 2° ORCA2	(1) Sea ice (2) (Fichefet and

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
	(2) (Dufresne et al., 2012)		equivalent to 1,25° x 2,5° LMDZ144x143 (3) 39 (4) 0.04 hPa (5) (Hourdin F. et al., 2012)			2005)	(2) (Aumont and Bopp, 2006; Aumont et al., 2003)	(3) 31 (4) 0m (5) depth (6) linear filtered (7) (Madec, 2008)	Maqueda, 1999)
(1) IPSL-CM5B-LR (2) 2010	(1) Institut Pierre Simon Laplace (2) (Dufresne et al., 2012)	Not implemented	(1) Atmosphere (2) 96x95 equivalent to 1,9° x 3,75° LMDZ96x95 (3) 39 (4) 0.04 hPa (5)(Hourdin et al., 2012)	Not implemented	Not implemented	(1) Land Surface (2) (Krinner et al., 2005)	(1) Ocean Biogeo Chemistry (2) (Aumont and Bopp, 2006; Aumont et al., 2003)	(1) Ocean (2) 2° ORCA2 (3) 31 (4) 0m (5) depth (6) linear filtered (7) (Madec, 2008)	(1) Sea ice (2) (Fichefet and Maqueda, 1999)
(1) MPI-ESM-LR (2) 2009	(1) Max Planck Institute for Meteorology (2) (Giorgetta et al., 2012)	Not implemented	(1) ECHAM6 (2) approx 1.8 deg T63 (3) 47 (4) 0.01 hPa (5) (Stevens et al., 2012)	Not implemented	Not implemented	(1) JSBACH (2) (Reick et al., 2012)	(1) HAMOCC (2) (Maier-Reimer et al., 2005), (Ilyina et al., 2012);	(1) MPIOM (2) average 1.5 deg GR15 (3) 40 (4) 6 m (5) depth (6) linear implicit (7) (Jungclaus et al., 2012)	(1) Sea Ice (2) (Notz et al., 2012)
(1) MPI-ESM-MR (2) 2009	(1) Max Planck Institute for Meteorology (2) (Giorgetta et al., 2012)	Not implemented	(1) ECHAM6 (2) approx 1.8 deg T63 (3) 95 (4) 0.01 hPa (5) (Stevens et al., 2012)	Not implemented	Not implemented	(1) JSBACH (2) (Reick et al., 2012)	(1) HAMOCC (2) (Maier-Reimer et al., 2005), (Ilyina et al., 2012);	(1) MPIOM (2) approx. 0.4 deg TP04 (3) 40 (4) 6 m (5) depth (6) linear implicit (7) (Jungclaus et al., 2012)	(1) Sea Ice (2) (Notz et al., 2012)
(1) MPI-ESM-P (2) 2009	(1) Max Planck Institute for Meteorology	Not implemented	(1) ECHAM6 (2) approx 1.8 deg T63	Not implemented	Not implemented	(1) JSBACH (2) (Reick et al., 2012)	(1) HAMOCC (2) (Maier-Reimer et al., 2005), (Ilyina et	(1) MPIOM (2) approx. 0.4 deg TP04	(1) Sea Ice (2) (Notz et al., 2012)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
	(2) (Giorgetta et al., 2012)		(3) 47 (4) 0.01 hPa (5) (Stevens et al., 2012)				al., 2012)	(3) 40 (4) 6 m (5) depth (6) linear implicit (7) (Jungclaus et al., 2012)	
(1) MRI-ESM1 (2) 2011	(1) Meteorological Research Institute (2)(Adachi et al., 2012; Yukimoto et al., 2011; Yukimoto et al., 2012)	(1) MASINGAR mk-2 (2) (Adachi et al., 2012; Yukimoto et al., 2011; Yukimoto et al., 2012)	(1)MRI-AGCM3.3 (2)TL159(320x160) (3)48 (4)0.01hPa (5) (Adachi et al., 2012; Yukimoto et al., 2011; Yukimoto et al., 2012)	(1) MRI-CCM2 (2) (Adachi et al., 2012; Deushi and Shibata, 2011; Yukimoto et al., 2011)	(1) SMIST (2) (Yukimoto et al., 2011; Yukimoto et al., 2012)	(1) HAL (2) (Yukimoto et al., 2011; Yukimoto et al., 2012)	(1) Ocean Biogeo Chemistry (MRI.COM3) (2) (Adachi et al., 2012; Nakano et al., 2011)	(1) MRI.COM3 (2) 1x0.5 (3) 50 + 1 Bottom Boundary Layer (4) 0m (5) hybrid sigma-z (6) non-linear split-explicit (7)(Tsujino et al., 2011; Yukimoto et al., 2011; Yukimoto et al., 2012)	(1) Sea Ice (MRI.COM3) (2) (Tsujino et al., 2011; Yukimoto et al., 2011; Yukimoto et al., 2012)
(1) MRI-AGCM3.2S (2) 2009	(1) Meteorological Research Institute (2) (Mizuta et al., 2012)	Not implemented	(1) Atmosphere (2) 1920x960 TL959 (3) 64 (4) 0.01hPa (5) (Mizuta et al., 2012)	Not implemented	Not implemented	(1) SiB0109 (2) (Hirai et al., 2007) (Yukimoto et al., 2011; Yukimoto et al., 2012)	Not implemented	Not implemented	Not implemented
(1) MRI-AGCM3.2H (2) 2009	(1) Meteorological Research Institute (2) (Mizuta et al., 2012)	Not implemented	(1) Atmosphere (2) 640x320 TL319 (3) 64 (4) 0.01hPa	Not implemented	Not implemented	(1) SiB0109 (2) (Hirai et al., 2007) (Yukimoto et al., 2011; Yukimoto et al., 2012)	Not implemented	Not implemented	Not implemented
(1) MRI-CGCM3 (2) 2011	(1) Meteorological Research Institute (2) (Yukimoto et al., 2011; Yukimoto et	(1) MASINGAR mk-2 (2) (Adachi et al., 2012; Yukimoto et	(1)MRI-AGCM3.3 (2)320x160 TL159 (3)48 (4)0.01hPa	Not implemented	(1) SMIST (2) (Yukimoto et al., 2011; Yukimoto et al., 2012)	(1) HAL (2) (Yukimoto et al., 2011; Yukimoto et al., 2012)	Not implemented	(1) MRI.COM3 (2) 1x0.5 (3) 50 + 1 Bottom Boundary Layer	(1) Sea Ice (MRI.COM3) (2) (Tsujino et al., 2011; Yukimoto et

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
	al., 2012)	al., 2011; Yukimoto et al., 2012)	(5) (Yukimoto et al., 2011; Yukimoto et al., 2012)					(4) 0m (5) hybrid sigma-z (6) non-linear split-explicit (7) (Tsujino et al., 2011; Yukimoto et al., 2011; Yukimoto et al., 2012)	al., 2011; Yukimoto et al., 2012)
(1) MIROC4h (2) 2009	(1) University of Tokyo, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology (2) (Sakamoto et al., 2012)	(1) SPRINTARS (2) (Takemura et al., 2002; Takemura et al., 2000)	(1) CCSR / NIES / FRCGC AGCM5.7 (2) 0.5625x0.5625 degree T213 (3) 56 (4) about 0.9 hPa	Not implemented	Not implemented	(1) MATSIRO (2) (Takata et al., 2003)	Not implemented	(1) COCO3.4 (2) 1/4° by 1/6° (average grid spacing is 0.28° and 0.19° for zonal and meridional directions) (3) 48 (4) 1.25m (5) hybrid z-s (6) non-linear split-explicit (7) (Hasumi and Emori, 2004)	(1) Sea Ice (2)
(1) MIROC5 (2) 2010	(1) University of Tokyo, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology (2) (Watanabe et al., 2010)	(1) SPRINTARS (2) (Takemura et al., 2005; Takemura et al., 2009)	(1) CCSR / NIES / FRCGC AGCM6 (2) 1.40625 x 1.40625 degree T85 (3) 40 (4) about 2.9 hPa	Not implemented	Not implemented	(1) MATSIRO (2) (Takata et al., 2003)	Not implemented	(1) COCO4.5 (2) 1.4degree (zonally) x 0.5-1.4 degree (meridionally) (3) 50 (4) 1.25m (5) hybrid s-z (6) linear split-explicit (7) (Hasumi and Emori, 2004)	(1) Sea Ice (2) (Komuro et al., 2012)
(1) MIROC-ESM (2) 2010	(1) University of Tokyo, National	(1) SPRINTARS (2) (Takemura et al.,	(1) MIROC-AGCM (2) 2.8125x2.8125	Not implemented	Not implemented	(1) MATSIRO (2) (Takata et al.,	(1) NPZD-type (2) (Schmittner et al.,	(1) COCO3.4 (2) 1.4degree	(1) Sea Ice (2)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
	Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology (2) (Watanabe et al., 2011)	2005; Takemura et al., 2009)	degree T42 (3) 80 (4) 0.003 hPa (5) (Watanabe, 2008)			2003)	2005)	(zonally) x 0.5-1.4 degree (meridionally) (3) 44 (4) 1.25m (5) hybrid s-z (6) linear split-explicit (7) (Hasumi and Emori, 2004)	
(1) MIROC-ESM-CHEM (2) 2010	(1) University of Tokyo, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology (2) (Watanabe et al., 2011)	(1) SPRINTARS (2) (Takemura et al., 2005; Takemura et al., 2009)	(1) MIROC-AGCM (2) 2.8125x2.8125 degree T42 (3) 80 (4) 0.003 hPa (5) (Watanabe, 2008)	(1) CHASER (2) Sudo et al. 2002	Not implemented	(1) MATSIRO (2) (Takata et al., 2003)	(1) NPZD-type (2) (Schmittner et al., 2005)	(1) COCO3.4 (2) 1.4degree (zonally) x 0.5-1.4 degree (meridionally) (3) 44 (4) 1.25m (5) hybrid s-z (6) linear split-explicit (7) (Hasumi and Emori, 2004)	(1) Sea Ice (2)
(1) NCEP-CFSv2 (2) 2011	(1) National Centers for Environmental Prediction (2)	Not implemented	(1) Global Forecast Model (2) 0.9375 T126 (3) 64 (4) 0.03 hPa (5) (Saha et al., 2010)	(1) Ozone chemistry (2) (McCormack et al., 2006)	Not Implemented	(1) Noah Land Surface Model (2) (Ek et al., 2003)	Not implemented	(1) MOM4 (2) 0.5° zonal resolution, meridional resolution varying from 0.25° at the equator to 0.5° north/south of 10N/10S. Tripolar. (3) 40 (4) 5.0m (5) depth (6) non-linear split explicit (7) (Griffies et al.,	(1) SIS (2) (Hunke and Dukowicz, 1997; Winton, 2000)

(1) **Model ID** (2) Vintage	(1) **Institution** (2) Main Reference(s)	**Aerosols** (1) Component Name (2) References	**Atmosphere** (1) Component Name (2) Horizontal Grid (3) Number of Vert Levels (4) Grid Top (5) References	**Atmos Chemistry** (1) Component Name (2) References	**Land Ice** (1) Component Name (2) References	**Land Surface** (1) Component Name (2) References	**Ocean** Biogeochemistry (1) Component Name (2) References	**Ocean** (1) Component Name (2) Horizontal Resolution (3) Number of Vertical Levels (4) Top Level (5) Z Co-ord (6) Top BC (7) References	**Sea Ice** (1) Component Name (2) References
								2004)	
(1) NorESM1-M (2) 2011	(1) Norwegian Climate Centre (2) (Bentsen et al., 2012; Iversen et al., 2012)	(1) CAM4-Oslo (2) (Kirkevåg et al., 2012)	(1) CAM4-Oslo (2) Finite Volume 1.9 degrees latitude, 2.5 degrees longitude (3) 26 (4) 2.194067 hPa (5) (Kirkevåg et al., 2012) (Neale et al., 2010)	(1) CAM4-Oslo (2) (Kirkevåg et al., 2012)	Not implemented	(1) CLM4 (2) (Lawrence et al., 2011; Oleson et al., 2010)	Not implemented	(1) NorESM-Ocean (2) 1.125 degrees along the equator (3) 53 (4) 1 m (5) hybrid Z-isopycnic (6) non-linear split-explicit (7) (Bentsen et al., 2012)	(1)CICE4 (2)(Hunke and Lipscomb, 2011) (Holland et al., 2012b)
(1) NorESM1-ME (2) 2012	(1) Norwegian Climate Centre (2) (Bentsen et al., 2012; Tjiputra et al., 2012)	(1) CAM4-Oslo (2) (Kirkevåg et al., 2012)	(1) CAM4-Oslo (2) Finite Volume 1.9 degrees latitude, 2.5 degrees longitude (3) 26 (4) 2.194067 hPa (5) (Kirkevåg et al., 2012) (Neale et al., 2010)	(1) CAM4-Oslo (2) (Kirkevåg et al., 2012)	Not implemented	(1) CLM4 (2) (Lawrence et al., 2011; Oleson et al., 2010)	(1) HAMOCC5 (2) (Assmann et al., 2010; Maier-Reimer et al., 2005; Tjiputra et al., 2012)	(1) NorESM-Ocean (2) (2) 1.125 degrees along the equator (3) 53 (4) 1 m (5) hybrid Z-isopycnic (6) non-linear split-explicit (7) (Bentsen et al., 2012)	(1) CICE4 (2) (Hunke and Lipscomb, 2011) (Holland et al., 2012b)

Table 9.2. Overview of observations that are used to evaluate climate characteristics in this chapter. The variable and CMIP5 output variable name are given along with references for the observations.

Variable	CMIP5 output variable name	Observations (default / alternates)	Reference	Figure and Section Number(s)
ATMOSPHERE				
Surface Air Temperature (°C)	tas (2m)	ERA-Interim	(Dee et al., 2011)	Figures. 9.2, Section 9.4.1, Figure 9.6, Section 9.4.1
		NCEP-NCAR	(Kalnay et al., 1996)	Figure 9.7^{A} , Section 9.4.1
		CRU TS 3.10	(Mitchell and Jones, 2005)	
		HadCRUT4	(Morice et al., 2012)	Figure 9.8, Section 9.4.1; Figures 9.38, 9.39, 9.40, Section 9.6.1.1
		GISTEMP	(Hansen et al., 2010)	
		NCDC	(Smith et al., 2008b)	
Temperature (°C)	ta (upper air)	ERA-Interim	(Dee et al., 2011)	Figures 9.40, Section 9.6.1, Figure 9.7^{D}, Section 9.4.1
		NCEP-NCAR	(Kalnay et al., 1996)	Figure 9.7^{A}, Section 9.4.1
		CRU TS 3.10	(Mitchell and Jones, 2005)	
		HadCRUT4	(Morice et al., 2012)	Figures 9.38, 9.39, 9.40 , Section 9.6.1.1
		GISTEMP	(Hansen et al., 2010)	
		NCDC	(Smith et al., 2008b)	
Zonal mean wind (m s^{-1})	uas (2m)	ERA-Interim	(Dee et al., 2011)	Figure 9.7^{D} Section 9.4.1
	ua (upper air)	NCEP-NCAR[1]	(Kalnay et al., 1996)	Figure 9.7^{A} Section 9.4.1

Meridional wind (m s^{-1})	vas (2m)	ERA-Interim	(Dee et al., 2011)	Figure 9.7^{D} Section 9.4.1
	va (upper air)	NCEP-NCAR	(Kalnay et al., 1996)	Figure 9.7^{A} Section 9.4.1
Geopotential height (m)	zg	ERA-Interim	(Dee et al., 2011)	Figure 9.7^{D} Section 9.4.1
		NCEP-NCAR	(Kalnay et al., 1996)	Figure 9.7^{A} Section 9.4.1
Specific humidity (kg kg^{-1})	hus	AIRS	(Chahine et al., 2006)	
		ERA-Interim	(Dee et al., 2011)	Figure 9.7^{D} Section 9.4.1
TOA reflected shortwave radiation (W m^{-2})	rsut	CERES	(Loeb et al., 2009)	Figure 9.7^{D} Section 9.4.1
		ERBE	(Barkstrom, 1984)	Figure 9.7^{A} Section 9.4.1
TOA longwave clear-sky radiation (W m^{-2})	rlut	CERES	(Loeb et al., 2009)	Figure 9.7^{D} Section 9.4.1
		ERBE	(Barkstrom, 1984)	Figure 9.7^{A} Section 9.4.1
Clear sky TOA reflected shortwave radiation (W m^{-2})	rsutcs	CERES	(Loeb et al., 2009)	Figure 9.7^{D} Section 9.4.1
		ERBE	(Barkstrom, 1984)	Figure 9.7^{A} Section 9.4.1
Clear sky TOA reflected shortwave radiation (W m^{-2})	rlutcs	CERES	(Loeb et al., 2009)	Figure 9.7^{D} Section 9.4.1
		ERBE	(Barkstrom, 1984)	Figure 9.7^{A} Section 9.4.1
Total precipitation (mm day^{-1})	precip	GPCP	(Adler et al., 2003)	Figure 9.7^{D}, Section 9.4.1
		CMAP	(Xie and Arkin, 1997)	Figure 9.40, Section 9.6.1.1, Figure 9.7^{D} Section 9.4.1
		CRU TS3.10.1		Figure 9.38, Section 9.6.1.1
3-hour precipitation fields		15 000 stations and corrected	Dai (2006)	

3h- surface temperature fields		15 000 stations and corrected Ta from COADS	(Dai and Deser, 1999; Dai et al., 2004a)	
			(Dai and Deser, 1999)	
Precipitable water		RSS	(Wentz, 1997)	Figure 9.7^{D}, Section 9.4.1
Total column ozone (DU)	tro3	Ground-based measurements NASA TOMS/OMI/SBUV(/2) merged satellite data, NIWA combined total column ozone database, Solar Backscatter Ultraviolet (SBUV, SBUV/2) retrievals, DLR GOME/SCIA/GOME-2	updated from Fioletov et al. (2002), (Stolarski and Frith, 2006), (Bodeker et al., 2005), updated from Miller et al. (2002), (Loyola and Coldewey-Egbers, 2012; Loyola et al., 2009).	Figure 9.10, Section 9.4.1.3.5

CARBON CYCLE

Atmospheric CO_2 (ppmv)	co2		(Masarie and Tans, 1995) (Meinshausen et al., 2011)	Figure 9.46, Section 9.8.3
Global Land Carbon Sink (PgC yr^{-1})	NBP	GCP	(Le Quere et al., 2009)	Figures 9.26, 9.27, Section 9.4.5
Global Ocean Carbon Sink (PgC yr^{-1})	fgCO2	GCP	(Le Quere et al., 2009)	Figures 9.26, 9.27, Section 9.4.5
Regional Land Sinks (PgC yr^{-1})	NBP	JAM	(Gurney et al., 2003)	Figure 9.27, Section 9.4.5
Regional Ocean Sinks (PgC yr^{-1})	fgCO2	JAM	(Gurney et al., 2003)	Figure 9.27, Section 9.4.5
			(Takahashi et al., 2009)	

OCEAN

Annual mean temperature	thetao		(Levitus et al., 2009)	Figure 9.13, Section. 9.4.2

Annual mean salinity	so		(Antonov et al., 2010)	Figure 9.13, Section. 9.4.2
Global ocean heat content (0–700 m)	OHC	Levitus Ishii Domingues	(Levitus et al., 2009) (Ishii and Kimoto, 2009) (Domingues et al., 2008)	Figure 9.17, Section. 9.4.2
Annual mean temperature and salinity		Paleoclimate reconstruction of temperature and salinity (Adkins et al. 2002)	(Otto-Bliesner et al., 2007a)	
CRYOSPHERE				
Total area (km^2) of grid cells where Sea Ice Area Fraction (%) is >15%			(Rayner et al., 2003b)	Figure 9.22
			(Comiso, 2008) (updated from 1999)	
Total area (km^2) of grid cells where Surface Snow Area Fraction (%) is 15% or Surface Snow Amount (kg/m^2) is >5 kg/m^2			(Robinson and Frei, 2000b)	Figure 9.23

Table 9.4: Climate sensitivity estimates (Andrews et al., 2012b) and feedback parameters (Soden et al., 2008) for the CMIP5 AOGCMs (see Table 9.1 for model details). The entries were calculated according to Hansen et al (2005) for radiative forcing using fixed SSTs; Gregory et al (2004) for radiative forcing and equilibrium climate sensitivity using regression; Soden et al (2008) for the feedback parameters using radiative kernel methods; and Taylor et al's (2012) reference CMIP5 experiment with 1% CO_2 increase per year for the transient climate response using the 20-year mean centred on the year of CO_2 doubling. Notice that the entries for radiative forcing and equilibrium climate sensitivity were obtained by dividing by two the original results, which were obtained for CO_2 quadrupling.

Model	**Radiative Forcing ($W\ m^{-2}$)**		**Water Vapour Feedback Parameter ($W\ m^{-2}\ °C^{-1}$)**	**Lapse Rate Feedback Parameter ($W\ m^{-2}\ °C^{-1}$)**	**Surface Albedo Feedback Parameter ($W\ m^{-2}\ °C^{-1}$)**	**Cloud Feedback Parameter ($W\ m^{-2}\ °C^{-1}$)**	**Climate Feedback Parameter ($W\ m^{-2}\ °C^{-1}$)**	**Climate Sensitivity Parameter ($°C\ (W\ m^{-2})^{-1}$)**	**Equilibrium Climate Sensitivity (°C)**	**Transient Climate Response (°C)**
	Fixed SST	**Regression**								
ACCESS1-0	n.a.	3.0	1.6	–0.5	0.4	0.5	0.8	1.3	3.8	2.0
CanESM2	3.7	3.8	1.7	–0.5	0.4	0.5	1.0	1.0	3.7	2.4
CCSM4	4.4	3.6	n.a.	n.a.	n.a.	n.a.	1.2	0.8	2.9	1.8
CNRM-CM5	n.a.	3.7	n.a.	n.a.	n.a.	n.a.	1.1	0.9	3.3	2.1
CSIRO-Mk3-6-0	3.1	2.6	n.a.	n.a.	n.a.	n.a.	0.6	1.6	4.1	1.8
FGOALS-s2	n.a.	3.8	n.a.	n.a.	n.a.	n.a.	0.9	1.1	4.2	2.4
GFDL-CM3	n.a.	3.0	1.6	–0.6	0.4	0.9	0.8	1.3	4.0	2.0
GFDL-ESM2G	n.a.	3.1	n.a.	n.a.	n.a.	n.a.	1.3	0.8	2.4	1.1
GFDL-ESM2M	n.a.	3.4	1.9	–0.8	0.3	0.2	1.4	0.7	2.4	1.3
HadGEM2-ES	3.5	2.9	1.6	–0.4	0.4	0.7	0.6	1.6	4.6	2.5
INM-CM4	3.1	3.0	1.6	–0.5	0.3	0.2	1.4	0.7	2.1	1.3
IPSL-CM5A-LR	3.2	3.1	1.9	–0.9	0.2	1.2	0.8	1.3	4.1	2.0

IPSL-CM5B-LR	n.a.	2.7	2.0	–1.0	0.2	1.2	1.0	1.0	2.6	1.5
MIROC-ESM	n.a.	4.3	1.8	–0.7	0.5	0.6	0.9	1.1	4.7	2.2
MIROC5	4.0	4.1	1.7	–0.5	0.5	-0.1	1.5	0.7	2.7	1.5
MPI-ESM-LR	4.3	4.1	1.7	–0.7	0.4	0.4	1.1	0.9	3.6	2.0
MPI-ESM-P	4.3	4.3	1.8	–0.8	0.4	0.3	1.3	0.8	3.5	2.0
MRI-CGCM3	3.6	3.2	1.5	–0.4	0.4	0.3	1.3	0.8	2.6	1.6
NorESM1-M	n.a.	3.1	1.5	–0.3	0.4	0.3	1.1	0.9	2.8	1.4
Model mean	3.7	3.4	1.7	–0.6	0.4	0.5	1.1	1.0	3.4	1.8
Standard deviation	0.5	0.5	0.2	0.2	0.1	0.4	0.3	0.3	0.8	0.4

Table 9.5: Features of Earth System Models of Intermediate Complexity (EMICs).

Model	Atmosphere[a]	Ocean[b]	Sea Ice[c]	Coupling[d]	Land Surface[e]	Biosphere[f]	Ice Sheets[g]	Sediment and Weathering[h]
Bern3D-LPJ (Ritz et al., 2011)	EMBM, 2-D(φ, λ), NCL, 10° x (3-19)°	FG with parameterized zonal pressure gradient, 3-D, RL, ISO, MESO, 10° x (3-19)°, L32 (Muller et al., 2006)	0-LT, DOC, 2-LIT	PM, NH, RW	Bern3D: 1-LST, NSM, RIV LPJ: 8-LST, CSM with uncoupled hydrology (Wania et al., 2009)	BO (Gangsto et al., 2011; Parekh et al., 2008; Tschumi et al., 2008), BT (Sitch et al., 2003; Stocker et al., 2011; Strassmann et al., 2008), BV (Sitch et al., 2003)	N/A	CS, SW (Tschumi et al., 2011)
CLIMBER-2.4 (Petoukhov et al., 2000)	SD, 3-D, CRAD, ICL, 10° x 51°, L10	FG, 2-D(φ,z), RL, 2.5°, L21 (Wright and Stocker, 1992)	1-LT, PD, 2-LIT (Petoukhov et al., 2000)	NM, NH, NW (Petoukhov et al., 2000)	1-LST, CSM, RIV (Petoukhov et al., 2000)	BO, BT, BV (Brovkin et al., 2002)	TM, 3-D, 0.75° x 1.5°, L20 (Calov et al., 2002)	N/A
CLIMBER-3α (Montoya et al., 2005)	SD, 3-D, CRAD, ICL, 7.5° x 22.5°, L10 (Petoukhov et al., 2000)	PE, 3-D, FS, ISO, MESO, TCS, DC, 3.75° x 3.75°, L24	2-LT, R, 2-LIT (Fichefet and Morales Maqueda, 1997)	AM, NH, RW	1-LST, CSM, RIV (Petoukhov et al., 2000)	BO (Six and Maier-Reimer, 1996), BT, BV (Brovkin et al., 2002)	N/A	N/A
DCESS (Shaffer et al., 2008)	EMBM, 2-box in φ, LRAD, CHEM (Shaffer et al., 2008)	2-box in φ, parameterized circulation and exchange, MESO, L55 (Shaffer and Sarmiento, 1995)	Parameterized from surface temperature (Shaffer et al., 2008)	NH, NW (Shaffer et al., 2008)	NST, NSM (Shaffer et al., 2008)	BO, BT (Shaffer et al., 2008)	N/A	CS, SW (Shaffer et al., 2008)
FAMOUS XDBUA (Smith et al., 2008a)	PE, 3-D, CRAD, ICL, 5° x 7.5°, L11 (Pope et al., 2000)	PE, 3-D, RL, ISO, MESO, 2.5° x 3.75°, L20 (Gordon et al., 2000)	0-LT, DOC, 2-LIT	NM, NH, NW	4-LST, CSM, RIV (Cox et al., 1999b)	BO (Palmer and Totterdell, 2001)	N/A	N/A
GENIE (Holden et al., 2010)	EMBM, 2-D(φ, λ), NCL, 10° x (3-19)° (Marsh et al., 2011)	FG, 3-D, RL, ISO, MESO, 10° x (3-19) °, L16 (Marsh et al., 2011)	1-LT, DOC, 2-LIT (Marsh et al., 2011)	PM, NH, RW (Marsh et al., 2011)	1-LST, BSM, RIV (Williamson et al., 2006)	BO, BT (Holden et al., 2010; Ridgwell et al., 2007b; Williamson et al., 2006)	N/A	CS, SW (Ridgwell and Hargreaves, 2007)
IAP RAS CM (Eliseev and Mokhov, 2011)	SD, 3-D, CRAD, ICL, 4.5° x 6°, L8 (Petoukhov et al., 1998)	PE, 3-D, RL, ISO, TCS, 3.5° x 3.5°, L32 (Muryshev et al., 2009)	0-LT, 2-LIT (Muryshev et al., 2009)	NM, NH, NW (Muryshev et al., 2009)	240-LST, CSM (Arzhanov et al., 2008)	BT (Eliseev and Mokhov, 2011)	N/A	N/A
IGSM 2.2 (Sokolov et al.,	SD, 2-D(φ, Z), ICL, CHEM,4° ×	Q-flux mixed-layer, anomaly diffusing,	2-LT, (Hansen et	Q-flux (Sokolov et	CSM (Oleson et al., 2008b)	BO (Holian et al., 2001), BT (Felzer et al., 2004;	N/A	N/A

Model	Atmosphere[a]	Ocean[b]	Sea Ice[c]	Coupling[d]	Land Surface[e]	Biosphere[f]	Ice Sheets[g]	Sediment and Weathering[h]
2005)	360° , L11 (Sokolov and Stone, 1998)	4°x5°, L11 (Hansen et al., 1984)	al., 1984)	al., 2005)		Liu, 1996; Melillo et al., 1993)		
LOVECLIM1.2 (Goosse et al., 2010)	QG, 3-D, LRAD, NCL, 5.6° x 5.6°, L3 (Opsteegh et al., 1998)	PE, 3-D, FS, ISO, MESO, TCS, DC, 3° x 3°, L30 (Goosse and Fichefet, 1999)	3-LT, R, 2-LIT (Fichefet and Maqueda, 1997)	NM, NH, RW (Goosse et al., 2010)	1-LST, BSM, RIV (Goosse et al., 2010)	BO (Mouchet and François, 1996), BT, BV (Brovkin et al., 2002)	TM, 3-D, 10 km x 10 km, L30 (Huybrechts, 2002)	N/A
MESMO 1.0 (Matsumoto et al., 2008)	EMBM, 2-D(φ, λ), NCL, 10° x (3-19)° (Fanning and Weaver, 1996)	FG, 3-D, RL, ISO, MESO, 10° x (3-19)°, L16 (Edwards and Marsh, 2005)	0-LT, DOC, 2-LIT (Edwards and Marsh, 2005)	PM, NH, RW	NST, NSM, RIV (Edwards and Marsh, 2005)	BO (Matsumoto et al., 2008)	N/A	N/A
MIROC-lite (Oka et al., 2011)	EMBM, 2-D(φ, λ), NCL, 4° x 4° (Oka et al., 2011)	PE, 3-D, FS, ISO, MESO, TCS, 4° x 4° (Hasumi, 2006)	0-LT, R, 2-LIT (Hasumi, 2006)	PM, NH, NW (Oka et al., 2011)	1-LST, BSM (Oka et al., 2011)	N/A	N/A	N/A
MIROC-lite-LCM (Tachiiri et al., 2010)	EMBM, 2-D(φ, λ), NCL, 6° x 6° (Oka et al., 2011), tuned for 3 K equilibrium climate sensitivity	PE, 3-D, FS, ISO, MESO, TCS, 6° x 6°, L15 (Hasumi, 2006)	0-LT, R, 2-LIT (Hasumi, 2006)	NM, NH (Oka et al., 2011), RW (Tachiiri et al., 2010)	1-LST, BSM (Oka et al., 2011)	BO (Palmer and Totterdell, 2001), loosely coupled BT (Ito and Oikawa, 2002)	N/A	N/A
UMD 2.0 (Zeng et al., 2004)	QG, 3-D, LRAD, ICL, 3.75° x 5.625°, L2 (Neelin and Zeng, 2000; Zeng et al., 2000)	Q-flux mixed-layer, 2-D surface (Hansen et al., 1983), deep ocean box model, 3.75° x 5.625°	N/A	Energy and water exchange only (Zeng et al., 2004)	2-LST with 2-layer soil moisture (Zeng et al., 2000)	BO (Archer et al., 2000), BT, BV (Zeng, 2003, 2006; Zeng et al., 2005)	N/A	N/A
UVic 2.9 (Weaver et al., 2001)	DEMBM, 2-D(φ, λ), NCL, 1.8° x 3.6° (Weaver et al., 2001)	PE, 3-D, RL, ISO, MESO, 1.8° x 3.6°, L19 (Weaver et al., 2001)	0-LT, R, 2-LIT (Weaver et al., 2001)	AM, NH, NW (Weaver et al., 2001)	1-LST, CSM, RIV (Meissner et al., 2003)	BO (Schmittner et al., 2005), BT, BV (Cox, 2001a)	TM, 3-D, 20 km x 20 km, L10 (Fyke et al., 2011)	CS, SW (Eby et al., 2009)

Notes:

(a) EMBM = energy moisture balance model; DEMBM = energy moisture balance model including some dynamics; SD = statistical-dynamical model; QG = quasi-geostrophic model; 2-D(φ, λ) = vertically averaged; 3-D = three-dimensional; LRAD = linearized radiation scheme; CRAD = comprehensive radiation scheme; NCL = non-interactive cloudiness; ICL = interactive cloudiness; CHEM = chemistry module; n° x m° = n degrees latitude by m degrees longitude horizontal resolution; Lp = p vertical levels.

(b) FG = frictional geostrophic model; PE = primitive equation model; 2-D(φ,z) = zonally averaged; 3-D = three-dimensional; RL = rigid lid; FS = free surface; ISO = isopycnal diffusion; MESO = parameterization of the effect of mesoscale eddies on tracer distribution; TCS = complex turbulence closure scheme; DC = parameterization of density-driven downward-sloping currents; n° x m° = n degrees latitude by m degrees longitude horizontal resolution; Lp = p vertical levels.
(c) n-LT = n-layer thermodynamic scheme; PD = prescribed drift; DOC = drift with oceanic currents; R = viscous-plastic or elastic-viscous-plastic rheology; 2-LIT = two-level ice thickness distribution (level ice and leads).
(d) PM = prescribed momentum flux; AM = momentum flux anomalies relative to the control run are computed and added to climatological data; NM = no momentum flux adjustment; NH = no heat flux adjustment; RW = regional freshwater flux adjustment; NW = no freshwater flux adjustment.
(e) NST = no explicit computation of soil temperature; n-LST = n-layer soil temperature scheme; NSM = no moisture storage in soil; BSM = bucket model for soil moisture; CSM = complex model for soil moisture; RIV = river routing scheme.
(f) BO = model of oceanic carbon dynamics; BT = model of terrestrial carbon dynamics; BV = dynamical vegetation model.
(g) TM = thermomechanical model; 3-D = three-dimensional; n° x m° = n degrees latitude by m degrees longitude horizontal resolution; n km x m km = horizontal resolution in kilometres; Lp = p vertical levels.
(h) CS = complex ocean sediment model; SW = simple, specified or diagnostic weathering model.

Chapter 9: Evaluation of Climate Models

Coordinating Lead Authors: Gregory Flato (Canada), Jochem Marotzke (Germany)

Lead Authors: Babatunde Abiodun (South Africa), Pascale Braconnot (France), Sin Chan Chou (Brazil), William Collins (USA), Peter Cox (UK), Fatima Driouech (Morocco), Seita Emori (Japan), Veronika Eyring (Germany), Chris Forest (USA), Peter Gleckler (USA), Eric Guilyardi (France), Christian Jakob (Australia), Vladimir Kattsov (Russia), Chris Reason (South Africa), Markku Rummukainen (Sweden)

Contributing Authors: Krishna AchutaRao, Alessandro Anav, Timothy Andrews, Johanna Baehr, Alejandro Bodas-Salcedo, Jennifer Catto, Ping Chang, Aiguo Dai, Clara Deser, Paul Durack, Michael Eby, Ramon de Elia, Thierry Fichefet, Bode Gbobaniyi, Clare Goodess, Fidel Gonzales-Rouco, Stephen Griffies, Alex Hall, Elizabeth Hunke, Tatiana Ilyina, Masa Kageyama, Viatcheslav Kharin, Stephen A. Klein, Jeff Knight, Reto Knutti, Felix Landerer, Tong Lee, Natalie Mahowald, Gerald Meehl, David Neelin, Tatiana Pavlova, Judith Perlwitz, Florian Rauser, Jouni Räisänen, Jeffrey Reid, Mark Rodwell, Adam A. Scaife, John Scinocca, David Sexton , Rym M'Sadek, Hideo Shiogama, Jana Sillmann, Ken Sperber, David Stephenson, Bjorn Stevens, Mark Webb, Keith Williams, Tim Woollings, Shang-Ping Xie, Jangling Zhang

Review Editors: Isaac Held (USA), Andy Pitman (Australia), Serge Planton (France), Zong-Ci Zhao (China)

Date of Draft: 5 October 2012

Notes: TSU Compiled Version

Figures

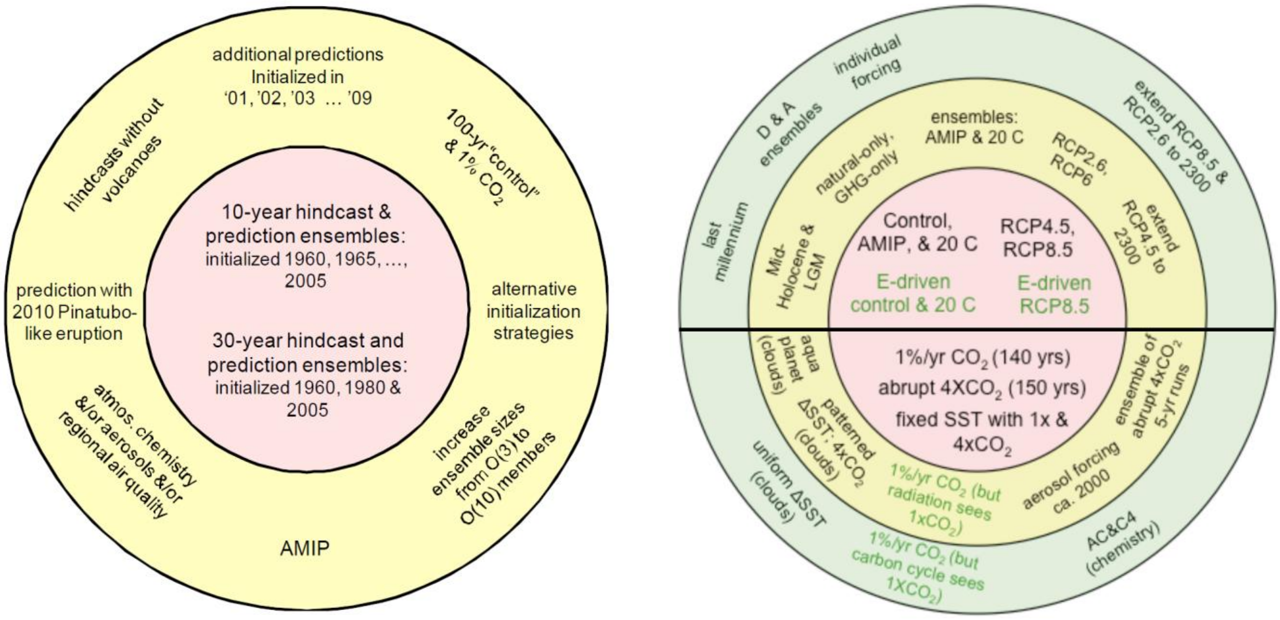


Figure 9.1: Left: Schematic summary of CMIP5 short-term experiments with tier 1 experiments (yellow background) organized around a central core (pink background). From (Taylor et al., 2012), their Figure 2. Right: Schematic summary of CMIP5 long-term experiments with tier 1 and tier 2 experiments organized around a central core. Green font indicates simulations to be performed only by models with carbon cycle representations, and "E-driven" means "emission-driven". Experiments in the upper hemisphere either are suitable for comparison with observations or provide projections, whereas those in the lower hemisphere are either idealized or diagnostic in nature, and aim to provide better understanding of the climate system and model behaviour. From (Taylor et al., 2012), their Figure 3.

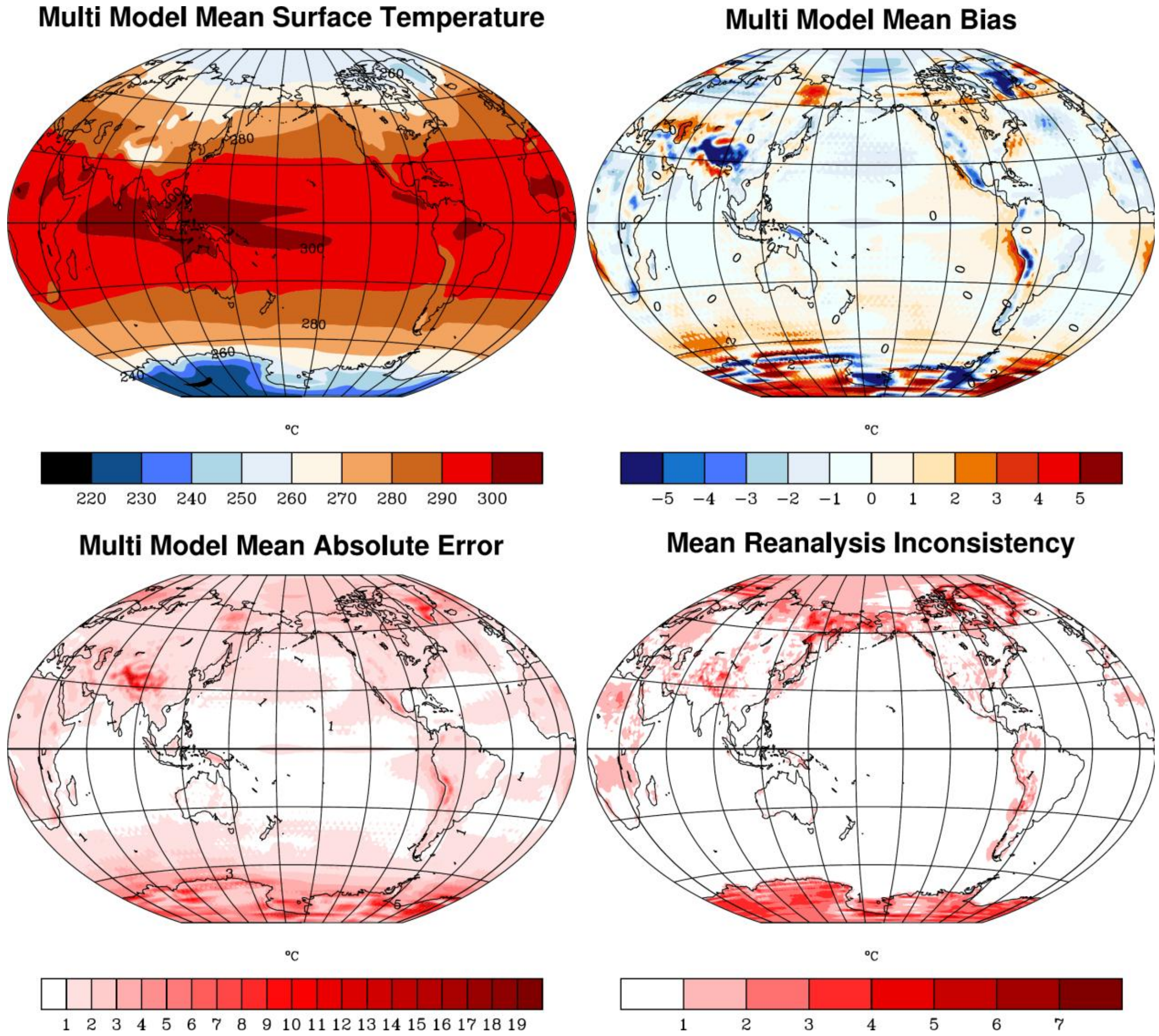


Figure 9.2: Annual mean surface (2 meter) air temperature (°C) for the period (1980–2005). Top left: Multi-model (ensemble) mean constructed with all available models used in the CMIP5 historical experiment. Top right: Multi-model mean bias as the difference between the CMIP5 multi-model mean and the climatology from ERA-Interim (1990–2005, (Dee et al., 2011)). Bottom left: Mean absolute model error with respect to the climatology from ERA-Interim. Bottom right: Mean inconsistency between three different reanalysis products as the mean of the absolute pair-wise differences between those fields.

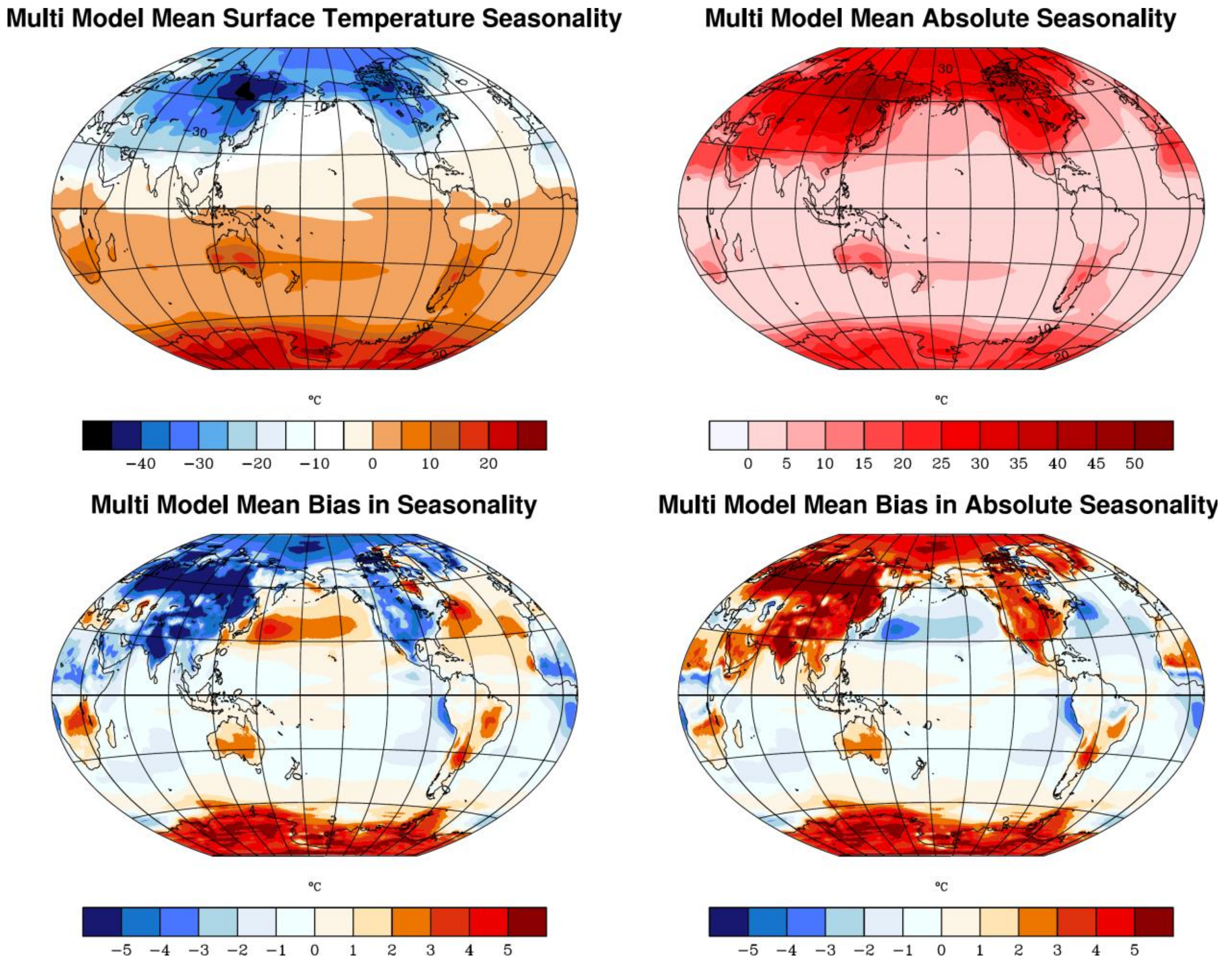


Figure 9.3: Annual surface (2 meter) air temperature (°C) range (DJF-JJA) for the period (1980–2005). Top left: multi-model mean of the DJF-JJA seasonality, calculated from all available CMIP5 models for the historical experiment. Top right: the mean absolute multi-model seasonality. Bottom left: the difference between the multi-model-mean and the ERA-Interim seasonality. Bottom right: the difference between the multi-model-mean and the ERA-Interim absolute seasonality.

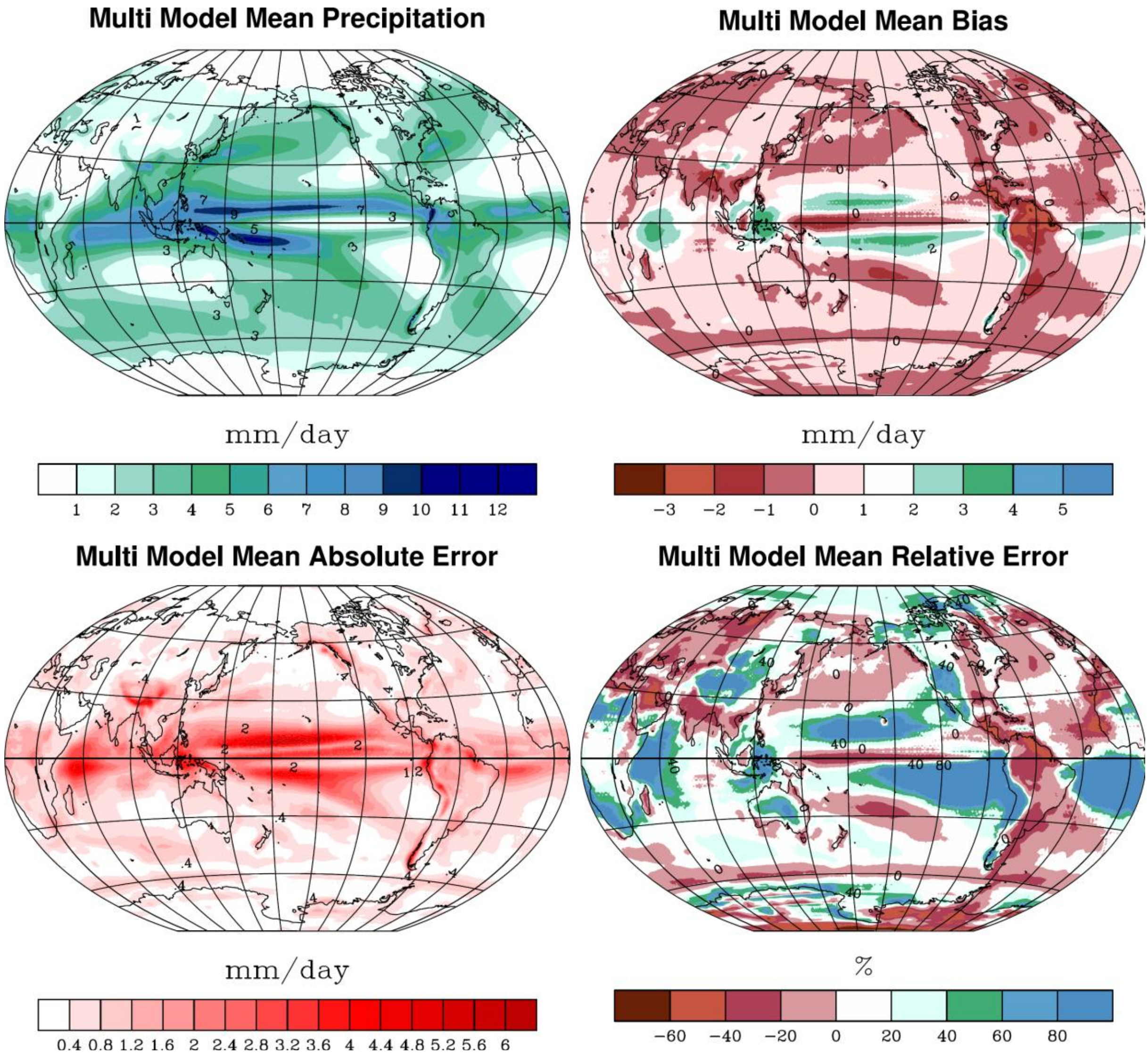


Figure 9.4: Annual mean precipitation for the period (1980–2005). Top left: Multi-model mean constructed with all available AOGCMs used in the CMIP5 historical experiment. Top right: difference between multi-model mean with and observations (Adler et al., 2003). Bottom left: Mean absolute model error with respect to observations. Bottom right: multi-model mean error relative to the multi model mean precipitation itself.

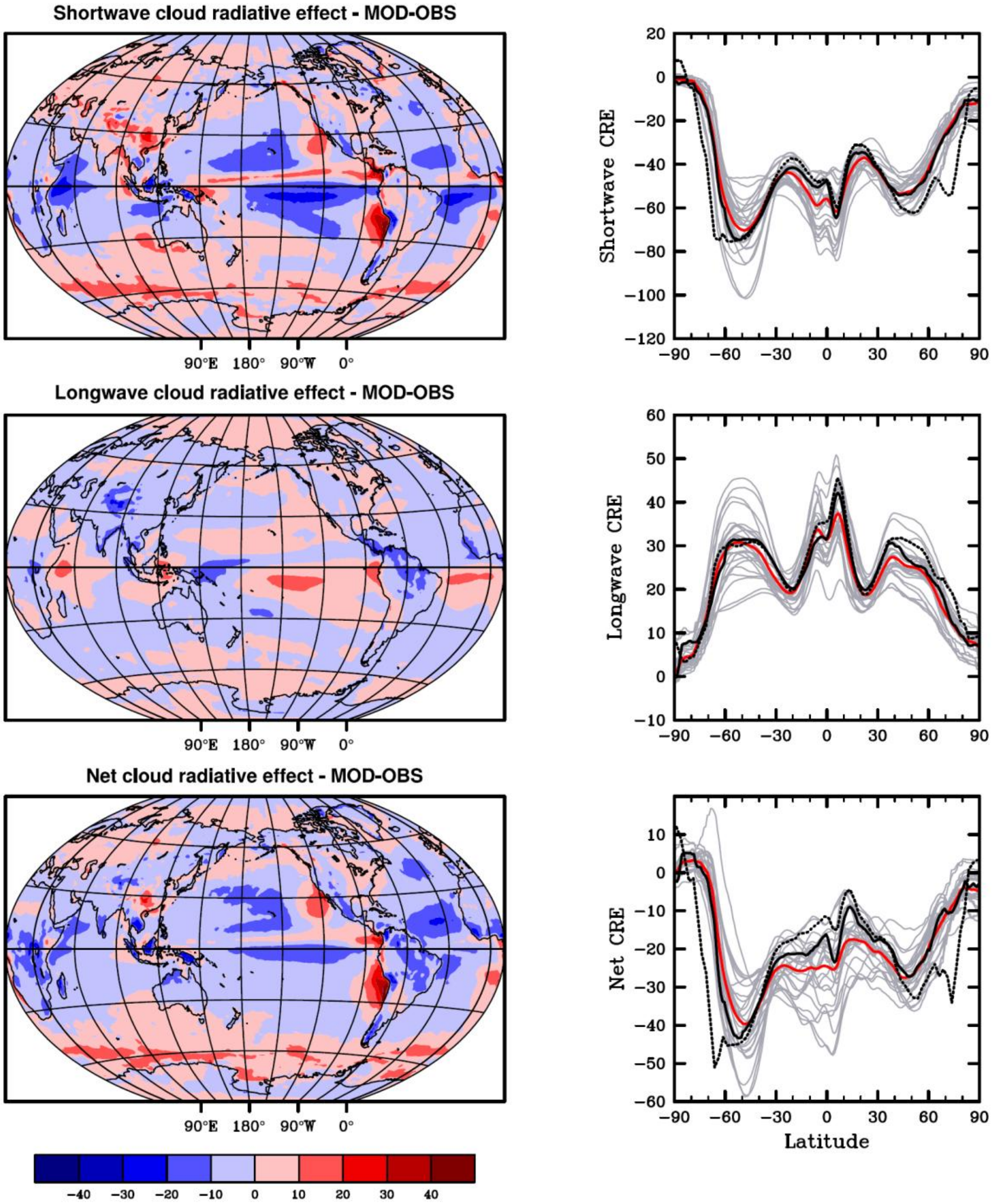


Figure 9.5: Annual mean errors in shortwave (top left), longwave (middle left) and net (bottom left) cloud radiative effect of the CMIP3 multi-model mean. Shown on the right are zonal averages of the absolute values of the same quantities from observations (solid black: CERES EBAF 2.6; dashed black: CERES ES-4), individual models (thin grey lines), and the multi-model mean (thick red line). For a definition of cloud radiative effect and maps of its absolute values, see Chapter 7.

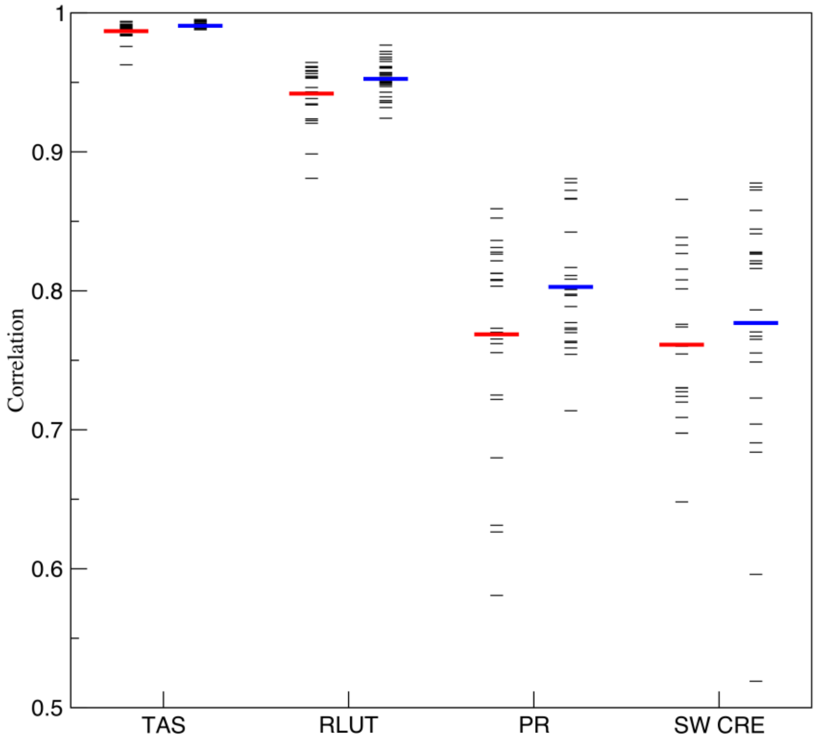


Figure 9.6: Global annual mean climatology (1980–1999) centred pattern correlations between models and observations. Results are shown for individual models from CMIP3 and CMIP5 (black dashes) and the average result for each (red for CMIP3, blue for CMIP5). The four variables shown are surface air temperature (TAS), top-of-atmosphere (TOA) outgoing longwave radiation (RLUT), precipitation (PR), and TOA shortwave cloud radiative effect (SW CRE). The observations used for each variable are the default products and climatological periods identified in Table 9.2. The centred pattern correlations are computed at a resolution of 5 degrees in longitude and latitude. Only one realization is used from each model from the CMIP3 20C3M and CMIP5 historical simulations.

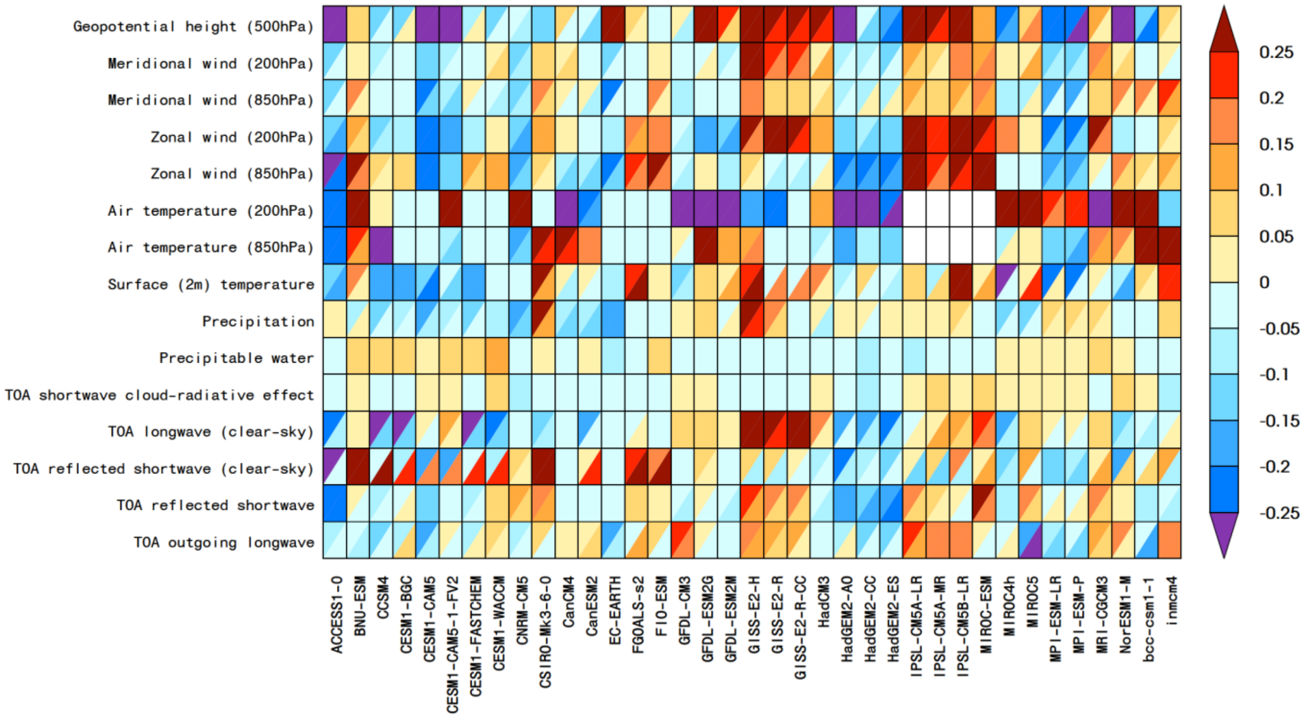


Figure 9.7: Relative error measures of CMIP5 model performance, based on the global annual cycle climatology (1980–2005) computed from the historical experiments. Rows and columns represent individual variables and models, respectively. The error measure is a global annual cycle space-time root-mean square error (RMSE), which, treating each variable separately, is portrayed as a relative error by normalizing the result by the median error of all model results (Gleckler at al., 2008). For example, a value of 0.20 indicates that a model's RMSE is 20% larger than the typical CMIP5 error for that variable, whereas a value of –0.20 means the error is 20% smaller than the typical error. No colour (white) denotes that data are currently unavailable. A diagonal splits each grid square, showing the relative error with respect to both the default (upper left triangle) and the alternate (lower right triangle) reference data sets. The relative errors are calculated independently for the default and alternate data sets. All reference data used in the diagram are summarized in Table 9.2.

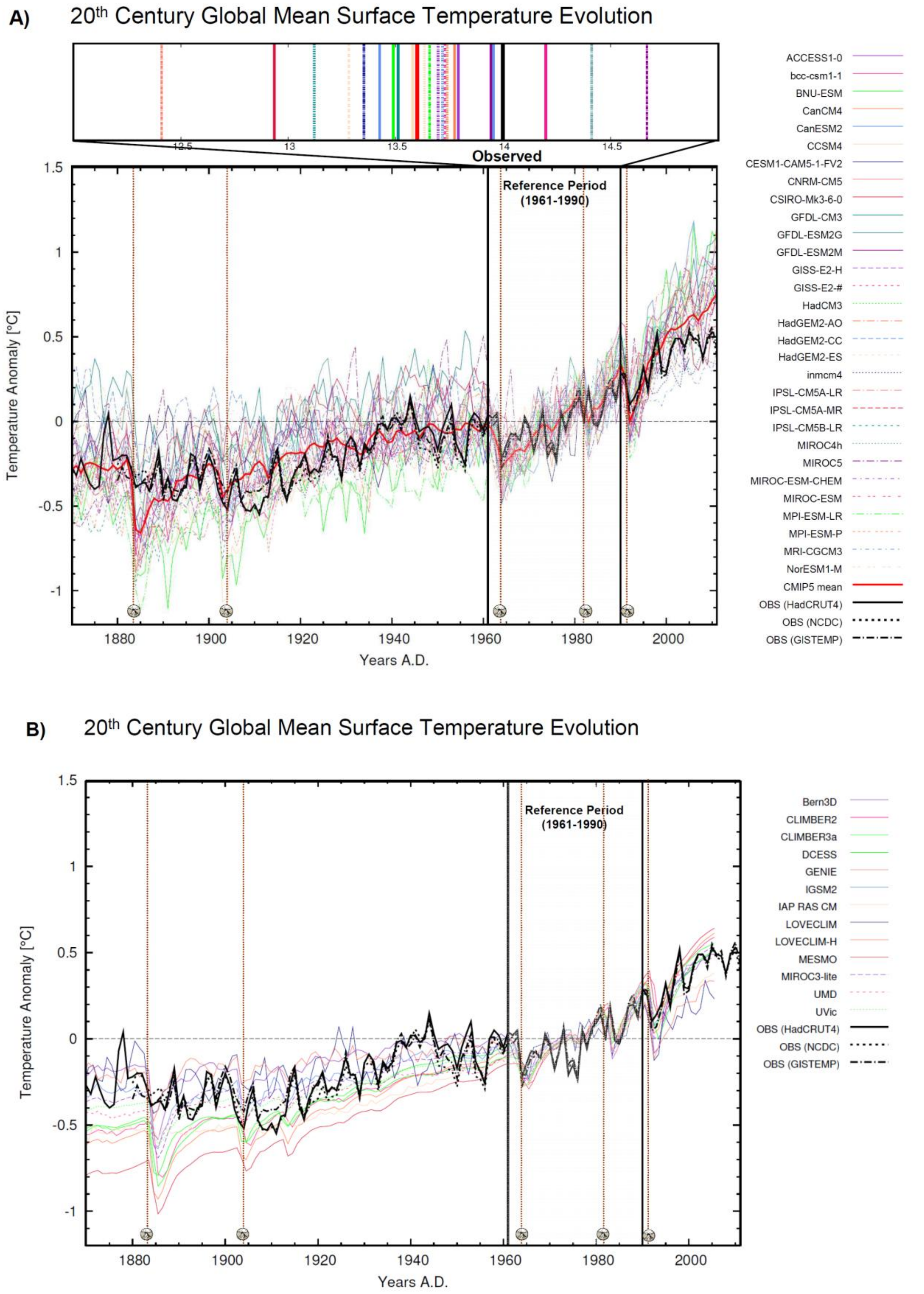


Figure 9.8: Observed and simulated annual mean global average anomaly time series of surface air temperature. All anomalies are differences from the 1961–1990 time-mean of each individual time series. Top: single simulations currently available for CMIP5 (thin lines); multi-model mean (thick red line); different observations (thick black lines).

Vertical dotted brown lines represent times of major volcanic eruptions. Observational data are HadCRUT4 (Morice et al., 2012), GISTEMP (Hansen et al., 2010), and NCDC (Smith et al., 2008b) and are merged surface temperature (2 m height over land and surface temperature over the ocean). Top, inset: the absolute global mean surface temperature for the reference period 1961–1990, for each individual model (colours) and the observations (black, (Jones et al., 1999). Bottom: single simulations from a variety of EMICs (thin lines). Vertical dotted brown bars represent times of major volcanic eruptions. Observational data are the same as for the top panel.

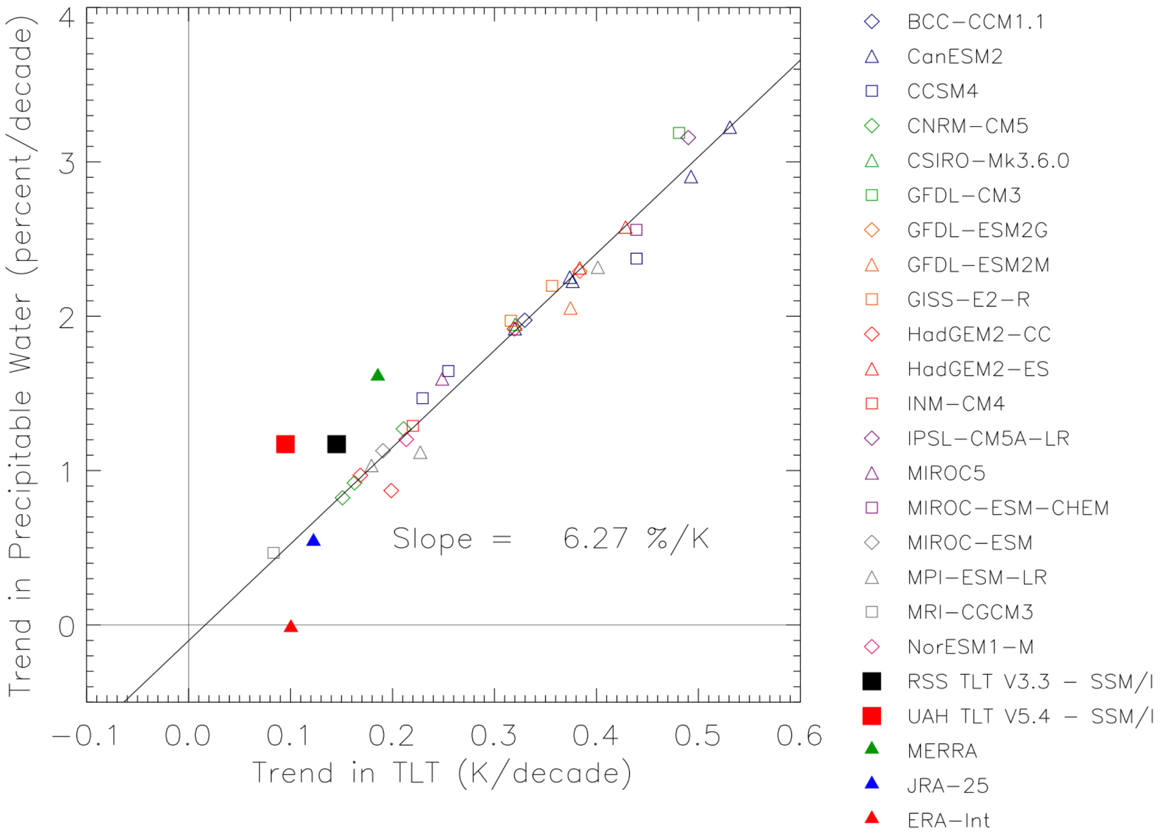


Figure 9.9: Scatter plot of decadal trends in tropical (20ºS to 20ºN) precipitable water as a function of trends in lower tropospheric temperature (TLT) over the world's oceans. Open symbols are from 19 CMIP5 models, filled squares are from satellite observations, and filled triangles are from reanalysis output. Trends are calculated over the 1988-2011 period, so CMIP5 historical runs, which typically end in December 2005, were extended using RCP8.5 simulations initialized using these historical runs. Figure updated from (Mears et al., 2007).

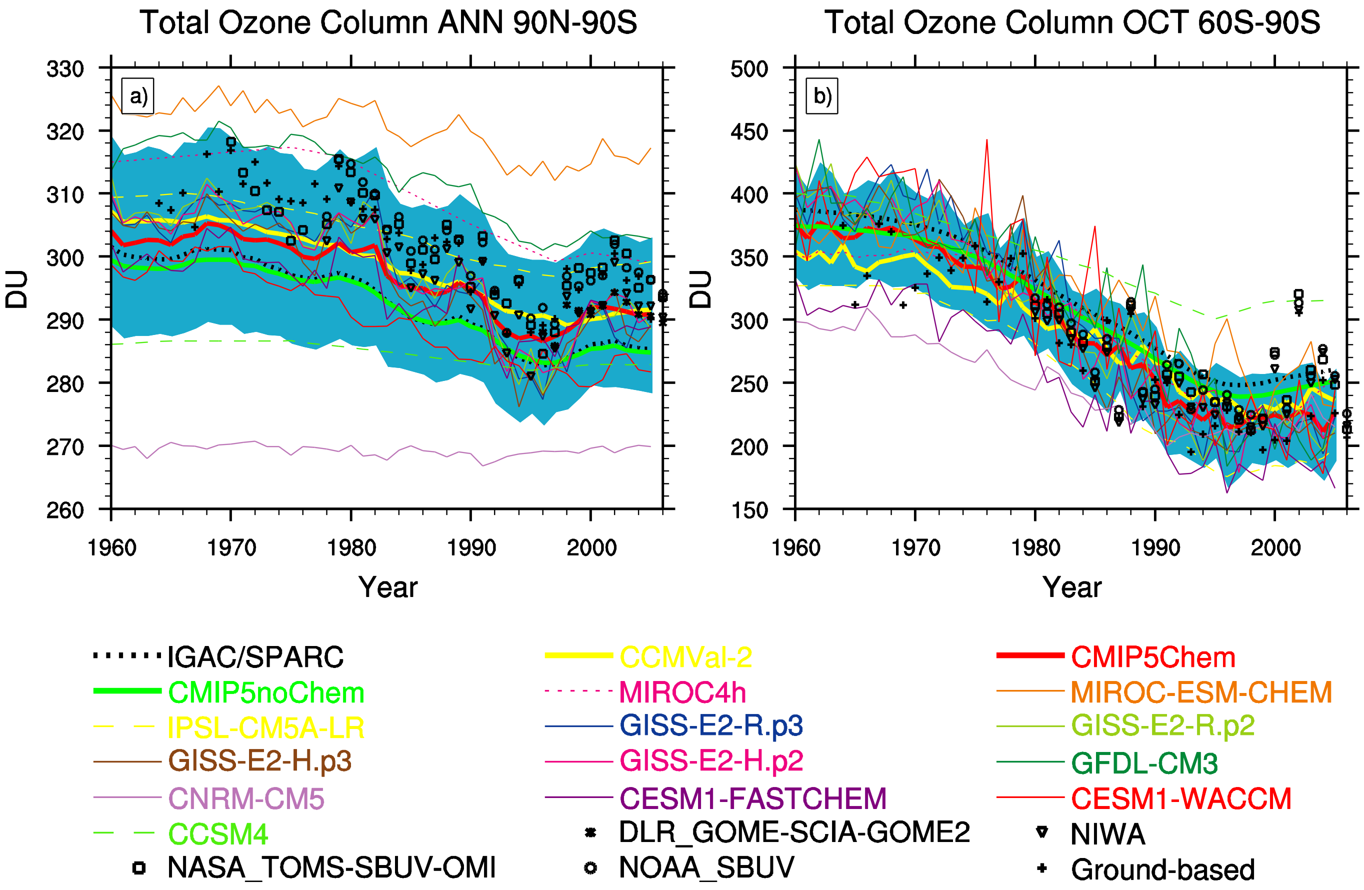


Figure 9.10: Time series of area-weighted total column ozone from 1960 to 2005 for (a) annual mean global mean (90°S–90°N) and (b) Antarctic October mean (60°S–90°S). The multi-model mean and individual CMIP5 models with interactive or semi-offline chemistry (CHEM, red solid and coloured lines) and standard deviation (blue shaded area) are compared to the multi-model mean of the CMIP5 models that prescribe ozone (NOCHEM, green solid line), the IGAC/SPARC ozone database (black dotted line), the CCMVal-2 multi-model mean (yellow solid line) and observations from five different sources (symbols). The observations include ground-based measurements (updated from Fioletov et al. (2002)), NASA TOMS/OMI/SBUV(/2) merged satellite data (Stolarski and Frith, 2006), the NIWA combined total column ozone database (Bodeker et al., 2005), Solar Backscatter Ultraviolet (SBUV, SBUV/2) retrievals (updated from Miller et al. (2002)), and DLR GOME/SCIA/GOME-2 (Loyola and Coldewey-Egbers, 2012; Loyola et al., 2009). Ozone depletion increased after 1960 as equivalent stratospheric chlorine values steadily increased throughout the stratosphere. Adapted from Figure 3 of Eyring et al. (2012b).

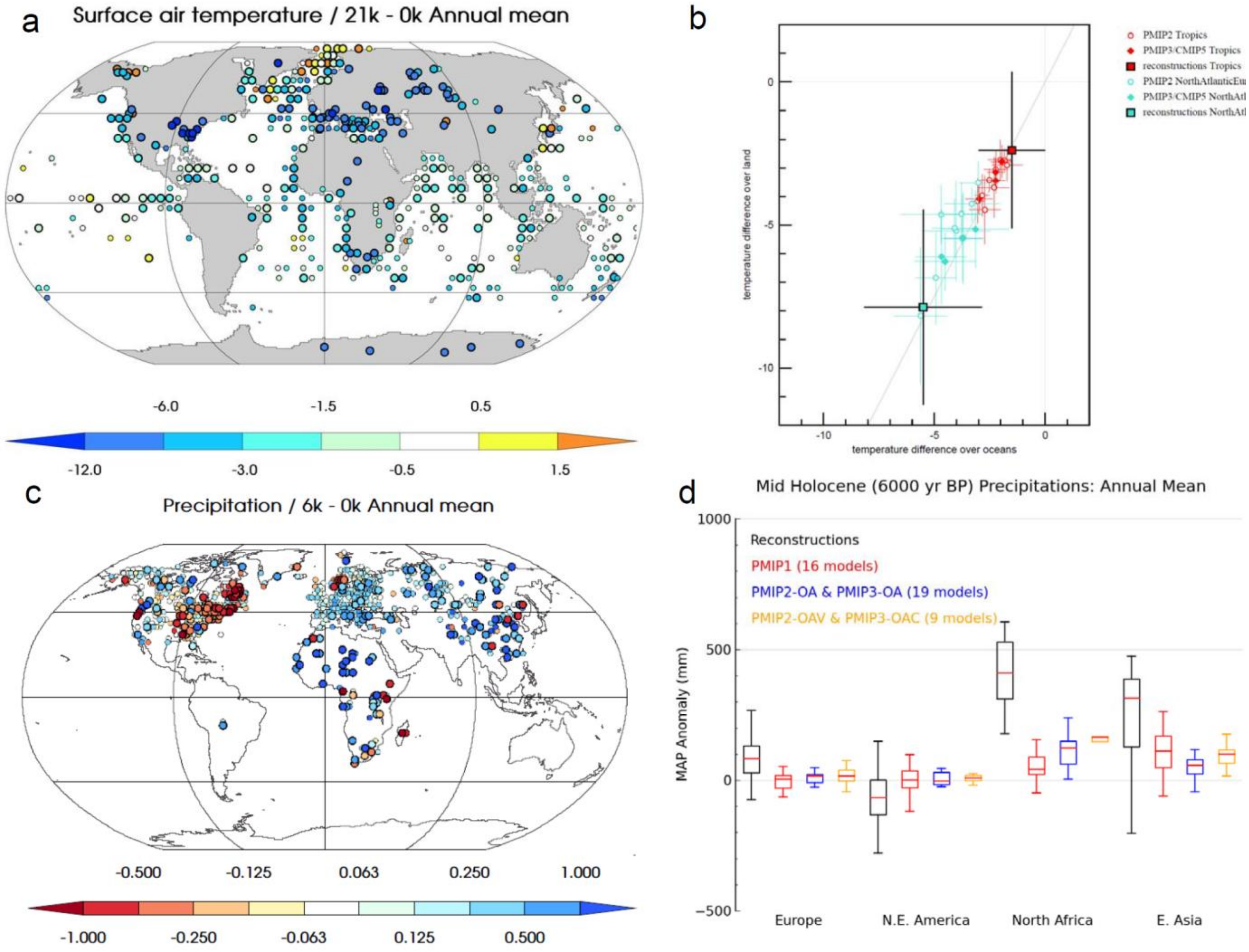


Figure 9.11: Ability of climate models to reproduce surface temperature during the Last Glacial Maximum (LGM, 21 ka BP, top) and precipitation during the mid-Holocene (6 ka BP, bottom), as shown by palaeo-environmental climate reconstructions from pollen and macrofossils over land (Bartlein et al., 2010b) and ice cores in a and c, and from different type of marine records for the 21 ka ocean (Waelbroeck et al., 2009) in a. In a and c, the size of the dots is proportional to the uncertainties at the different sites as provided in the reconstructions. Panel b provides a synthetic view of the ability of climate models to reproduce the relationship between changes in annual-mean temperature over ocean and land in the tropics (red) and changes in annual-mean temperature over the North Atlantic and in Europe (cyan). The squares show the mean value of the reconstructions, with the range shown in black. The empty symbols are the results of individual coupled ocean–atmosphere general circulation models from the second phase of the Paleoclimate Modeling Intercomparison Project (PMIP2) (Braconnot et al., 2007c), and the filled symbols are the results from CMIP5. Panel d shows how model skill in reproducing changes in annual-mean precipitation in different data rich regions has evolved during the different phases of PMIP. Box plots for reconstructions provide estimates of the scatter between different sites, whereas for models they represent the spread of model results. The limits of the boxes are as follows: W Europe (40°N–50°N, 10°W–30°E); North East America (35°N–60°N, 95°W–60°W); North Africa (10°N–25°N, 20°W–30°) and E Asia (25°N–40°N, 75°E–105°E). Adapted from Braconnot et al. (2012).

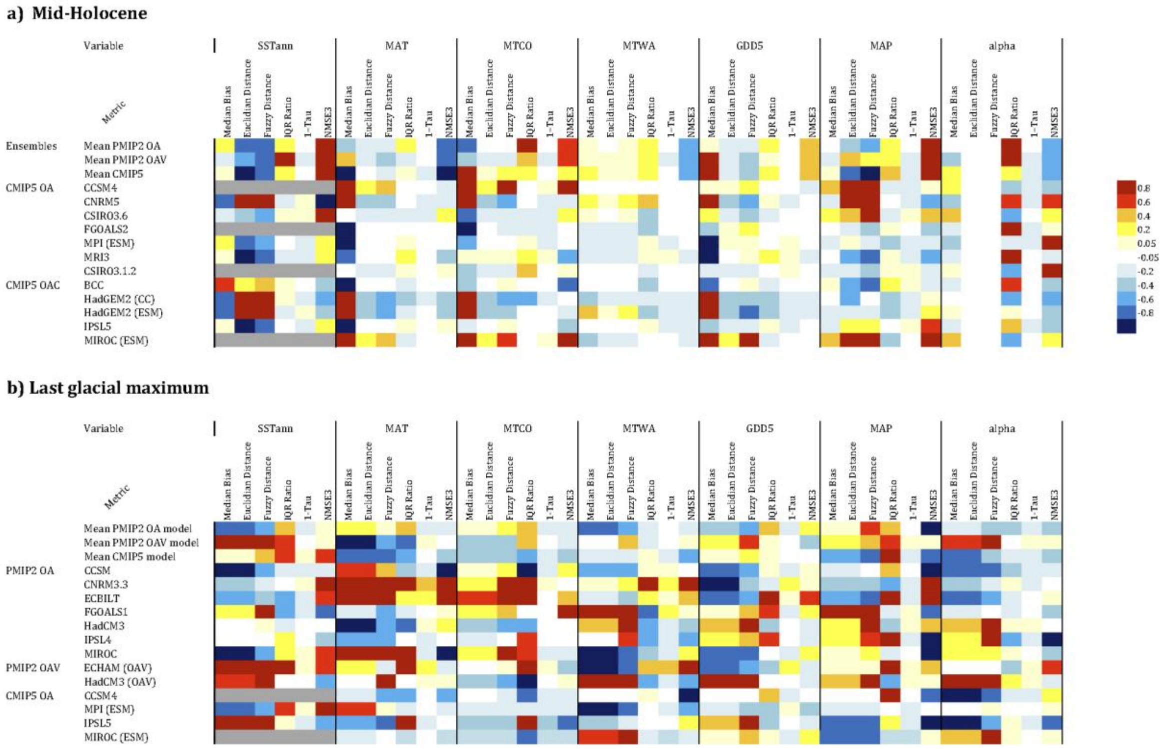


Figure 9.12: Summary of benchmark metrics for the Last Glacial Maximum (LGM, ca 21,000 yr BP) and the mid-Holocene (MH, ca 6000 yr BP). For each variable, six different metrics are considered. The median and the inter-quartile range (IQR) characterize the global distribution of the values considering only grid cells where there are observations. The agreement between the simulated and reconstruction maps is then characterised by the difference between the median values (median bias) and the ratio of the simulated and reconstructed IQR (IQR ratio). Other metrics consider the Euclidian distance between the simulated and reconstructed maps and a variant based on fuzzy logic that takes into account an estimate of simulated and reconstructed field uncertainties. The Kendall rank correlation (1-Tau) measures the similarities or differences in the spatial patterns without regard for magnitude and is used as an alternative to the Normalized Mean Square Error (NMSE) based on standardised variables. In this graph all the values have been normalised following Gleckler et al. (2008), so as to highlight model spread as in Figure 9.7. For the mid-Holocene we only plotted the values for all the CMIP5 simulations and consider only the PMIP2 results as ensembles to show how the performances of the model evolved. For the LGM there are fewer simulations and results of both PMIP2 and CMIP5 simulations are included. Adapted from Harrison et al. (2012).

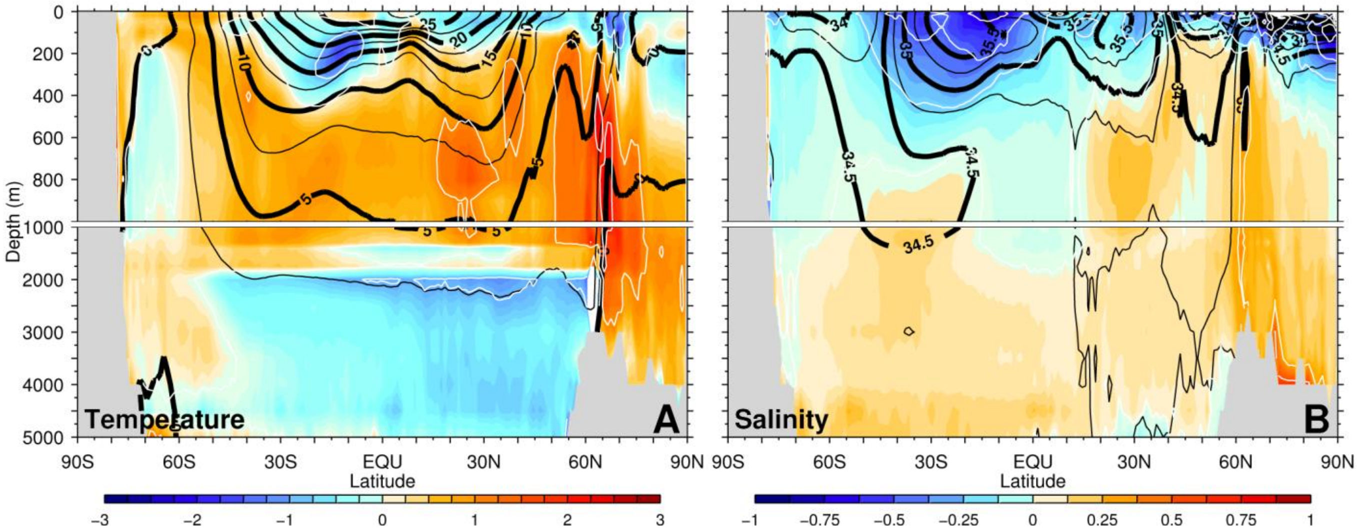


Figure 9.13: Time-mean differences between CMIP5 multi-model ensemble-mean and observed (A) potential temperature (°C) and (B) salinity (PSS-78) (colour). The observed climatological values (Antonov et al., 2010; Levitus et al., 2009) are zonally averaged for the global ocean (excluding marginal and regional seas) and are shown as labelled black contours. The simulations cover the period 1975 to 2005 from available historical simulations, whereas the observations are from 1874 to 2008. Multiple realizations from individual models are first averaged to form a single model climatology, before the construction of the multi-model ensemble mean. 21 available CMIP5 models have contributed to the temperature panel (A) and 20 models to the salinity panel (B).

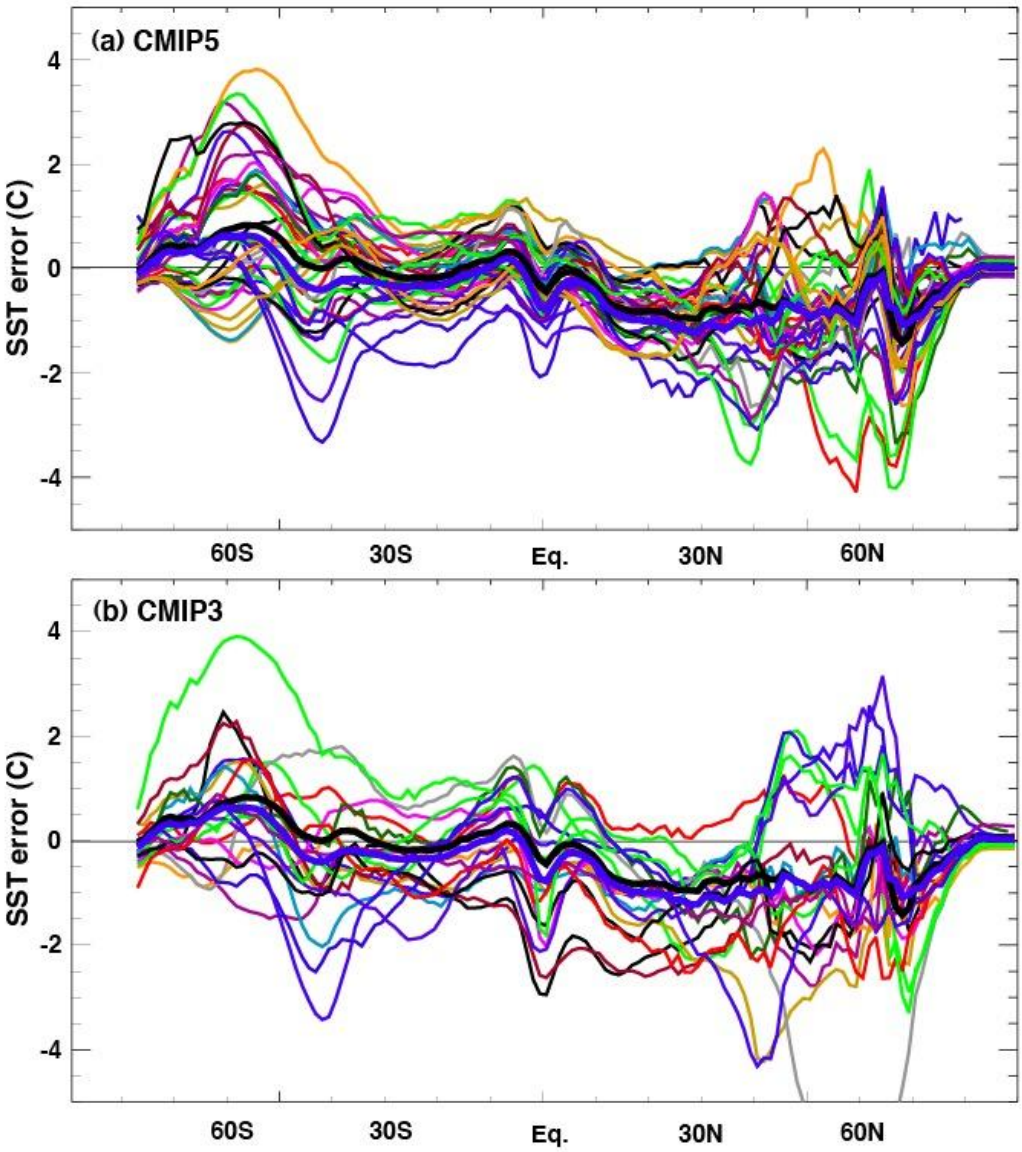


Figure 9.14: Annual mean, zonally averaged SST error, simulated minus observed climatology in (a) CMIP5 and (b) CMIP3. The Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) (Rayner et al., 2003a) observational climatology for 1850 to 2010 is the reference used here, and the model results are for the same period in the historical simulations. The multi-model mean of CMIP5 and CMIP3 are shown in both panels (bold black and blue respectively). [PLACEHOLDER FOR FINAL DRAFT: legend that identifies the individual models to be included.]

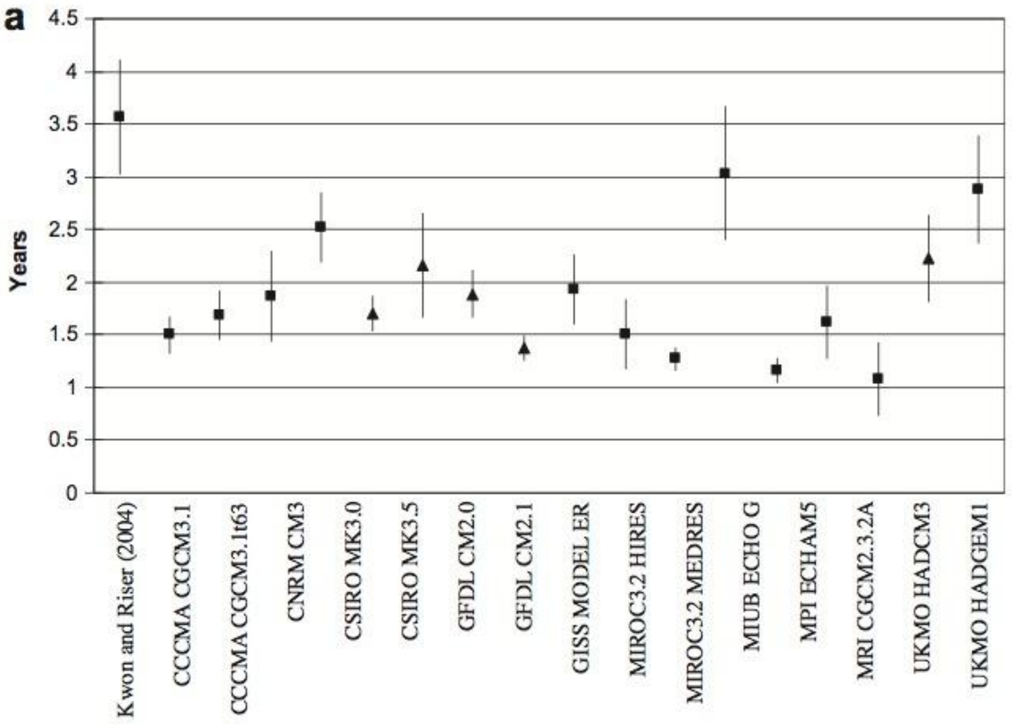


Figure 9.15: Sub-Tropical Mode Water (STMW) turnover time for various models compared with (Kwon and Riser, 2004); time is calculated by annual maximum volume divided by annual production. Values are means; error bars give ranges of one standard deviation. Square data symbols indicate those models with a distinct (if small) secondary water mass transformation rate peak corresponding to STMW formation. Triangular data symbols indicate those models with broad, diffuse, or indiscernible STMW formation peak (from McClean and Carman, 2011).

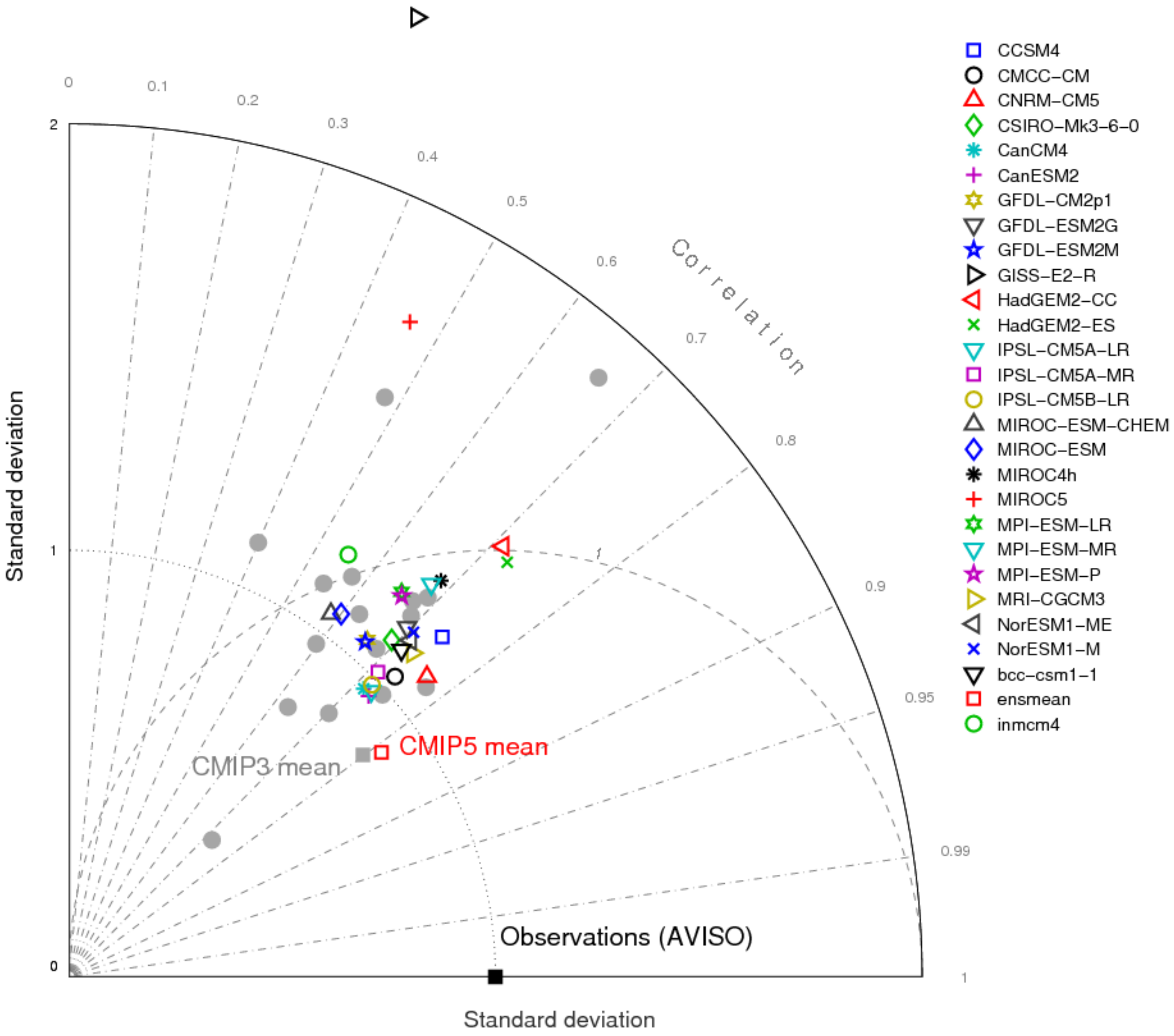


Figure 9.16: Taylor diagram of the dynamic sea-level height seasonal cycle climatology (1987–2000). The radial coordinate shows the standard deviation of the spatial pattern, normalised by the observed standard deviation. The azimuthal variable shows the correlation of the modelled spatial pattern with the observed spatial pattern. The root-mean square error is indicated by the dashed grey circles about the observational point. Analysis is for the global ocean, 50°S–50°N. The reference dataset is AVISO, a merged satellite product (Ducet et al., 2000), which is described in Chapter 3. Figure currently shows results for the CMIP3 models and the CMIP5 data currently available.

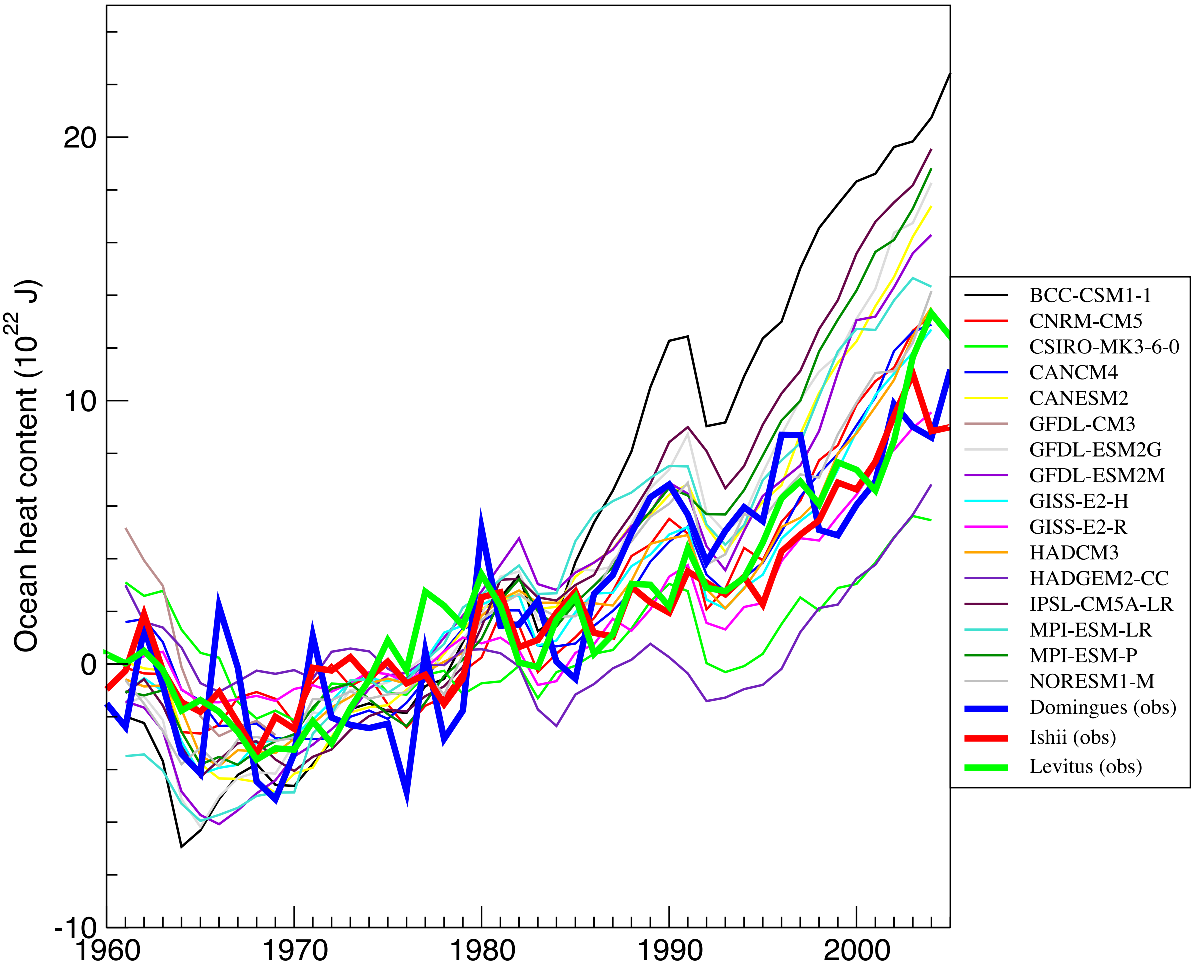


Figure 9.17: Time series of observed and simulated (CMIP5 historical) global 0–700 m ocean heat content anomalies (with respect to 1957–1990 climatologies), with units of 10^{22} Joules. The three observational estimates (thick lines) are discussed in Chapter 3. Individual simulations (one per model) are shown for the historical period. [PLACEHOLDER FOR FINAL DRAFT: to be updated; model simulations will be corrected for simulation drift, which for some models may be significant (Sen Gupta et al., 2012).]

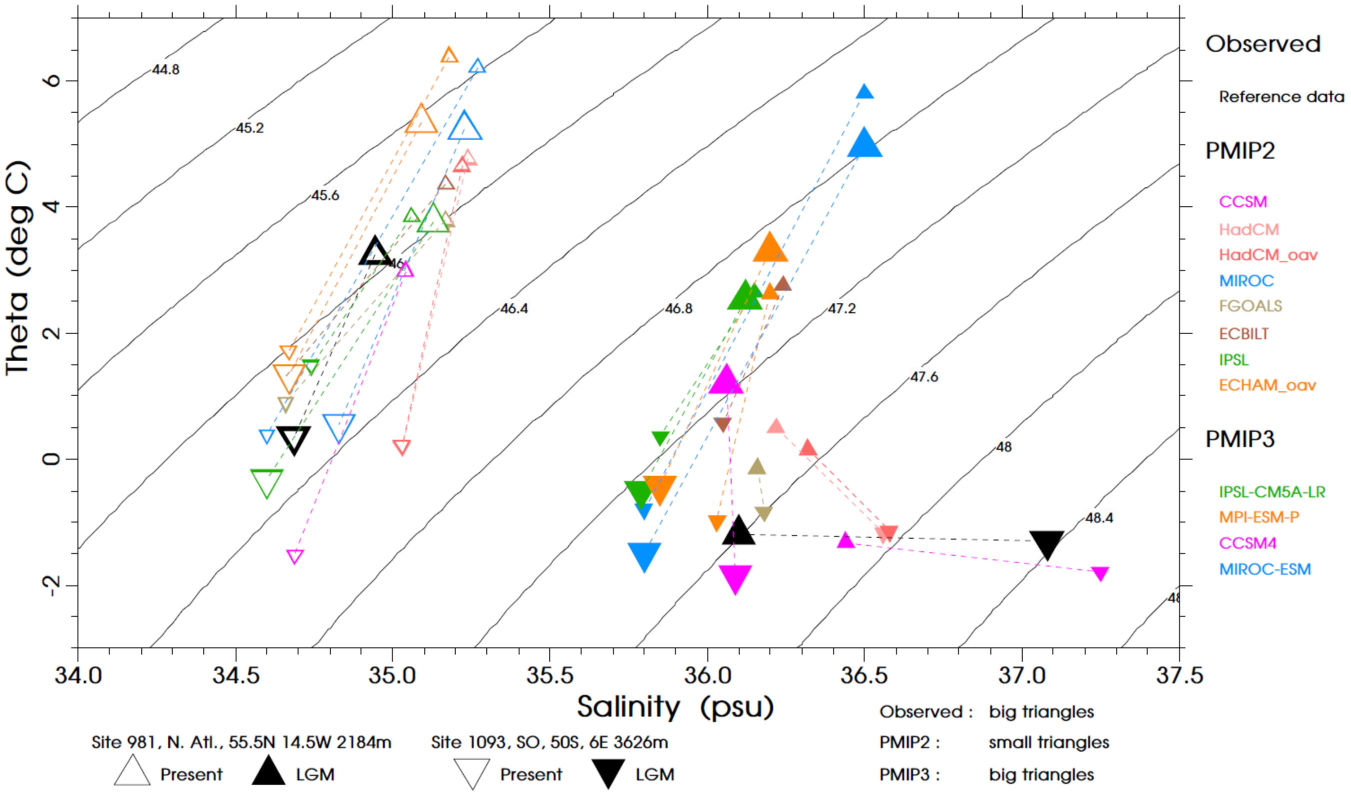
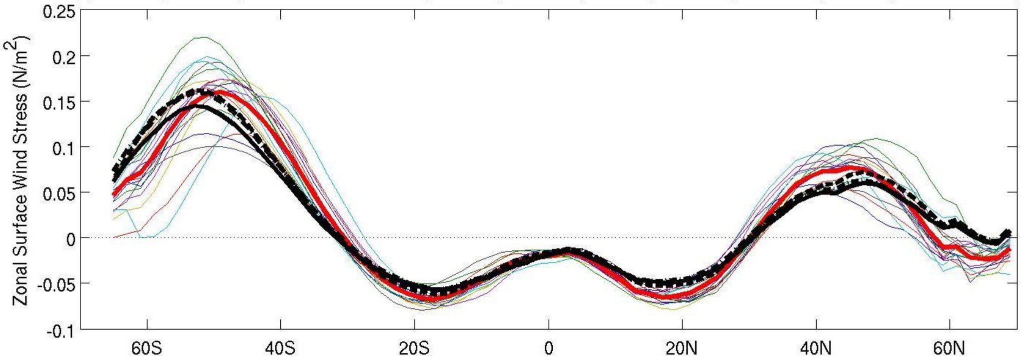
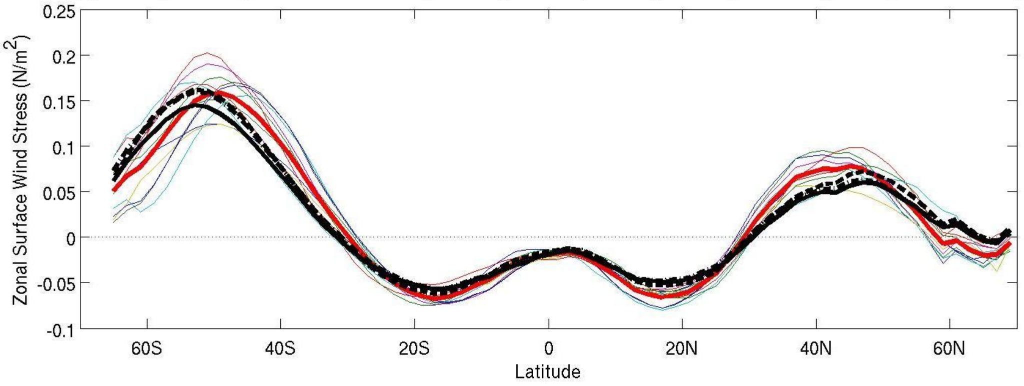


Figure 9.18: Temperature and salinity for the modern period (open symbols) and the Last Glacial Maximum (LGM, filled symbols) as estimated from data (black symbols) at ODP sites (Adkins et al., 2002) and predicted by the PMIP2 (small triangles) and PMIP3/CMIP5 (big triangles) simulations. Site 981 (triangles) is located in the North Atlantic (Feni Drift, 55'N, 15'W, 2184 m). Site 1093 (upside-down triangles) is located in the South Atlantic (Shona Rise, 50'S, 6'E, 3626 m). In PMIP2 only CCSM included a 1 psu adjustment of ocean salinity at initialization to account for freshwater frozen into LGM ice sheets; the other PMIP2 model-predicted salinities have been adjusted to allow a comparison. In PMIP3 all simulations include the 1 psu adjustment as required in the PMIP2/CMIP5 protocol (Braconnot et al., 2012), T he dotted lines allow a comparison of the values at the northern and the southern sites for a same model. This figure is adapted from Otto-Bliesner et al. (2007b) and shows quantitatively how deep-ocean properties can be evaluated for both modern and palaeo climates. In particular, models have difficulties to reproduce the cold and salty water found at depth at the LGM.



(a) CMIP3 (thin color), their ensemble mean (thick red), QSCAT (blk solid), NCPE1 (blk dahsed), ERA-I (blk dash-dotted)



(b) CMIP5 (thin color), their ensemble mean (thick red), QSCAT (blk solid), NCPE1 (blk dahsed), ERA-I (blk dash-dotted)

Figure 9.19: Zonal mean zonal wind stress over the oceans in CMIP3 and CMIP5 simulations during 1970–1999. Individual simulations: thin coloured; ensemble mean: thick red. The black solid, dashed, and dash-dotted curves represent QuikSCAT satellite measurements (Risien and Chelton, 2008), NCEP/NCAR reanalysis I (Kalnay et al., 1996), and ERA-Interim (Dee et al., 2011), respectively. The figure shows that CMIP3 & CMIP5 mid-latitude westerly winds are too strong and that circumpolar westerly wind are too weak compared to QuikSCAT, NCEP1, and ERA-Interim. CMIP3 and CMIP5 ensemble means are very similar.

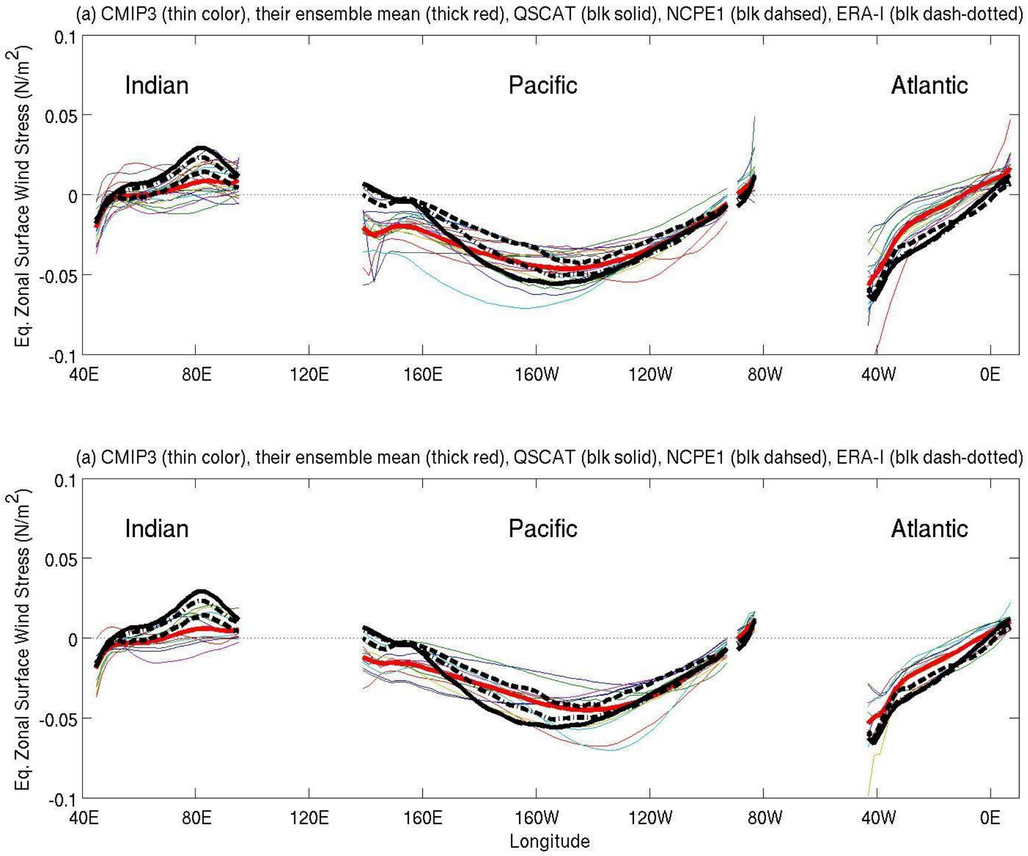


Figure 9.20: Equatorial (2ºS–2ºN averaged) zonal wind stress for the Indian, Pacific, and Atlantic Oceans; comparison of simulations by CMIP3 (a) and CMIP5 (b) with QuikSCAT observations. Individual simulations: thin coloured; ensemble mean: thick red. The black solid, dashed, and dash-dotted curves represent QuikSCAT satellite measurements (Risien and Chelton, 2008), NCEP/NCAR reanalysis I (Kalnay et al., 1996), and ERA-Interim (Dee et al., 2011), respectively. The figure shows that in both CMIP3 and CMIP5, the ensemble-mean equatorial zonal wind stress is too weak in the Atlantic and Indian Oceans and too strong in the western Pacific. CMIP3 and CMIP5 ensemble means are very similar.

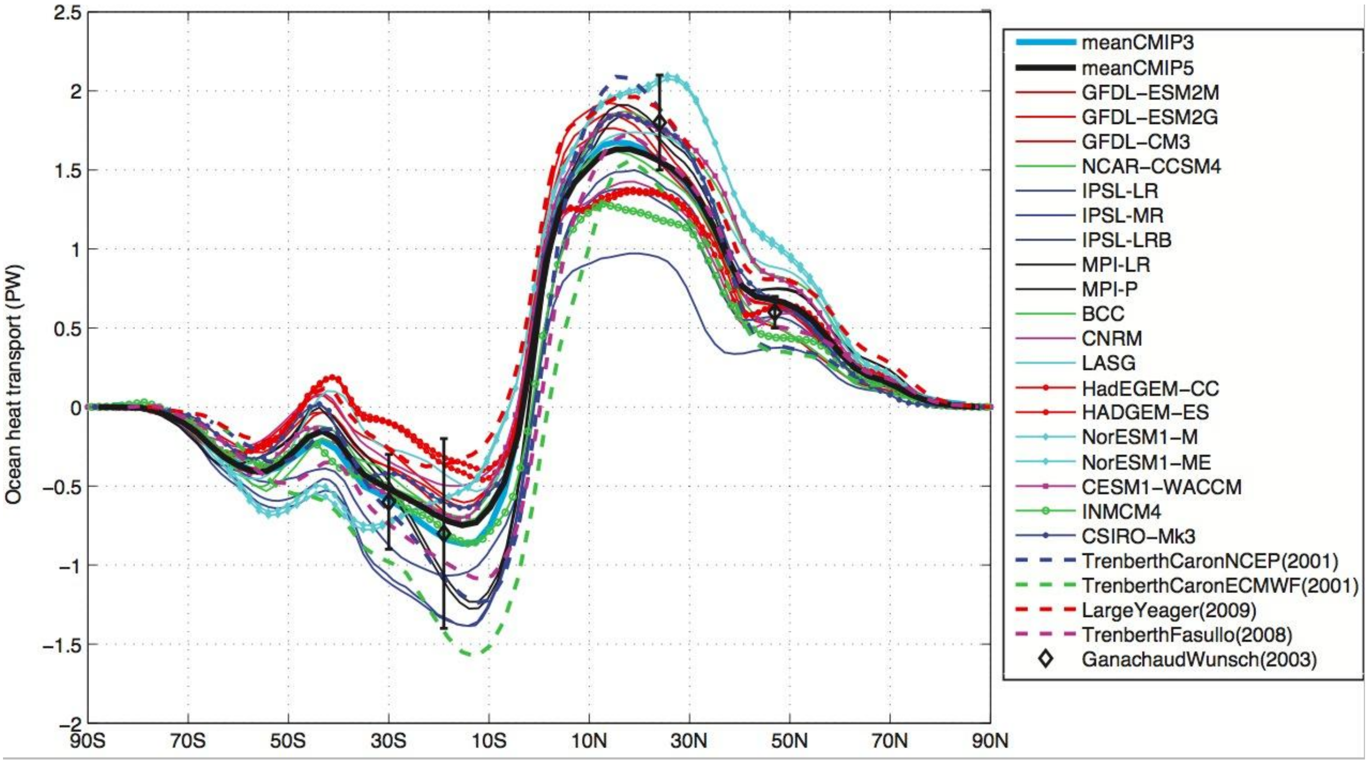


Figure 9.21: Annual mean, zonally averaged oceanic heat transport implied by net heat flux imbalances at the sea surface for CMIP5 simulations, under an assumption of negligible changes in oceanic heat content. Observational estimates include: (blue and green dashed lines) the dataset from Trenberth and Caron (2001) for the period February 1985 to April 1989, derived from reanalysis products from the National Centers for Environmental Prediction (NCEP/NCAR, Kalnay et al., 1996) and European Centre for Medium Range Weather Forecasts 40-year reanalysis (ERA40, Uppala et al., 2005), (purple dashed line) an updated version by Trenberth and Fasullo (2008) with improved TOA radiation data from the Clouds and Earth's Radiant Energy System (CERES) for March 2000 to May 2004, and updated NCEP reanalysis (Kistler et al., 2001) up to 2006, (red dashed line) the Large and Yeager (2009) analysis based on the range of annual mean transport estimated over the years 1984–2006, computed from air-sea surface fluxes adjusted to agree in the mean with a variety of satellite and in situ measurements, and (black diamonds) direct estimates by Ganachaud and Wunsch (2003) obtained from hydrographic sections during the World Ocean Circulation Experiment combined with inverse models. The model climatologies are derived from the years 1986 to 2005 in the historical simulations in CMIP5. The multi-model mean is shown as a thick black line. The CMIP3 multi-model mean is added as a cyan blue line. Note: climate models should feature a vanishing net energy balance when long time averages are considered but unphysical sources and sinks lead to energy biases (Trenberth and Fasullo, 2009; Trenberth and Fasullo, 2010a; Lucarini and Ragone, 2011) that are also found in reanalysis constrained by observations (Trenberth et al., 2009). When correcting for the imperfect closure of the energy cycle, as done here, comparison between models and observational estimates become possible.

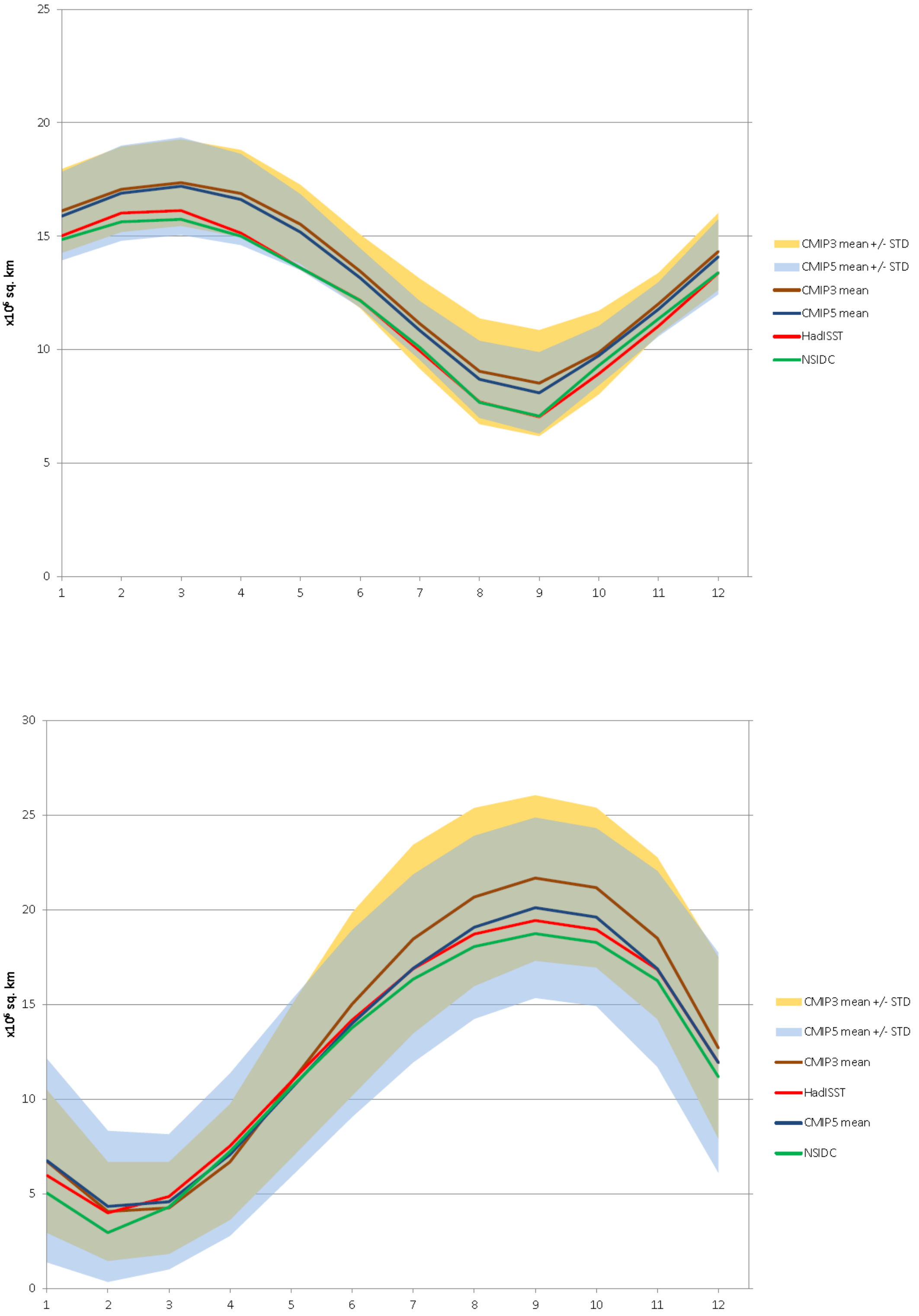


Figure 9.22: Mean sea ice extent (the ocean area with a sea ice concentration of at least 15%) seasonal cycle in the Northern (upper panel) and Southern (lower panel) Hemispheres as simulated by 40 CMIP5 and 17 CMIP3 models. The observed sea-ice extent cycles (1980–1999) are based on the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST; Rayner et al., 2003a) and the National Snow and Ice Data Center (NSIDC; Fetterer et al., 2002) data sets. The shaded areas show the inter-model standard deviation for each ensemble. Adapted from Pavlova et al. (2011).

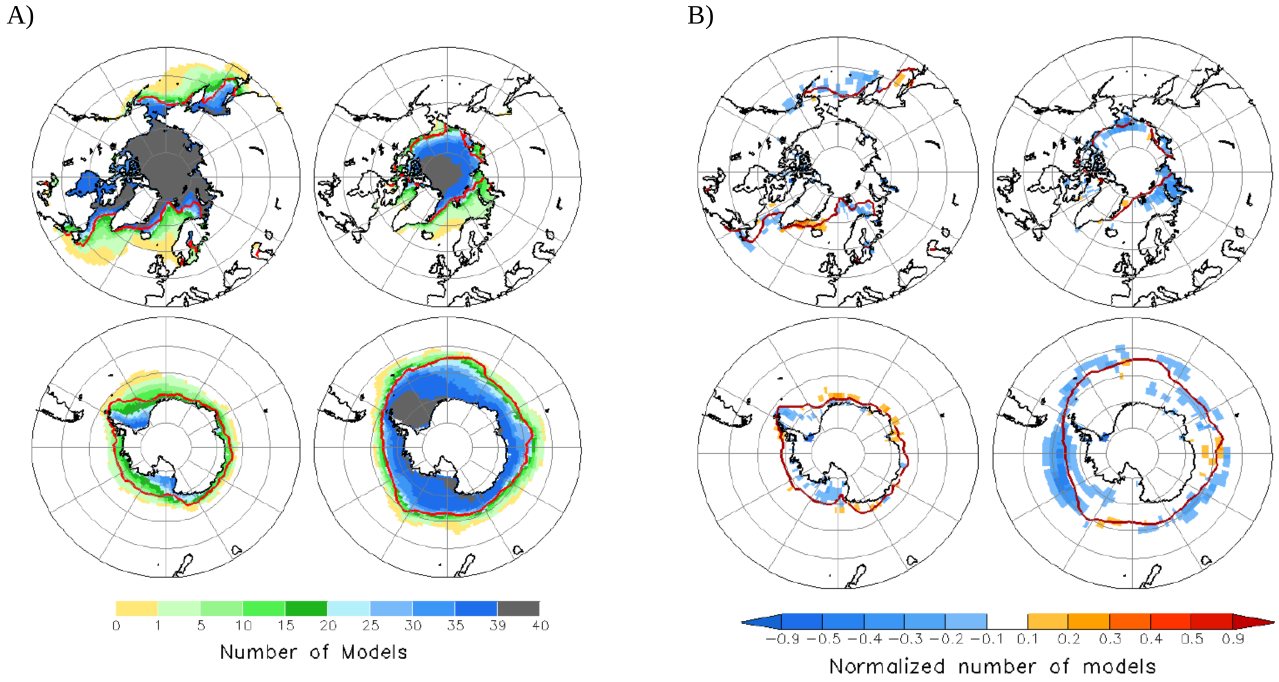


Figure 9.23: Sea ice distribution in the Northern Hemisphere (upper panels) and the Southern Hemisphere (lower panels) for March (left) and September (right). A) AR5 baseline climate (1986–2005) simulated by 40 of CMIP5 AOGCMs. For each 1° × 1° longitude-latitude grid cell, the figure indicates the number of models that simulate at least 15% of the area covered by sea ice. B) AR4 baseline climate (1980–1999) differences between 40 CMIP5 and 17 CMIP3 (AR4 (Randall et al., 2007) Figure 8.10) AOGCMs. For each 2.5° × 2.5° longitude-latitude grid cell, the figure indicates the difference in the normalized number of CMIP5 minus CMIP3 models that simulate at least 15% of the area covered by sea ice. The observed 15% concentration boundaries (red line) are based on the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) data set (Rayner et al., 2003a). Adapted from Pavlova et al. (2011).

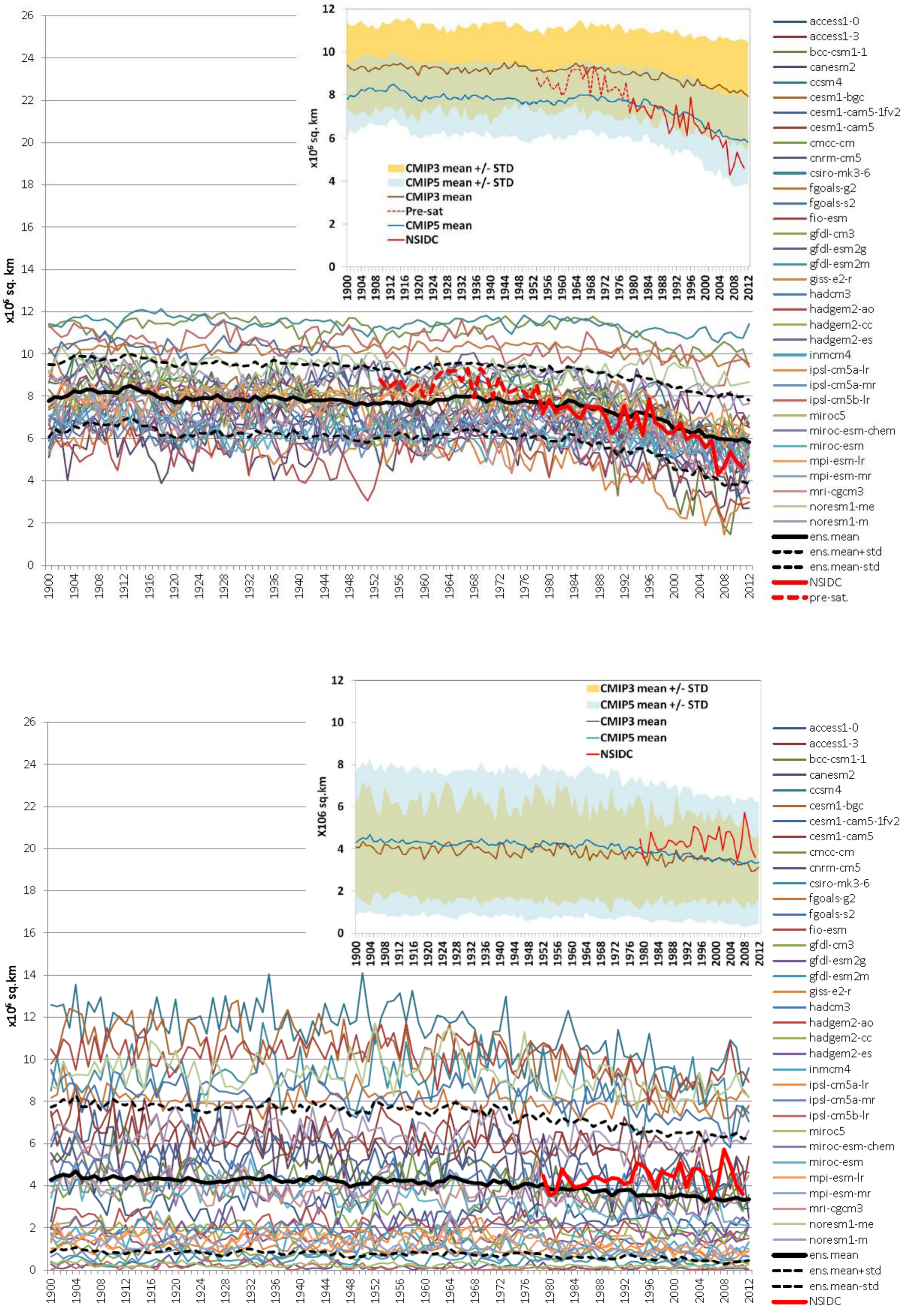


Figure 9.24: Time-series of CMIP5 modelled (coloured lines) and NSIDC (Fetterer et al., 2002) observed (thick red line) Arctic September (upper panel) and Antarctic March (lower panel) sea-ice extent from 1900 to 2012. The CMIP5 multi-model ensemble mean (black) is based on 34 CMIP5 models, with ±1 standard deviation shown as dashed black lines. The dashed thick red line for the Arctic (upper panel) relates to the pre-satellite period of observationally based time series. The upper and lower panel insets are based on the corresponding multi-model ensembles from CMIP5 and CMIP3, ±1 standard deviation. Note that these are monthly means, not yearly minima. Adapted from Stroeve et al. (2012).

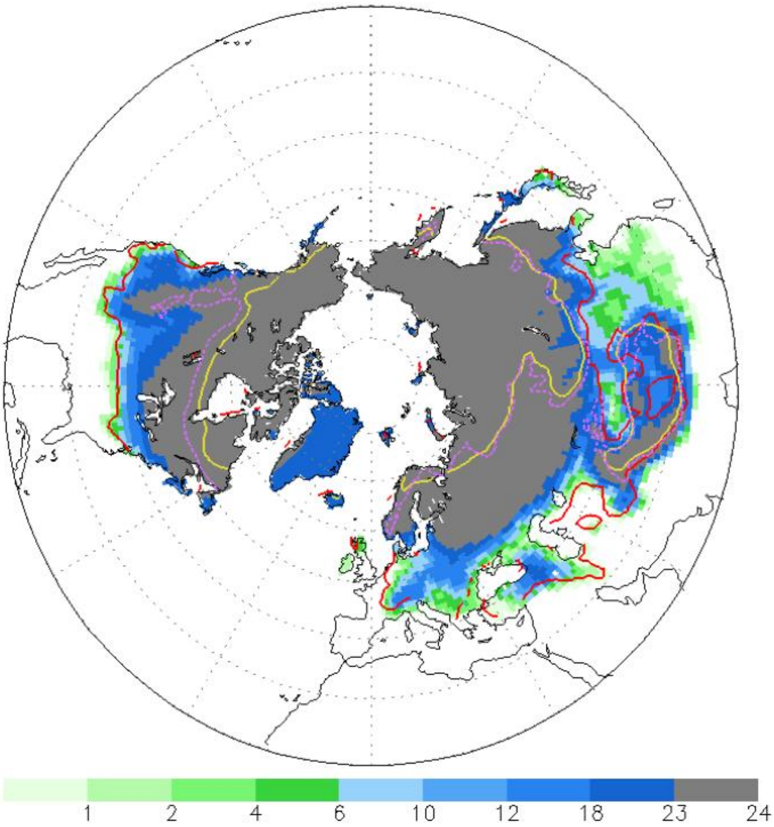


Figure 9.25: Terrestrial snow-cover distribution in the Northern Hemisphere simulated by 24 CMIP5 AOGCMs for February, updated for CMIP5 from (Pavlova et al., 2007). For each 1° × 1° longitude-latitude grid cell, the figure indicates the number of models that simulate at least 5 kg m^{-2} of snow water equivalent. The observationally based boundaries (red line) mark the territory with at least 20% of the days per month with snow cover (Robinson and Frei, 2000a) over the period 1986–2005. The annual mean 0°C isotherm at 3.3 m depth averaged across the 24 AOGCMs (yellow line) is a proxy for the permafrost boundary. Observed permafrost extent in the Northern hemisphere (magenta dashed line) is based on Brown et al. (1997, 1998).

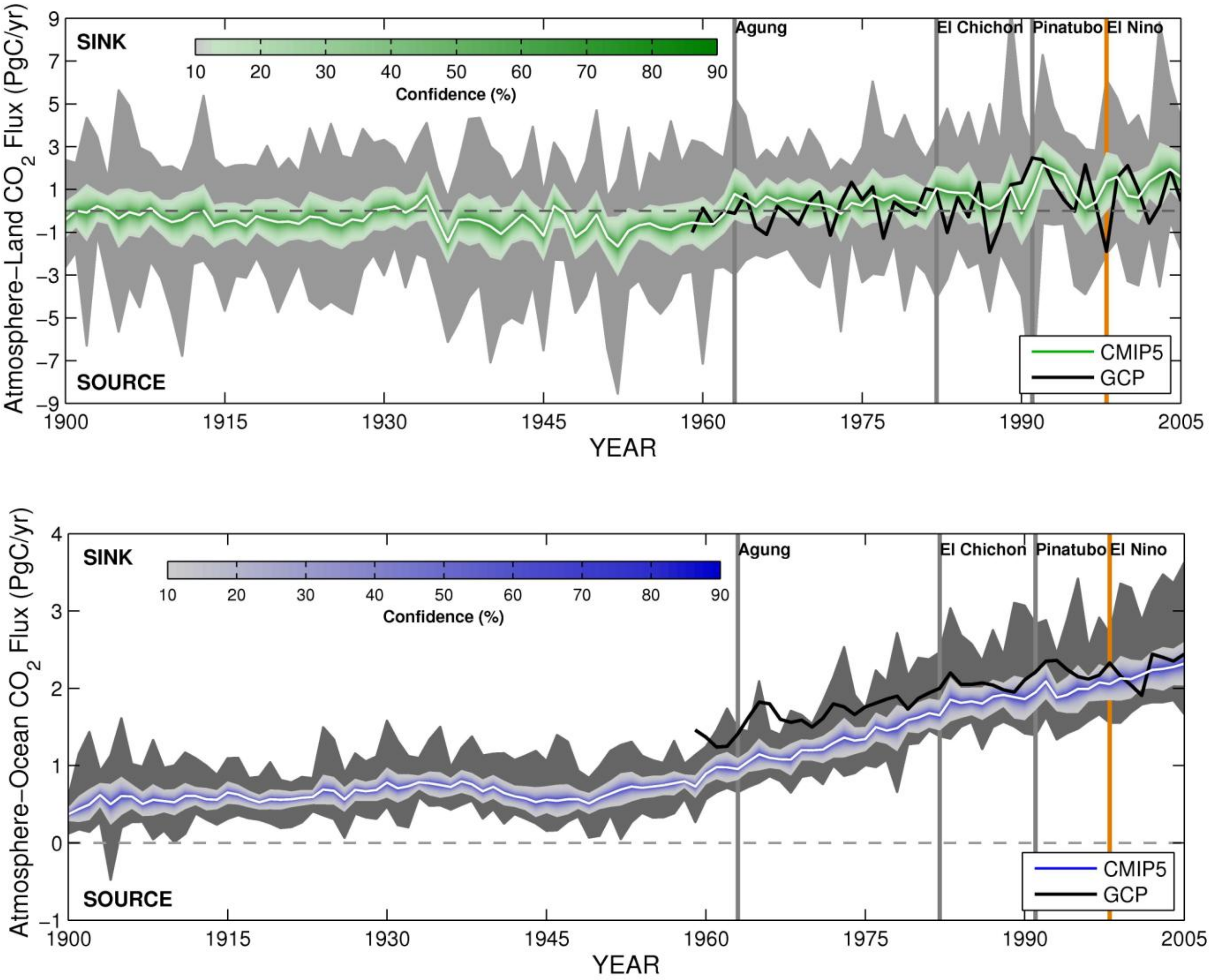
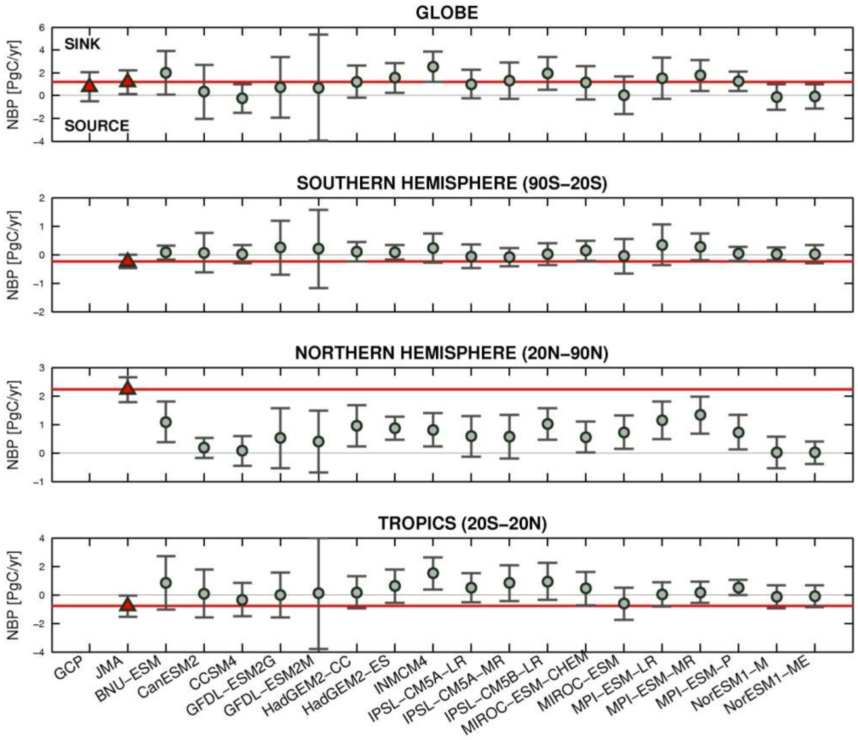


Figure 9.26: Ensemble mean of annual global land carbon uptake (top) and annual global ocean carbon uptake (bottom panel) in the CMIP5 ESMs for the historical period 1900–2005. For comparison, the observation-based estimates provided by the Global Carbon Project ("GCP", Le Quere et al., 2009) are also shown (black line). The confidence limits on the ensemble mean are derived by assuming that the CMIP5 models come from a t-distribution. The grey areas show the range of annual fluxes simulated across the model ensemble.

a) Atmosphere-Land CO_2 Flux (PgC/yr)



b) Atmosphere-Ocean CO_2 Flux (PgC/yr)

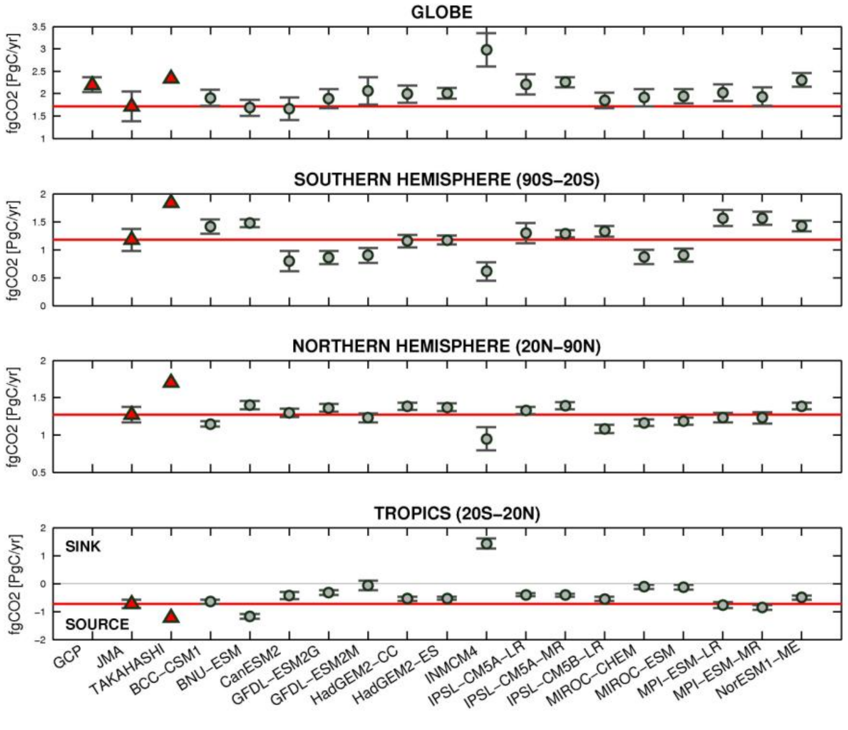


Figure 9.27: Simulation of net (a) atmosphere-land CO_2 fluxes ("NBP", top) and (b) atmosphere-ocean CO_2 fluxes ("fgCO2", bottom) in the CMIP5 ESMs, for the period 1986–2005. The error bars represent the interannual variability in the fluxes calculated as the standard deviation of the annual fluxes. . For comparison, the observation-based estimates provided by the Global Carbon Project ("GCP", Le Quere et al., 2009), the "JMA" atmospheric inversion (Gurney et al., 2003), and the Takahashi ocean pCO_2 dataset (Takahashi et al., 2009) are also shown as the red symbols, with the horizontal red line representing the estimate from JMA.

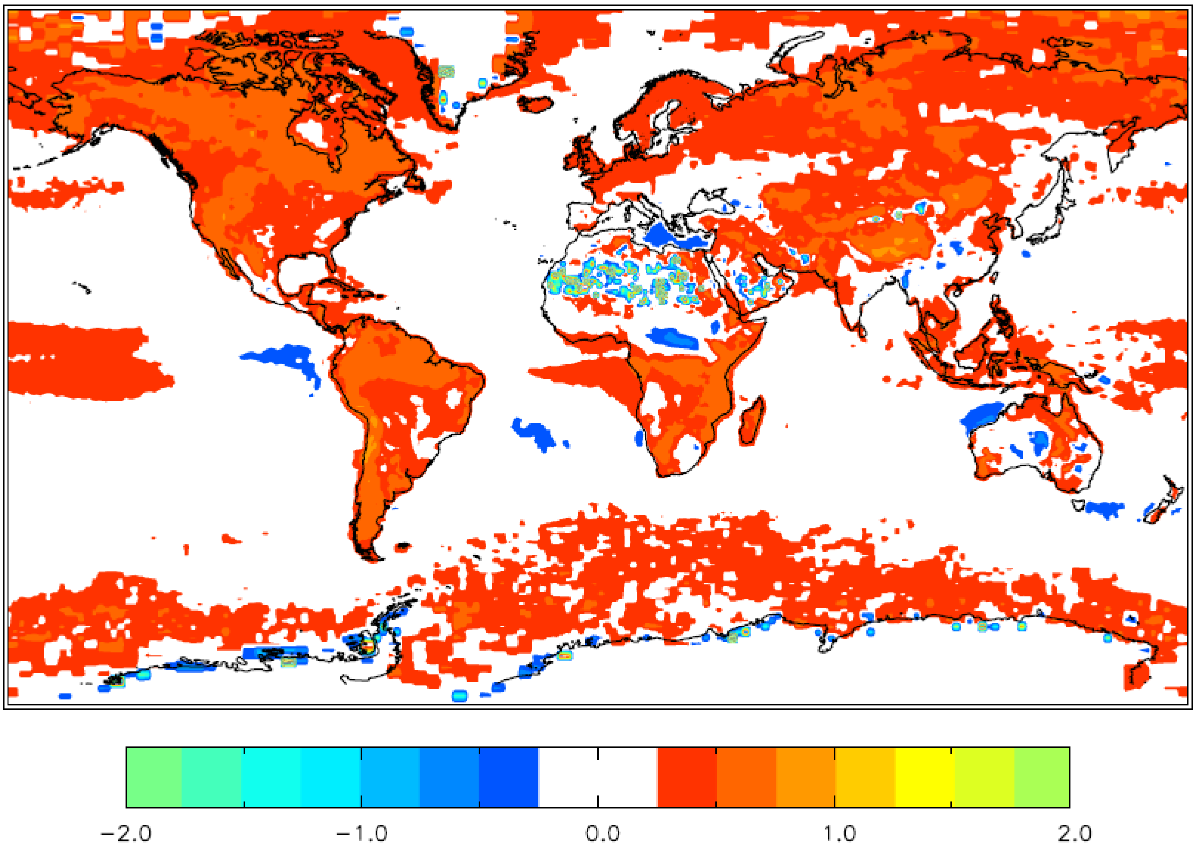


Figure 9.28: The relative error in visible aerosol optical thickness (AOT) from the median of a subset of CMIP5 models' historical simulations, relative to satellite retrievals of AOT. Units are such that +1 is equivalent to satellite AOT exceeding model AOT by 100%. The figure is constructed following (Kinne et al., 2006). The satellite AOT is from the MODIS instrument on the NASA Terra satellite from 2001 through 2005. The data version is MODIS 4; the model outputs are from ACCESS1-0, ACCESS1-3, BNU-ESM, CESM1-CAM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadGEM2-CC, HadGEM2-ES, IPSL-CM5A-LR IPSL-CM5A-MR, IPSL-CM5B-LR, MIROC-ESM, MIROC-ESM-CHEM, MIROC5, MRI-CGCM3, NorESM1-M, and NorESM1-ME.

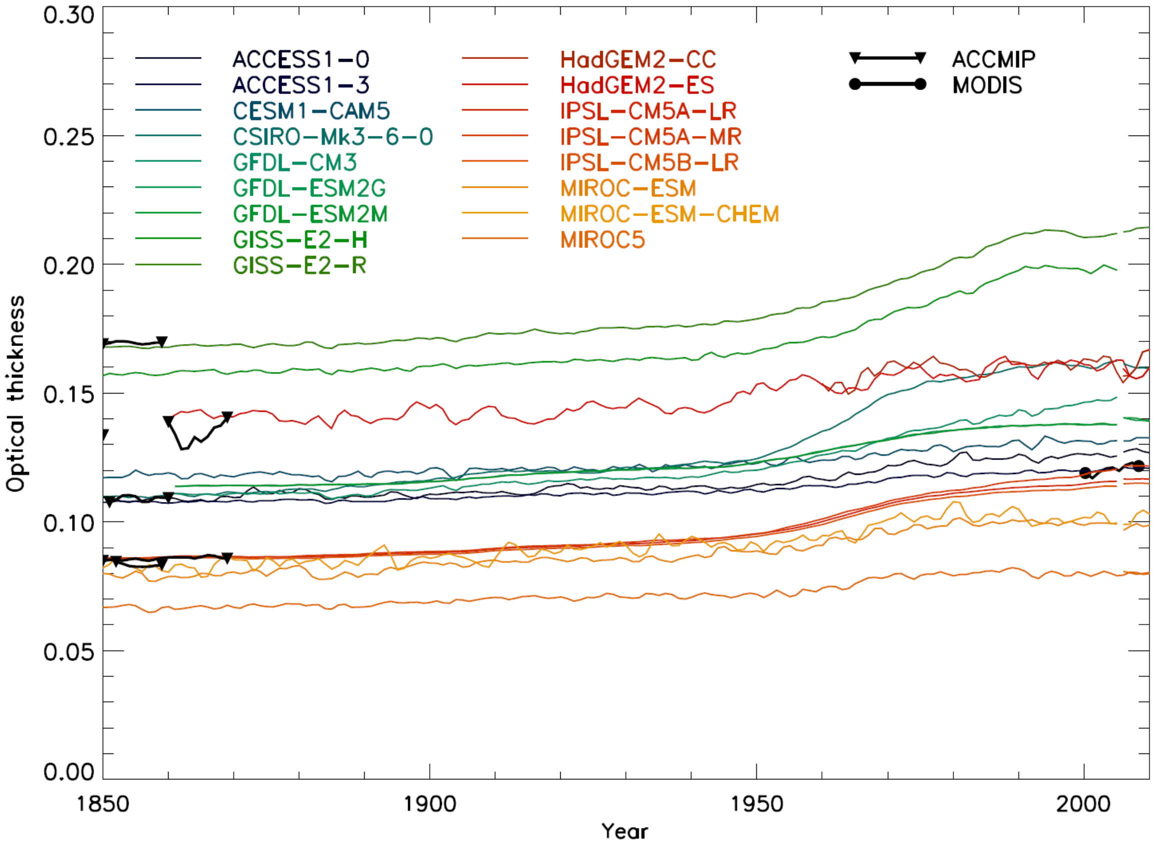


Figure 9.29: Time series of global oceanic-mean AOT from individual CMIP5 models' historical (1850–2005) and RCP4.5 (2006–2010) simulations, corrected MODIS satellite observations by Shi et al. (2011) and Zhang et al. (2008a), and the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) simulations for the 1850s by Shindell et al. (2012b). ACCMIP model output is from CICERO-OsloCTM2, GISS-E2-R, HadGEM2, LMDzORINCA, NCAR-CAM3.5, and NCAR-CAM5.1.

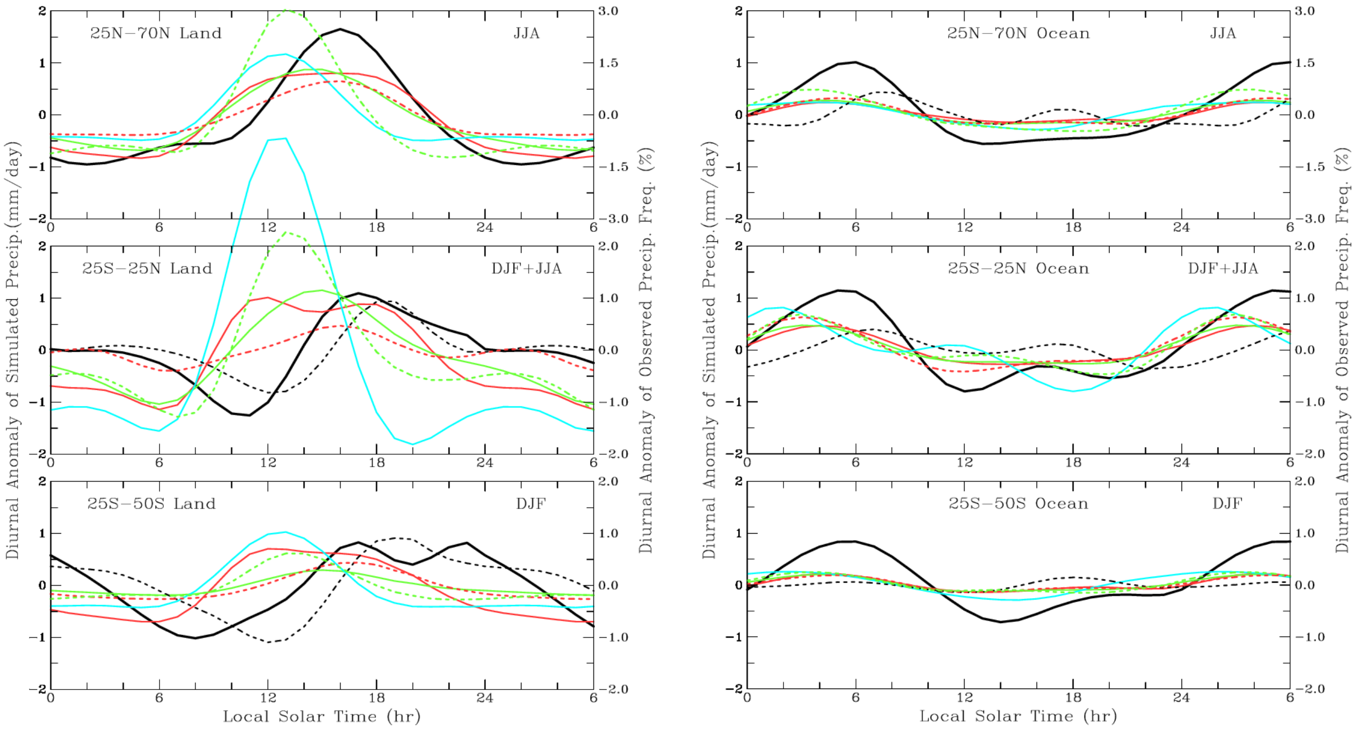


Figure 9.30: Composite diurnal cycle precipitation from observations (black) and a subset of CMIP3 models (coloured lines) averaged over land (left) and ocean (right) areas for three different zones at each local time and seasons (JJA, DJF). Black solid line: surface-observed precipitation frequency; black dashed line: TRMM 3B42 dataset, 1998–2003 mean]. Each colored curve is for one model: CCSM2 (red solid), GFDL-CM2.0 (green solid), GISS-ER (blue), MIROC3.2 (red dashed), and MRI-CGCM2.3.2a (green dashed). Adapted from (Dai, 2006).

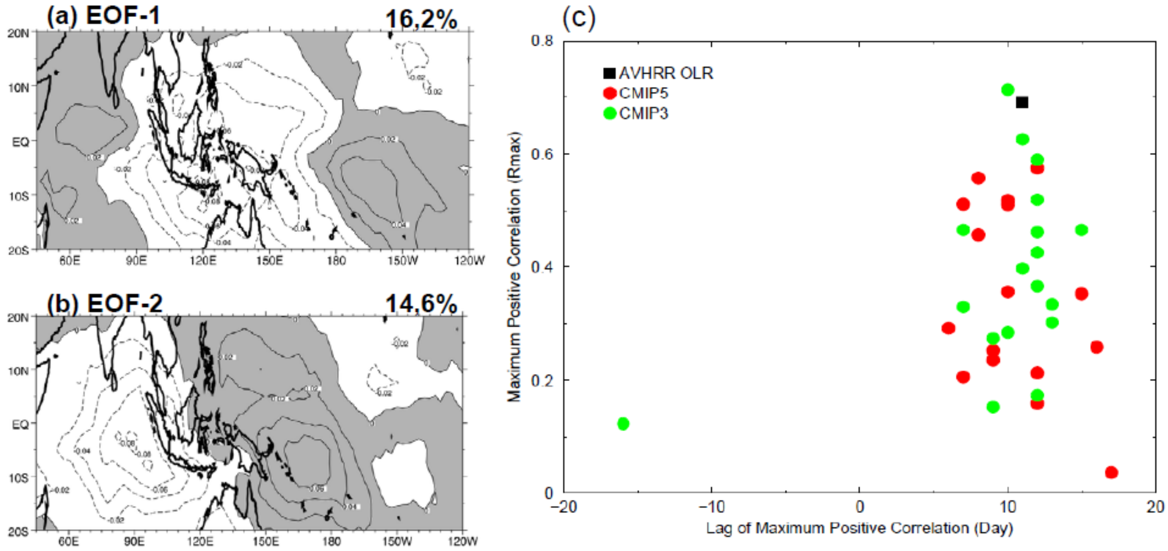


Figure 9.31: Two leading Empirical Orthogonal Functions (EOF's) of Outgoing Longwave Radiation (OLR) from years of strong MJO variability (Sperber, 2003). The 20–100 day filtered OLR from observations and each of the CMIP5 historical simulations and the CMIP3 simulations of 20th century climate is projected on these two leading EOF's to obtain MJO Principal Component time series. The scatter plot shows the maximum positive correlation between the resulting MJO Principal Components and the time lag at which it occurred for all winters (November-March). The maximum positive correlation is an indication of the coherence with which the MJO convection propagates from the Indian Ocean to the Maritime Continent/western Pacific, and the time lag is approximately 1/4 of the period of the MJO. Most models have weaker coherence in the MJO propagation (smaller maximum positive correlation), and some have periods that are too short compared to observations. One CMIP3 model has its maximum positive correlation at Day –16, indicating that this model is incorrectly dominated by westward propagation. Constructed following Sperber and Kim (2012).

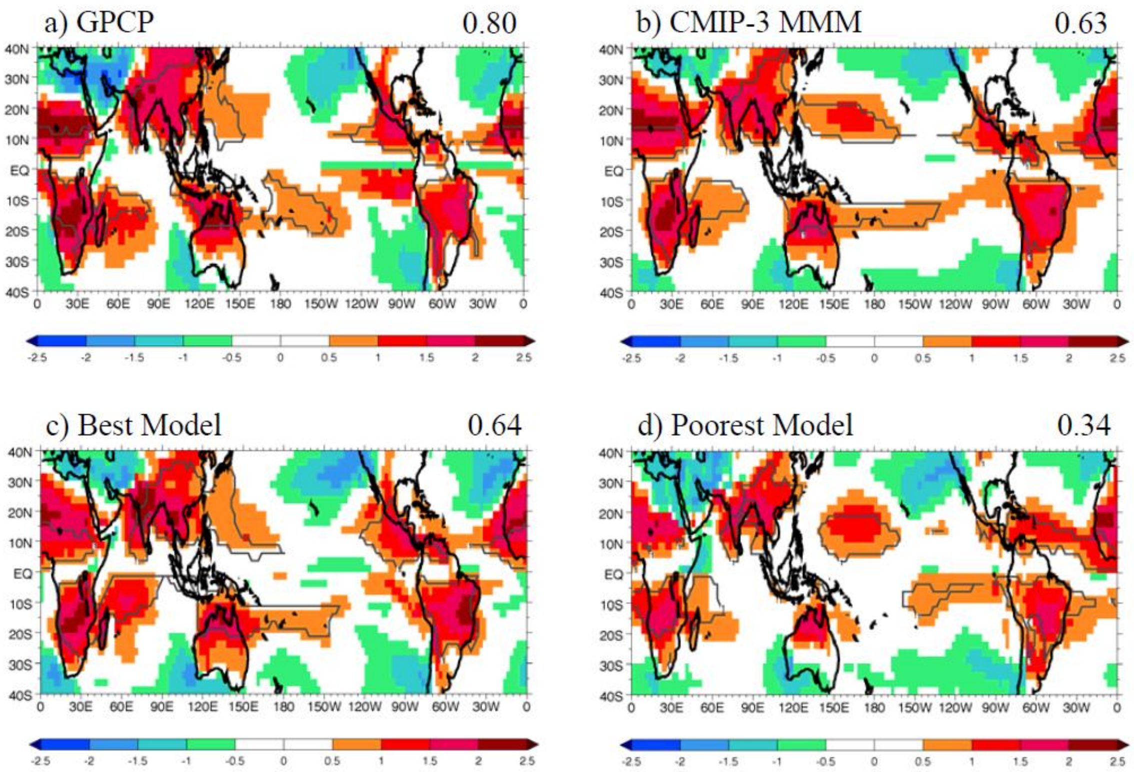


Figure 9.32: Monsoon precipitation intensity (shading, mm/day) and monsoon precipitation domain (lines) are shown for (a) observations from GPCP, (b) the CMIP3 multi-model mean, (c) the best model, and (d) the worst model in terms of the threat score for this diagnostic. The threat scores indicate how well the models represent the monsoon precipitation domain compared to the GPCP data. The threat score in panel (a) is between GPCP and CMAP rainfall to indicate observational uncertainty. A threat score of 1.0 would indicate perfect agreement between the two datasets. See Wang and Ding (2008); Wang et al. (2011a); and Kim et al. (2011) for details of the calculations.

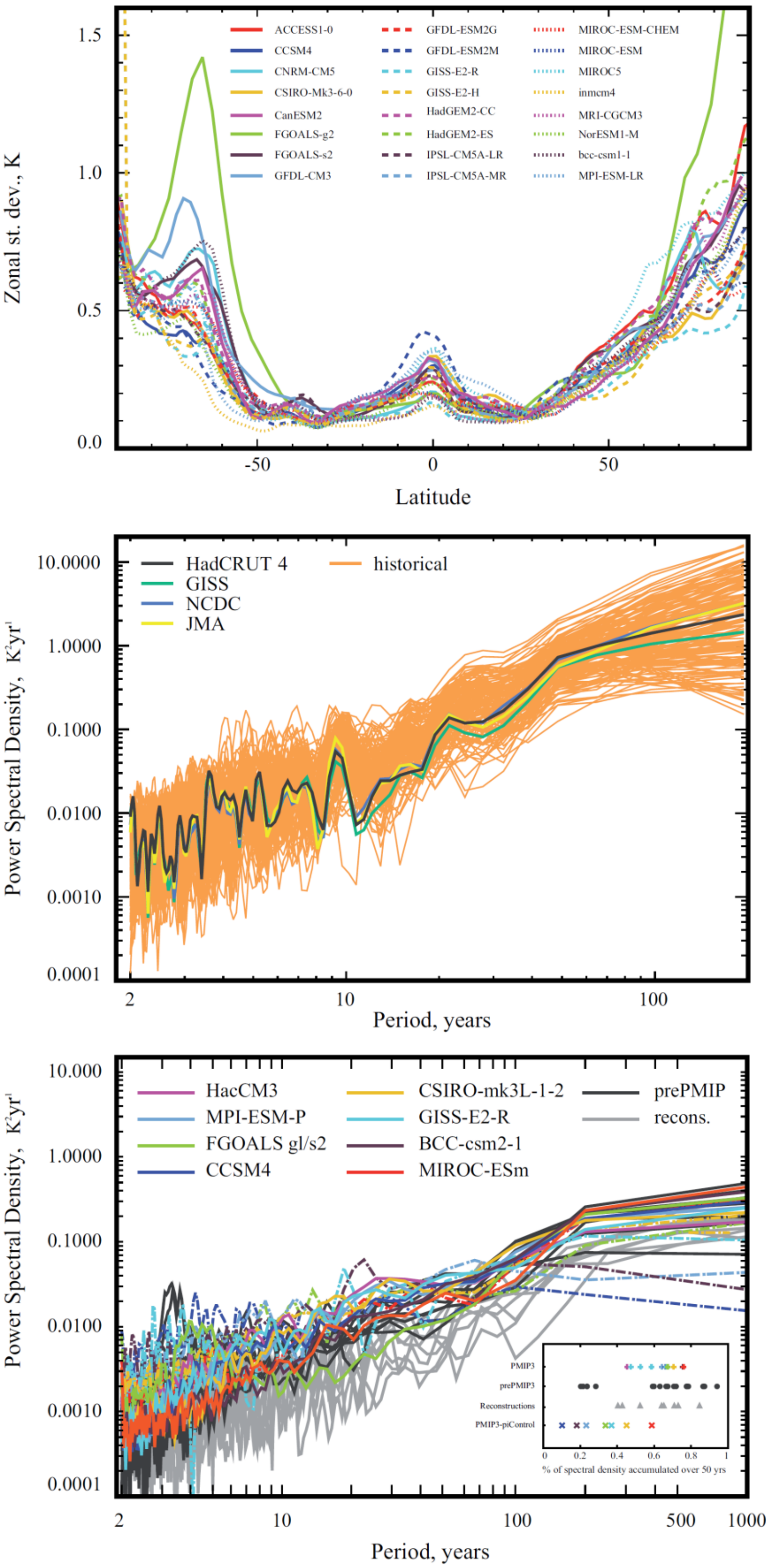


Figure 9.33: Global climate variability represented as (top): Standard deviation of zonal-mean surface temperature of the CMIP5 pre-industrial control simulations (after Jones et al., 2012). (Middle): Power spectral density for 1901–2010 global-mean surface temperature for both historical CMIP5 simulations and the observations (after Jones et al., 2012). (Bottom): Power spectral density for Northern-Hemisphere surface temperature from the CMIP5-PMIP3 last-millennium simulations using common external forcing configurations (colour, (Schmidt et al., 2012), together with the

corresponding pre-industrial simulations (colour, dashed), previous last-millennium AOGCM simulations (black, (Fernandez-Donado et al., 2012), and temperature reconstructions from different proxy records (see Section 5.3.5). The small panel included in the bottom panel shows for the different models and reconstructions the percentage of spectral density cumulated for periods above 50 year, so as to better highlight the differences between unforced (pre-industrial control) and forced (PMIP3 and pre-PMIP3) simulations, compared to temperature reconstruction for the longer time scales.

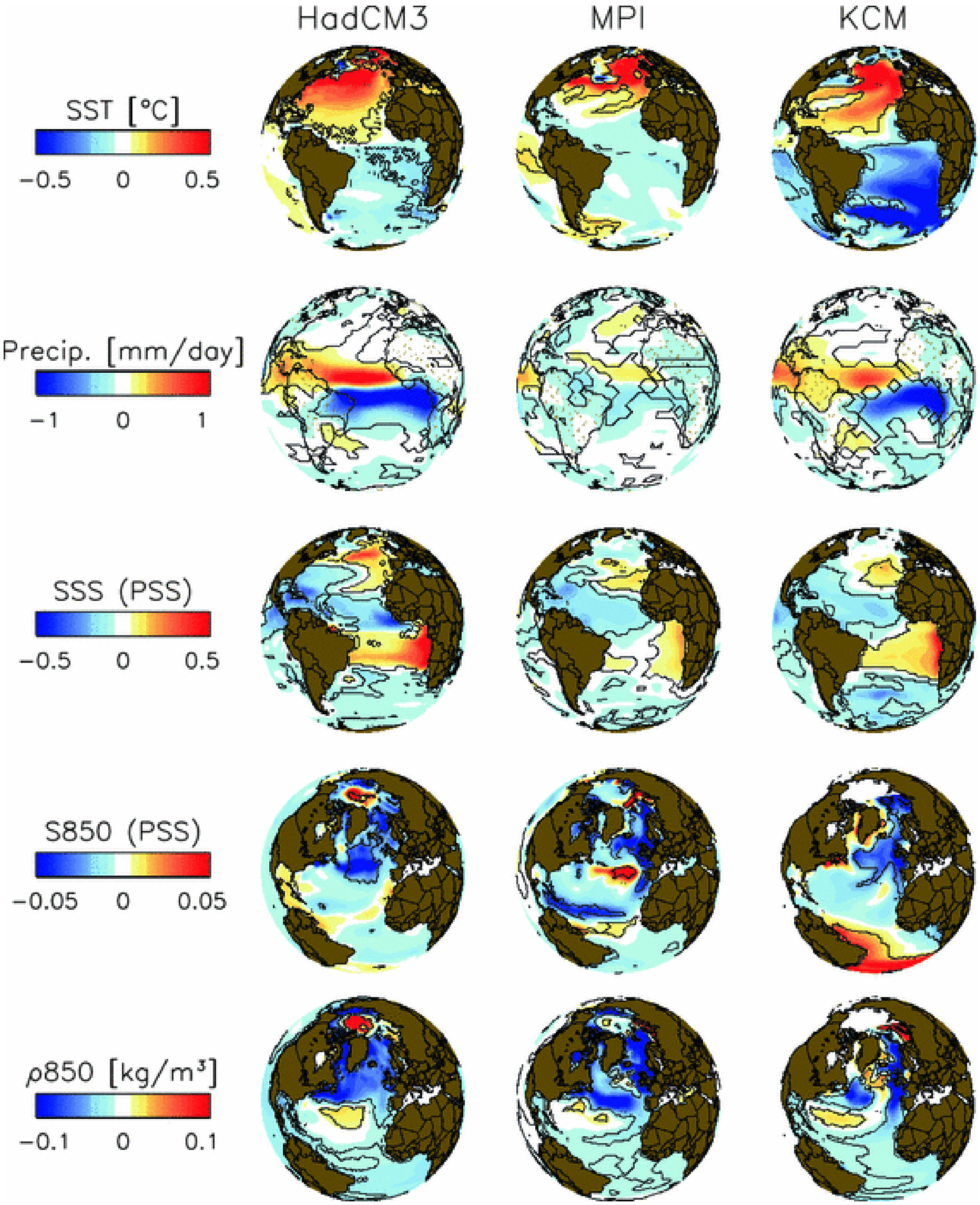


Figure 9.34: From top to bottom: SST composites using AMOC time series; precipitation composites using cross-equatorial SST difference time series; equatorial salinity composites using ITCZ-strength time series; subpolar-gyre depth-averaged salinity (top 800–1,000 m) using equatorial salinity time series; subpolar gyre depth averaged density using subpolar gyre depth averaged salinity time series. From left to right: HadCM3, MPI-ESM, and KCM. Black outlining signifies areas statistically significant at the 5% level for a two-tailed t test using the moving-blocks bootstrapping technique (Wilks, 1995). Figure 3 from Menary et al. (2011).

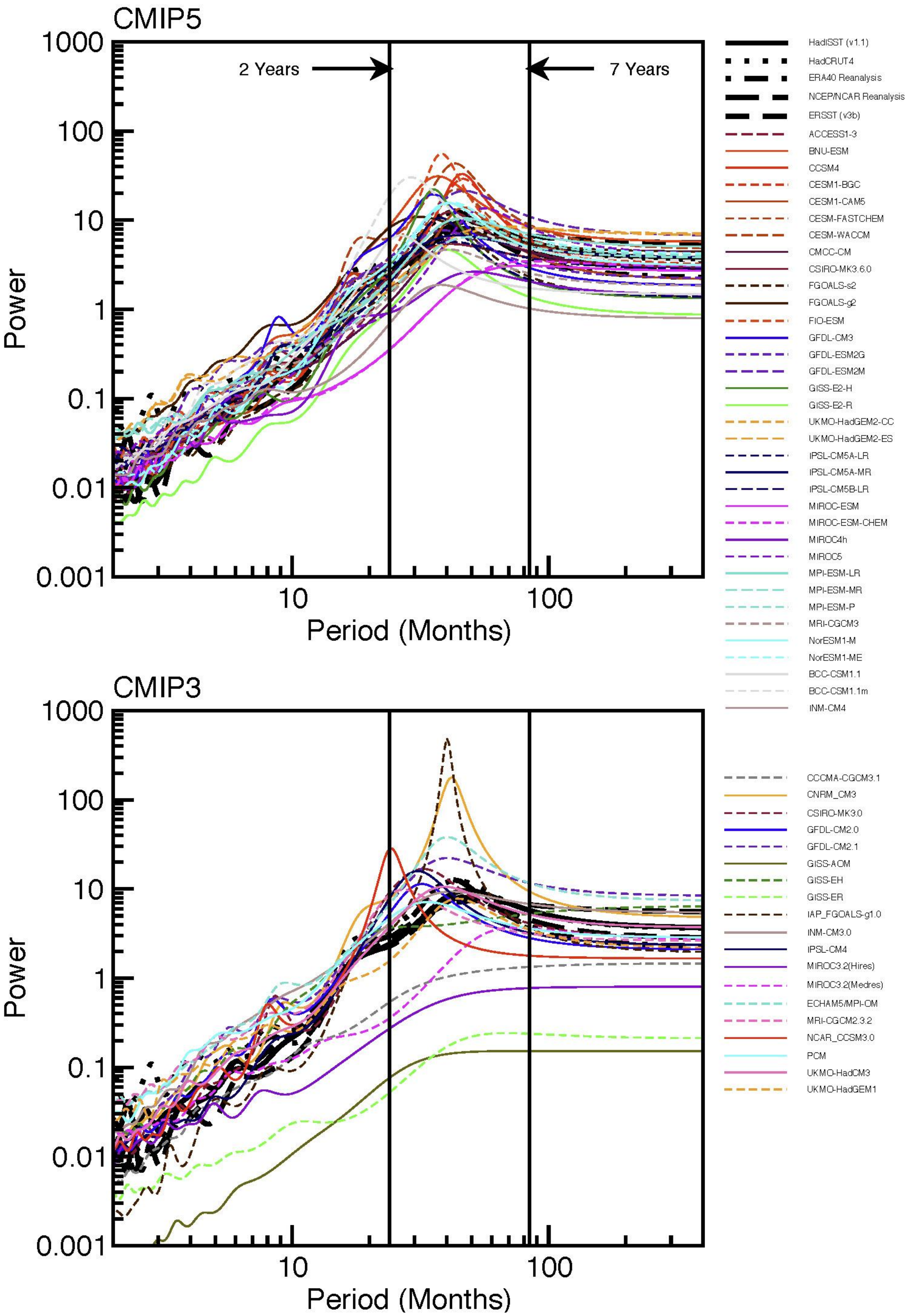


Figure 9.35: Maximum entropy power spectra of surface air temperature averaged over the NINO3 region (i.e., 5°N to 5°S, 150°W to 90°W) for (a) the CMIP3 models and (b) the CMIP5 models. Note that ECMWF reanalysis in (a) refers

to the European Centre for Medium Range Weather Forecasts (ECMWF) 15-year reanalysis (ERA15). The vertical lines correspond to periods of two and seven years. The power spectra from the reanalyses and for SST from the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) version 1.1, HadCRU 4, ERA40 and NCEP/NCAR data set are given by the series of solid, dashed and/or dotted black curves. Adapted from (AchutaRao and Sperber, 2006).

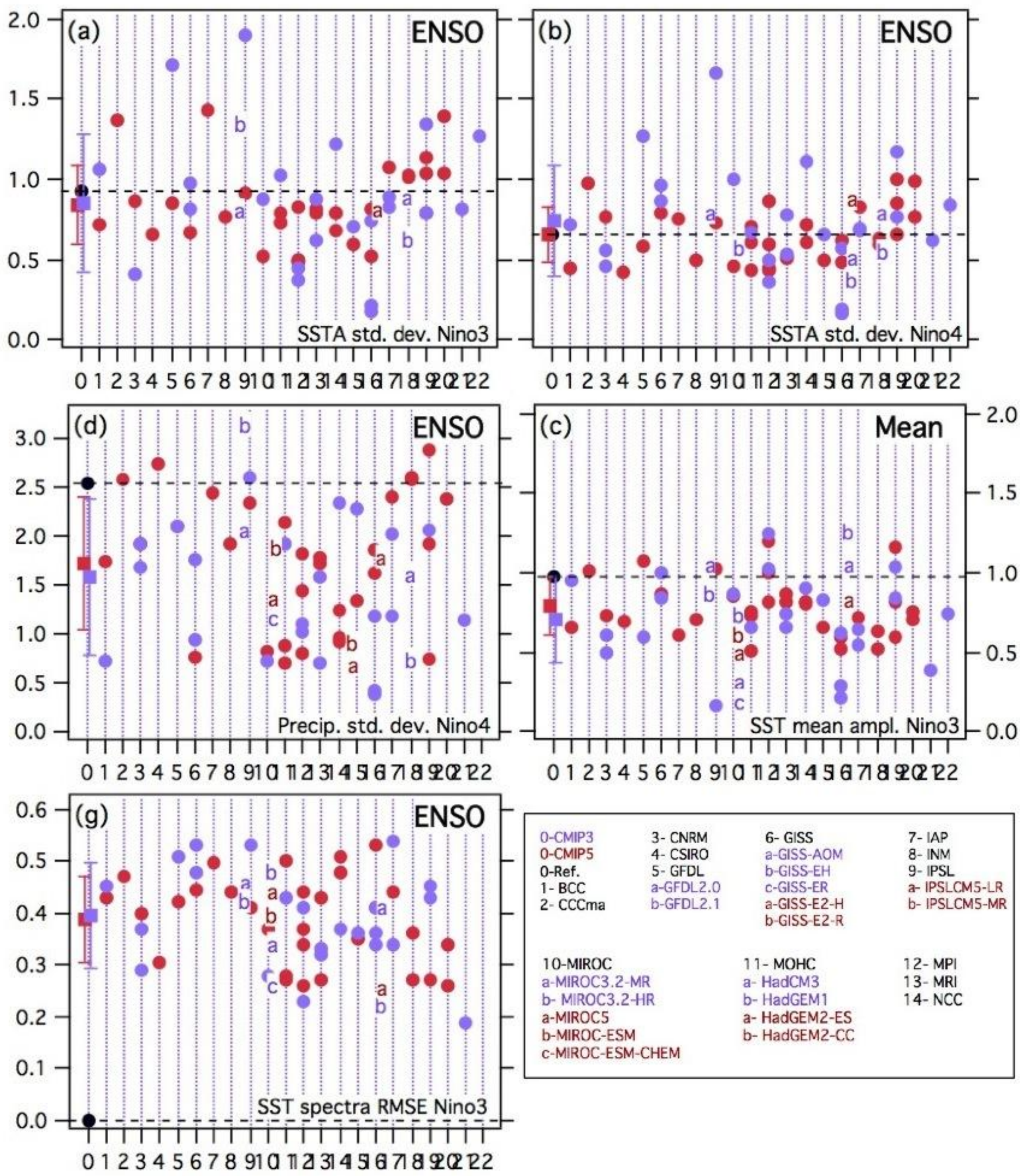


Figure 9.36: ENSO and mean tropical Pacific metrics for pre-industrial control simulations in CMIP3 (blue) and CMIP5 (red). (a) and (b): SST anomaly standard deviation in Niño 3 and Niño 4 (°C), (c) SST annual cycle amplitude in Niño3, (°C), (d) precipitation response (standard deviation) in Niño4 (mm/day), (g) ENSO power spectrum (Niño3) RMS error, (°C). Reference datasets, shown as black solid circles and dashed lines: HadISST1.1 for (a), (b), (c), and (g); CMAP for (d). The CMIP3 and CMIP5 multi-model mean are shown as squares on the left of each panel with the whiskers representing the model standard deviation.

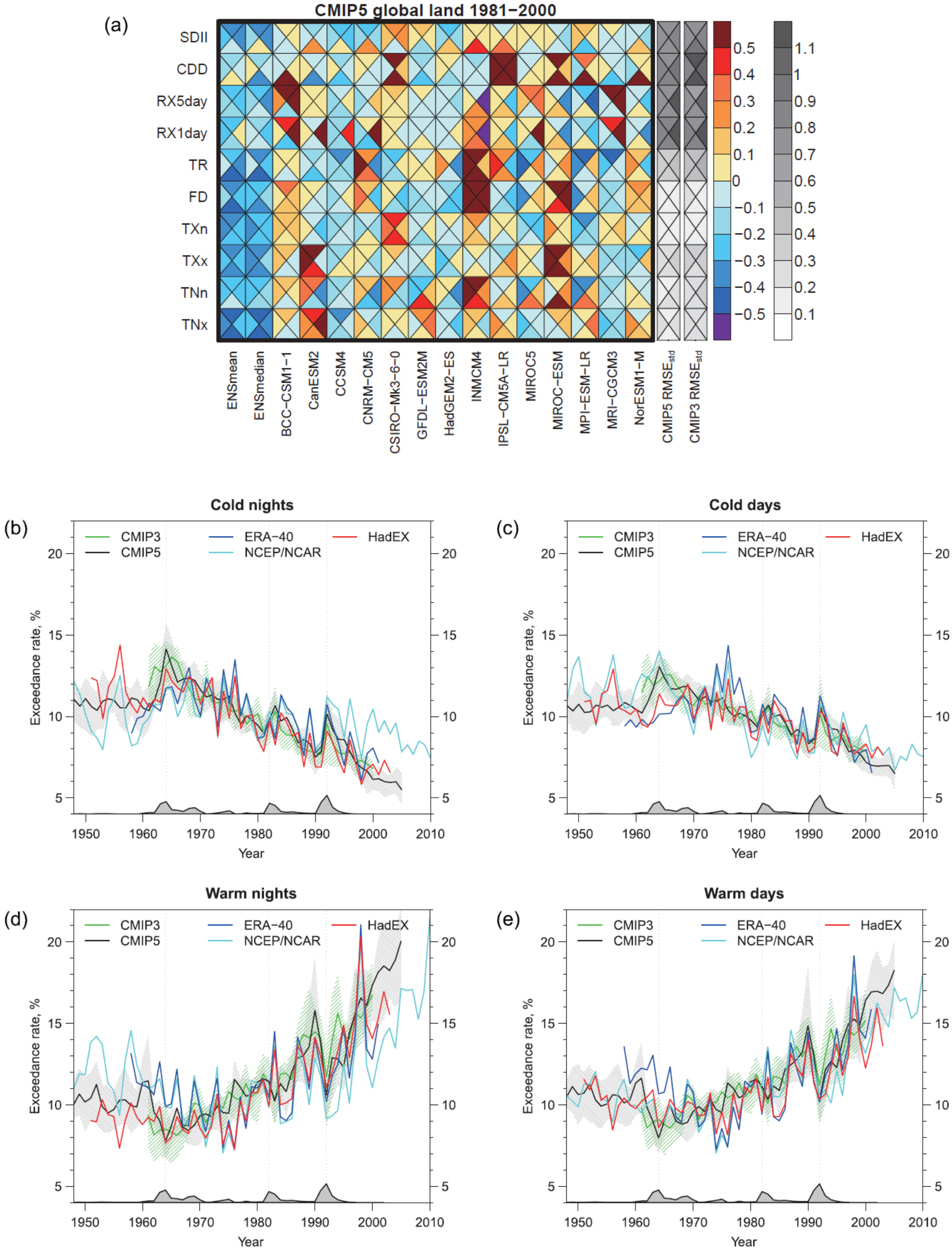


Figure 9.37: (a) Portrait plot display of relative error metrics for the CMIP5 temperature and precipitation extreme indices based on Sillmann et al. (2012). Reddish and bluish colours indicate, respectively, larger and smaller RMS errors for an individual model relative to the median model. The gray-shaded columns on the right side indicate the RMS error for the multi-model median standardized by the spatial standard deviation of the index climatology in the reanalysis, representing absolute errors for CMIP3 and CMIP5 ensembles. Results for four different reference datasets, ERA-interim (top), ERA40 (left), NCEP/NCAR (right) and NCEP-DOE (bottom) reanalyses, are shown in each box. The analysis period is 1981–2000 and only land areas are considered. The indices shown are simple daily precipitation intensity index (SDII), consecutive dry days (CDD), annual maximum 5-day/1-day precipitation (RX5day/RX1day), tropical nights (TR), frost days (FD), annual minimum/maximum daily maximum surface air temperature (TXn/TXx)

and annual minimum/maximum daily minimum surface air temperature (TNn/TNx). Note that only a small selection of the indices analysed in Sillmann et al. (2012) is shown, preferentially those that appear in other Chapters (2, 10, 11, 12, 14). (b)-(e) Time series of global mean temperature extreme indices over land from 1948 to 2010 for CMIP3 (green) and CMIP5 (black) models, ERA40 (blue) and NCEP/NCAR (cyan) reanalyses and HadEX station-based observational dataset (red) based on Sillmann et al. (2012). Shading for model results indicates the 25th to 75th quantile range of inter-model spread. Grey shading along the horizontal axis indicates the evolution of globally averaged volcanic forcing according to Sato et al. (1993). The indices shown are the frequency of daily minimum/maximum surface air temperature below 10th percentile (b: Cold nights/c: Cold days) and that above 90th percentile (d: Warm nights/e: Warm days) of 1961–1990 base period. Note that, as these indices essentially represent changes relative to the base period, they are particularly suitable for being shown in time series and not straightforward for being shown in (a).

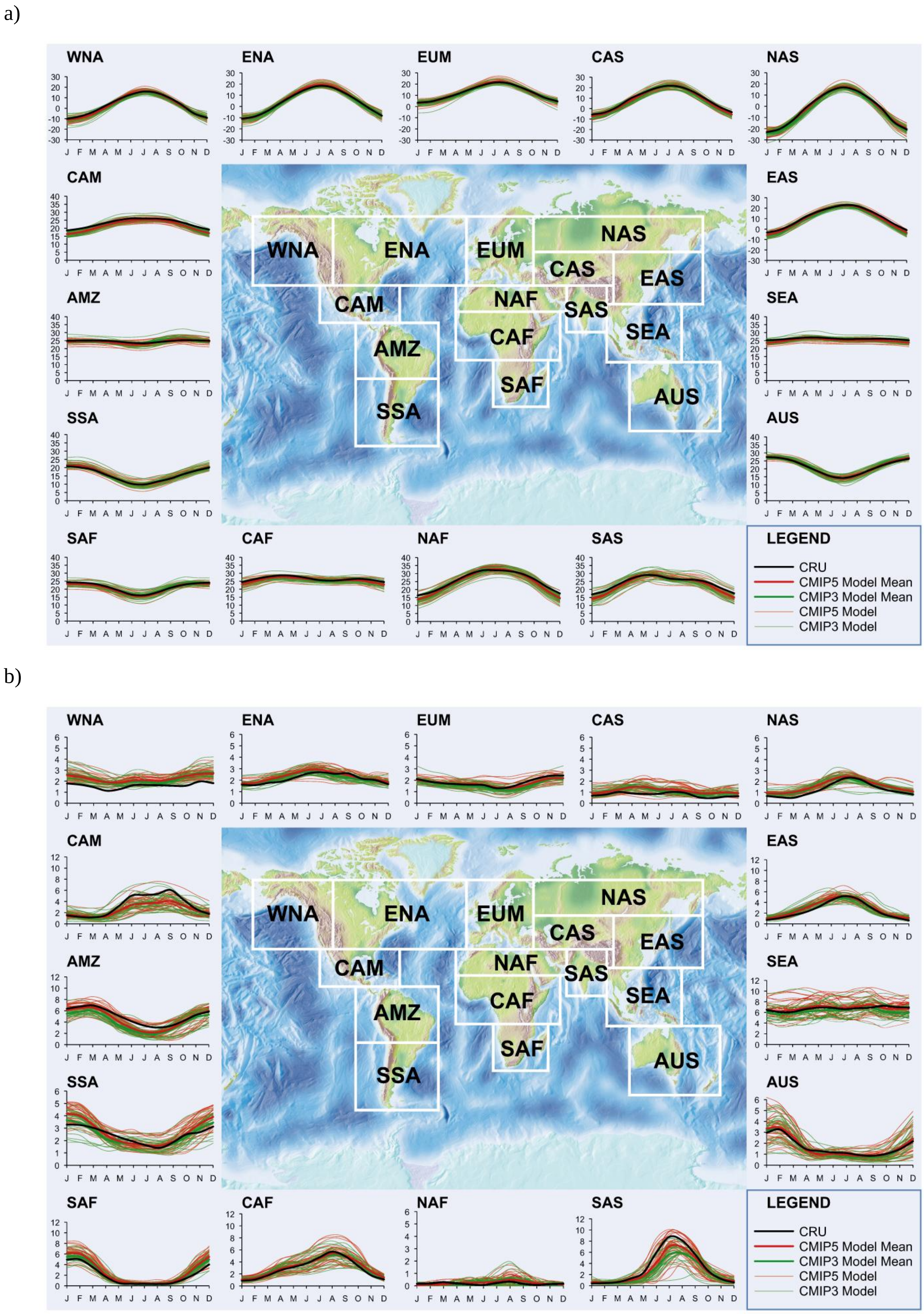


Figure 9.38: Mean annual cycle of (a) temperature (°C) and (b) precipitation (mm day^{-1}). The average is taken over land areas within the indicated rectangles, and over the period 1980–1999. The red thick line is the average over 37 CMIP5 models (red thin lines), the green thick line is the average over 22 CMIP3 models (green thin lines) and the black thick line is the CRU TS3.10 observational data for temperature and CRU TS3.10.1 for precipitation. Note that some of the sub-plots have different axis-ranges.

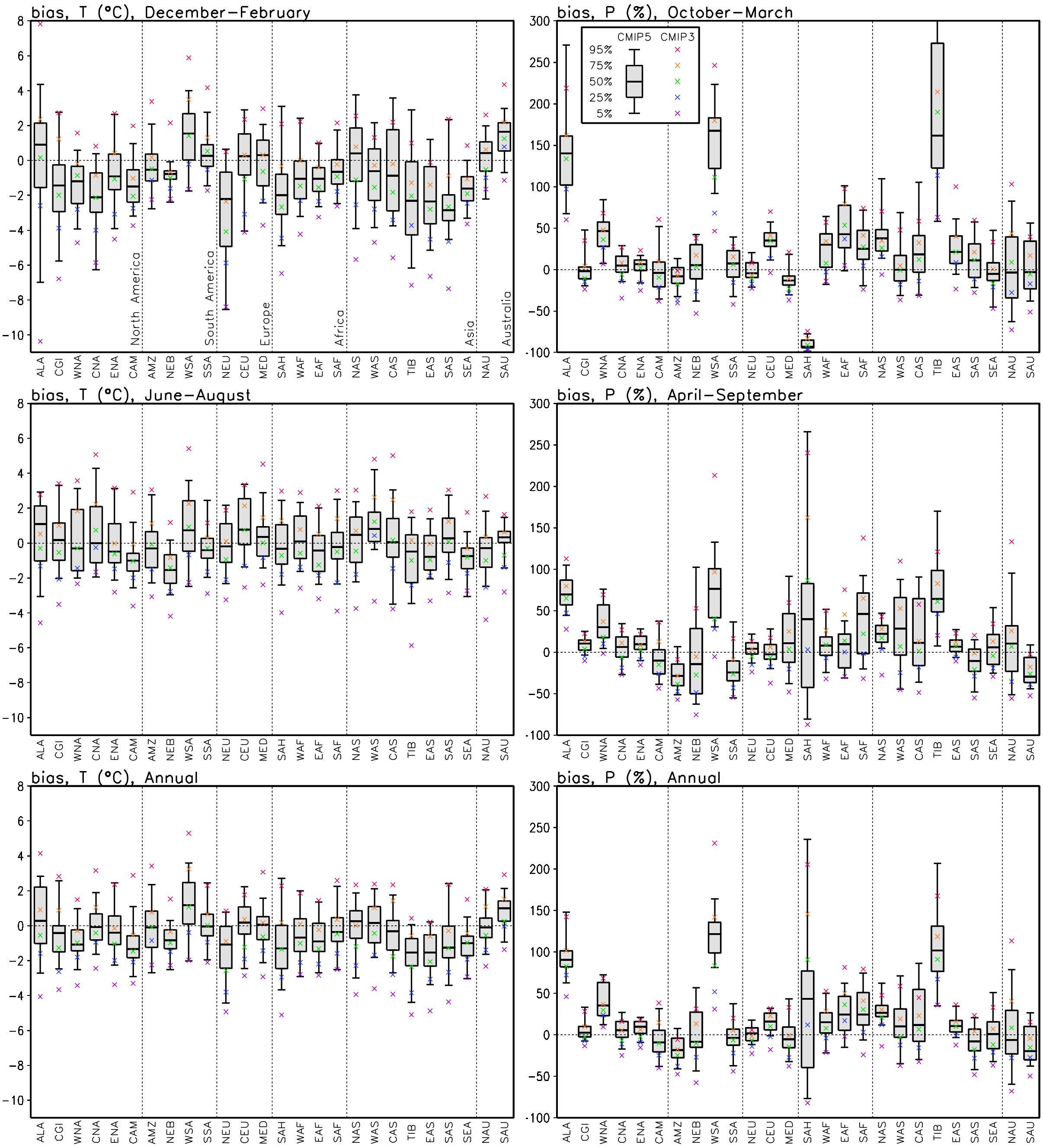


Figure 9.39: Seasonal and annual mean biases of (left) temperature (in °C) and (right) precipitation (in %) in the SREX land regions (cf. Seneviratne et al., 2012, page 12. The region's coordinates can be found from their online Appendix 3.A). The 5th, 25th, 50th, 75th and 95th percentiles of the biases in 34 CMIP5 models are given in box-whisker format, and the corresponding values for 23 CMIP3 models with crosses, as indicated in the legend in the top-right panel. The CMIP3 models' 20C3M simulations are complemented with the corresponding A1B runs for the 2001–2005 period. The biases are calculated over the period 1986–2005, using CRU T3.10 as the reference for temperature and CRU TS 3.10.01 for precipitation. The regions are: Alaska/NW Canada (ALA), Eastern Canada/Greenland/Iceland (CGI), Western North America (WNA), Central North America (CNA), Eastern North America (ENA), Central America/Mexico (CAM), Amazon (AMZ), NE Brazil (NEB), West Coast South America (WSA), South-Eastern South America (SSA), Northern Europe (NEU), Central Europe (CEU), Southern Europe/the Mediterranean (MED), Sahara (SAH), Western Africa (WAF), Eastern Africa (EAF), Southern Africa (SAF), Northern Asia (NAS), Western Asia (WAS), Central Asia (CAS), Tibetan Plateau (TIB), Eastern Asia (EAS), Southern Asia (SAS), South-Eastern Asia (SEA), Northern Australia (NAS) and Southern Australia/New Zealand (SAU). Note that the region WSA is poorly resolved in the models.

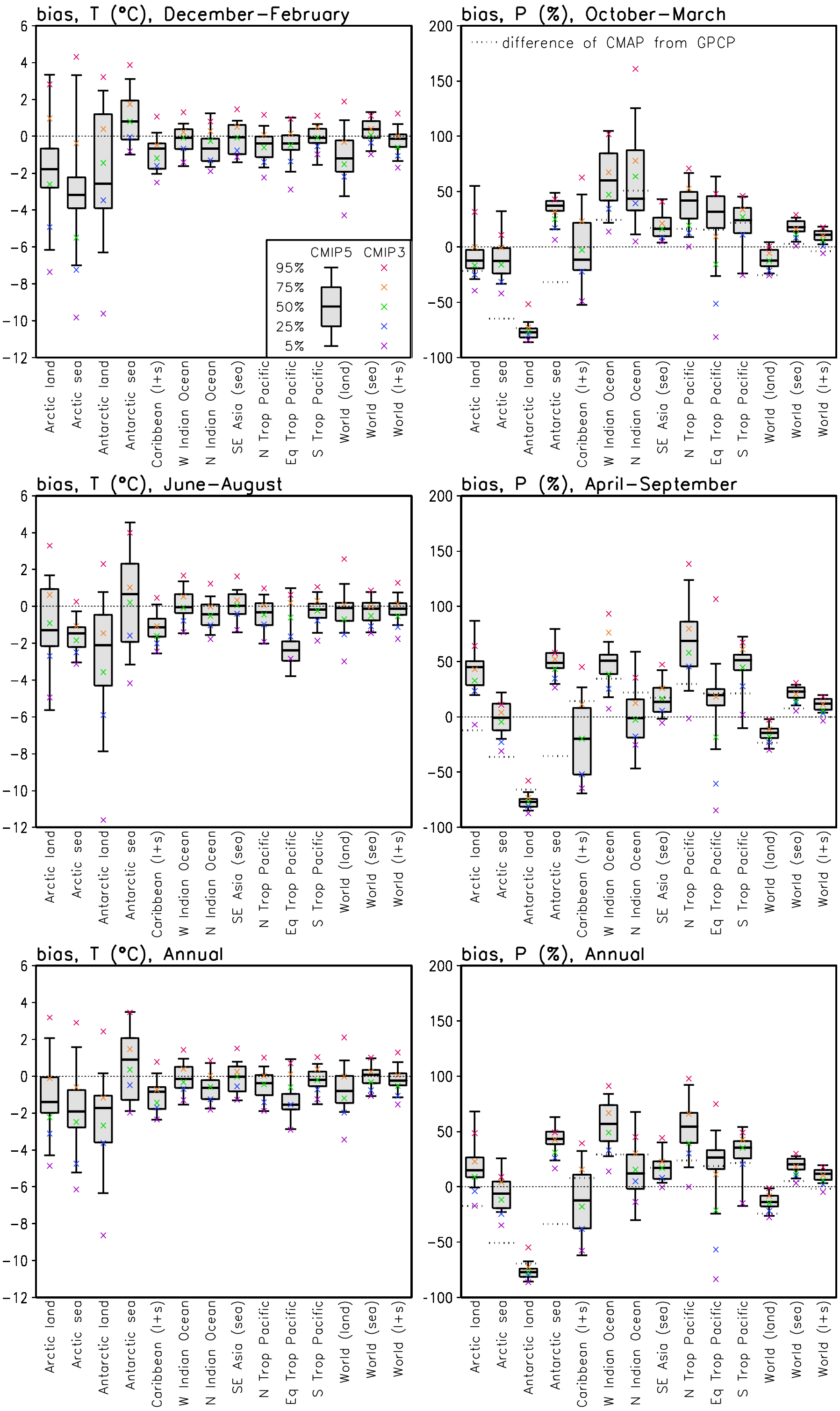


Figure 9.40: As Figure 9.39, but for various polar and ocean regions, and using ERA Interim reanalysis as the reference for temperature and GPCPCMAP for precipitation. Global land, ocean and overall means are also shown. The regions shown are defined as; Arctic: 67.5-90°N, Caribbean: 10°N–25°N, 85°W–60°W, West Indian Ocean: 25°S–5°N, 52°E–75°E; North Indian Ocean: 5°N–30°N, 60°E–95°E; Northern Tropical Pacific: 5°N–25°N, 155°E–150°W; Equatorial Tropical Pacific: 5°S–5°N, 155°E–130°W; Southern Tropical Pacific: 5°S–25°S, 155°E–130°W; Antarctic: 50°S–90°S. As an indicator of observational uncertainty, the normalised difference between CMAP and GPCP precipitation is shown with dotted lines.

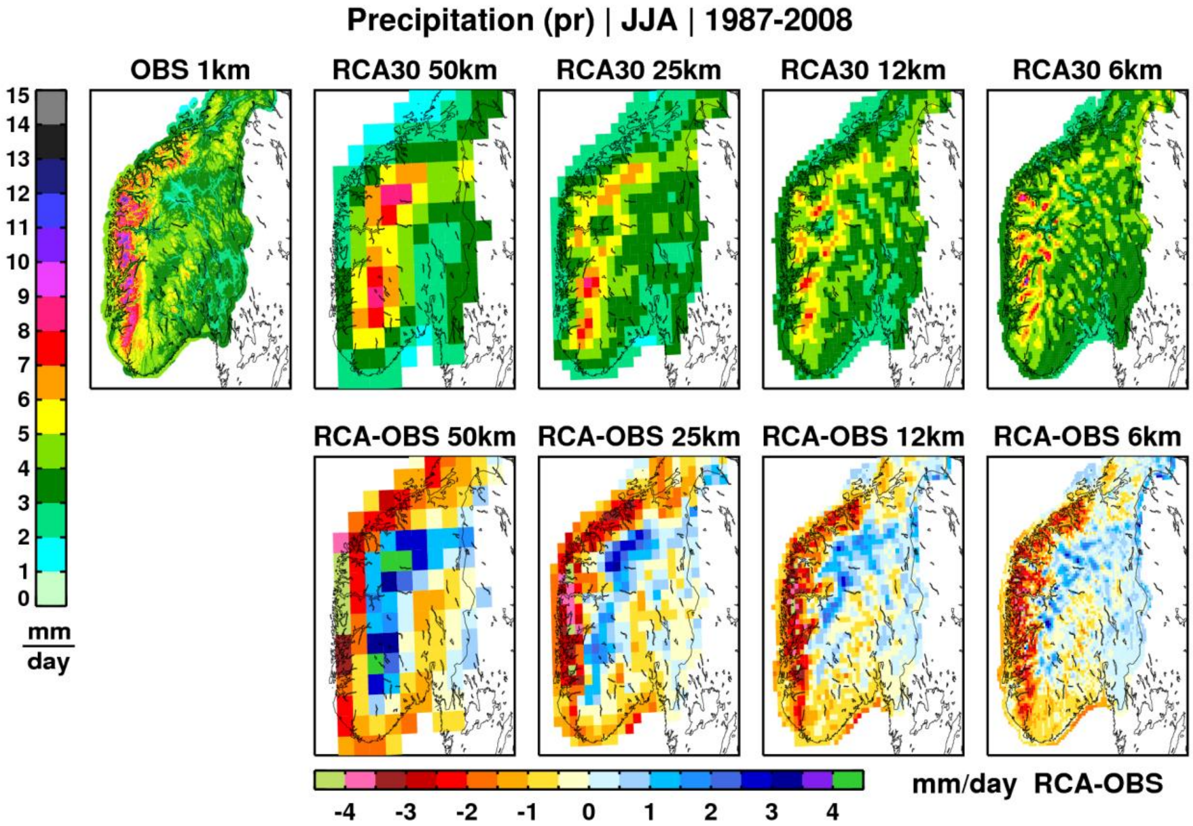


Figure 9.41: Summer seasonal mean (JJA, 1987–2008) in southern-Norway gridded observational precipitation with 1 km resolution from Met.no (Mohr, 2008), and RCM-simulated precipitation with boundary conditions from the ERA40 reanalysis and ECMWF operational analysis (top row). The RCM was run at four different resolutions ranging from 50 to 6 km. Differences between the simulated precipitation and the gridded observations aggregated from 1 km to respectively 50, 25, 12 and 6 km grids are shown in the bottom row. After Güttler et al. (2012).

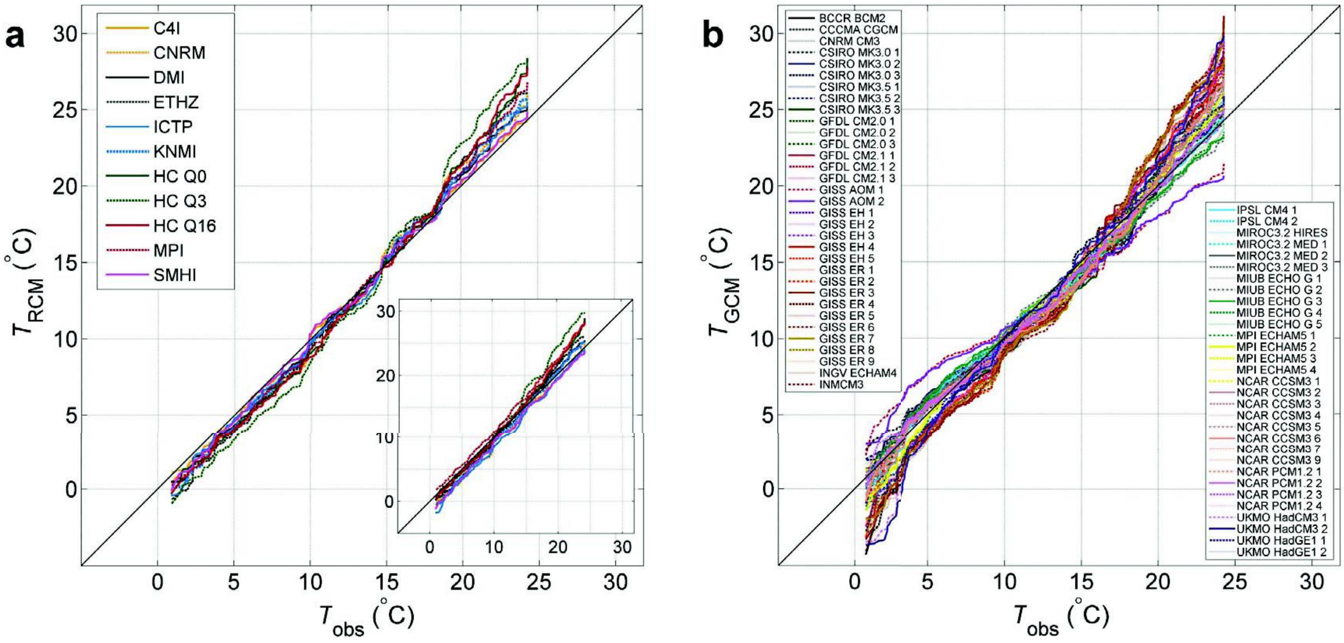


Figure 9.42: Ranked modelled versus observed monthly mean temperature for a Mediterranean region for the 1961–2000 period. The RCM data (panel a) are from Christensen et al. (2008) and are adjusted to get a zero mean in model temperature with respect to the diagonal. The smaller insert shows uncentred data. The GCM data in panel b are from CMIP3 and adjusted to get a zero mean in model temperature with respect to the diagonal. Figure after Boberg and Christensen (2012).

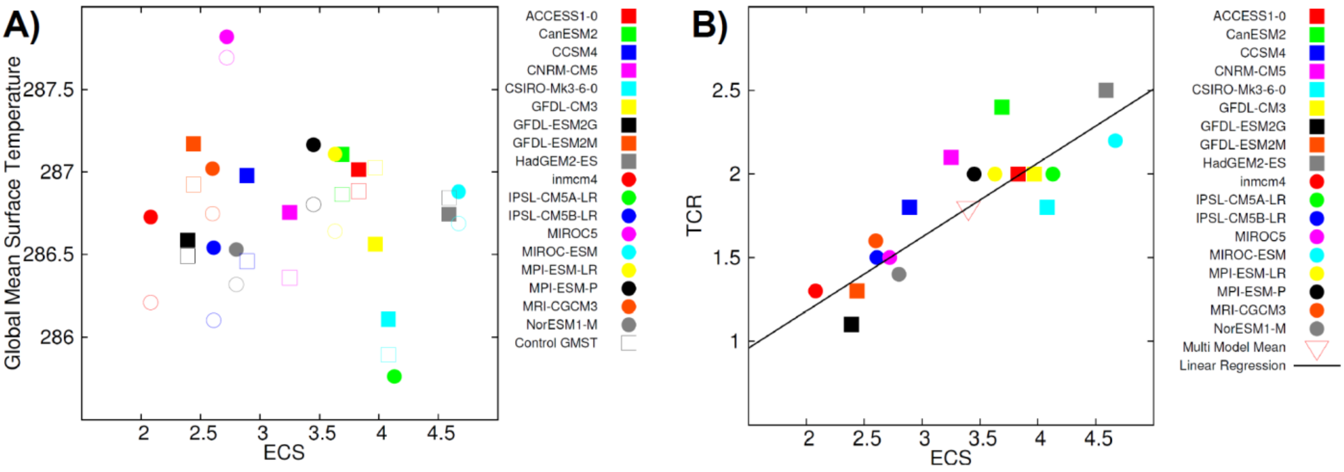


Figure 9.43: A) Equilibrium climate sensitivity (ECS) against the global mean surface air temperature of CMIP5 models for the period of 1961 to 1990. B) Equilibrium climate sensitivity against transient climate response (TCR). The ECS and TCR information is taken from Andrews et al. (2012b).

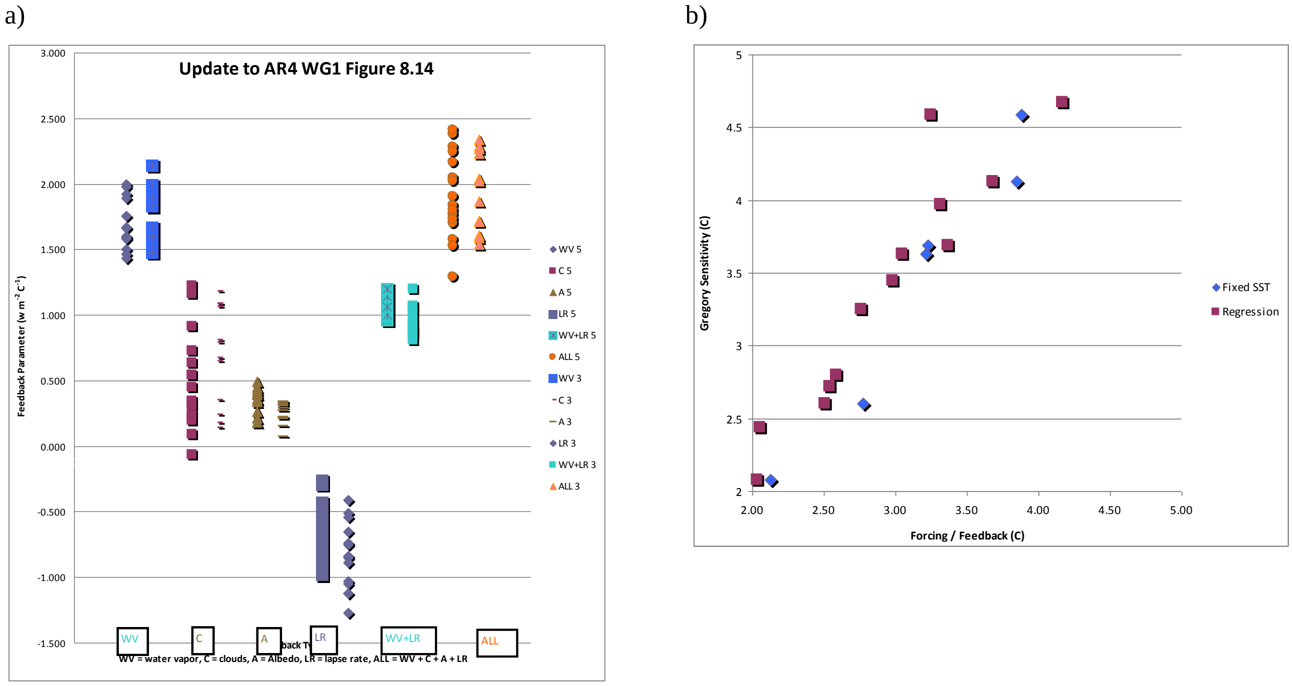


Figure 9.44: a) Feedback parameters for CMIP3 and CMIP5 models (left and right columns of symbols) for water vapour (WV), clouds (C), albedo (A), lapse rate (LR), combination of water vapour and lapse rate (WV+LR), and sum of all feedbacks (ALL) updated from Soden and Held (2006). CMIP5 feedbacks are derived from CMIP5 simulations for abrupt four-fold increases in CO_2 concentrations (4 × CO_2). b) ECS obtained using fixed-SST and regression techniques by Andrews et al. (2012b) against ECS estimated from the ratio of CO_2 radiative forcing to the sum of all feedbacks. The CO_2 radiative forcing is one-half the 4 × CO_2 forcings from Andrews et al. (2012b), and the sum of feedbacks (ALL + Planck) is updated from Soden and Held (2006).

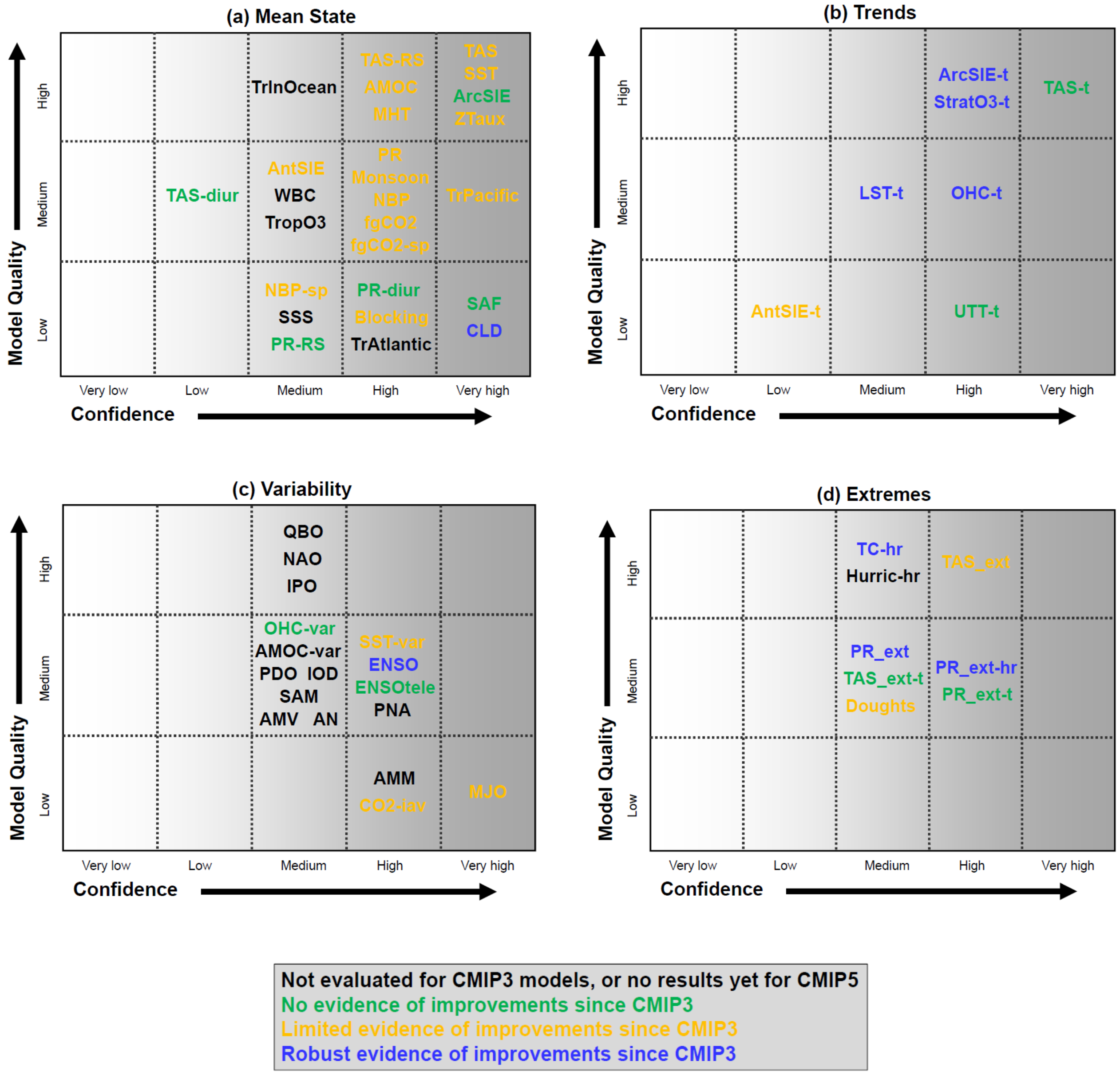


Figure 9.45: Summary of the findings of Chapter 9 with respect to how well the CMIP5 models simulate important features of the climate of the 20th century. Confidence in the assessment increases towards the right as suggested by the increasing strength of shading. Model quality increases from bottom to top. The color coding indicates improvements from CMIP3 (or models of that generation) to CMIP5. The assessment is mostly based on the multi-model mean, not excluding that deviations for individual models could exist. Note that assessed model quality is simplified for representation in the figure and it is referred to the text for details of each assessment. The figure highlights the following key features, with the sections that back up the assessment added in brackets:

PANEL a:

AMOC	Atlantic Meridional Overtuning Circulation mean (Section 9.4.2.3)
AntSIE	Annual cycle Antarctic sea ice extent (Section 9.4.3)
ArctSIE	Annual cycle Arctic sea ice extent (Section 9.4.3)
Blocking	Blocking events (Section 9.5.2.2)
CLD	Clouds (Section 9.4.1.1.2)
fgCO2	Global ocean carbon sink (Section 9.4.5.3)
fgCO2-sp	Spatial pattern of ocean-atmosphere CO_2 fluxes (Section 9.4.5.3)
MHT	Meridional heat transport (Section 9.4.2.4)
Monsoon	Global monsoon (Section 9.5.2.3)
NBP	Global land carbon sink (Section 9.4.5.3)
NBP-sp	Spatial pattern of land-atmosphere CO_2 fluxes

PR	Large scale precipitation (Section 9.4.1)
PR-diur	Diurnal cycle precipitation (Section 9.5.2.2)
PR-RS	Regional scale precipitation (Section 9.6.1.1)
SAF	Snow albedo feedbacks (Section 9.8.3)
SSS:	Sea surface salinity (Section 9.4.2.1)
SST:	Sea surface temperature (Section 9.4.2.1
TAS:	Large scale surface air temperature (Section 9.4.1)
TAS-diur	Diurnal cycle surface air temperature (Section 9.5.2.1)
TAS-RS	Regional scale surface air temperature (Section 9.6.1.1)
TropO3	Tropospheric column ozone climatology (Section 9.4.1.3.5)
TrAtlantic	Tropical Atlantic / Pacific mean state (Section 9.4.2.5)
TrInOcean	Tropical Indian Ocean mean state (Section 9.4.2.2.5)
TrPacific	Tropical Pacific mean state (Section 9.4.2.5)
WBC	Western boundary current (Section 9.4.2.3)
ZTaux	Zonal mean zonal wind stress (Section 9.4.2.4)

PANEL b (Trends)

AntSIE-t:	Trend in Antarctic sea ice extent (Section 9.4.3)
ArctSIE-t:	Trend in Arctic sea ice extent (Section 9.4.3)
LST-t	Lower stratospheric temperature trends (Section 9.4.1.3.5)
OHC-t	Global ocean heat content trends (Section 9.4.2.2)
StratO3-t	Total column ozone trends (Section 9.4.1.3.5)
TAS-t	Surface air temperature trends (Section 9.4.1)
UTT-t	Upper tropospheric temperature trends (Section 9.4.1.3.2)

PANEL c (Variability)

AMM	Atlantic Meridional Mode (Section 9.5.3.3)
AMOC-var	Atlantic Meridional Overtuning Circulation (Section 9.5.3.3)
AMV	Atlantic Multi-decadal Variability (Section 9.5.3.3)
AN	Atlantic Niño (Section 9.5.3.3)
CO2-iav	Interannual variability of atmospheric CO2 (Section 9.4.5)
ENSO	El Niño Southern Oscillation (Section 9.5.3.4)
ENSOtele	Tropical ENSO teleconnections (Section 9.5.3.5)
IOD	Indian ocean dipole (Section 9.5.3.4)
IPO	Indian ocean dipole (Section 9.5.3.4)
MJO	Madden Julian Oscillation (Section 9.5.2.2)
NAO	North Atlantic Oscillation (Section 9.5.3.2)
OHC-var	Global ocean heat content variability (Section 9.4.2.2)
PDO	Pacific Decadal Oscillation (Section 9.5.3.4)
PNA	Pacific North American (Section 9.5.3.5)
QBO	Quasi-Biennal Oscillation (Section 9.5.3.5)
SAM	Southern Annual Mode (Section 9.5.3.2)
SST-var	Global sea surface temperature variability (Section 9.5.3.1)

PANEL d (Extremes):

Hurric-hr	Year-to-year counts of Atlantic hurricanes in high-resolution AGCMs (Section 9.5.4.3)
PR_ext	Global distributions of precipitation extremes (Section 9.5.4.2)
PR_ext-hr	Global distribution of precipitation extremes in high-resolution AGCMs (Section 9.5.4.2)
PR_ext-t	Global trends in precipitation extremes (Section 9.5.4.2)
TAS_ext	Global distributions of surface air temperature extremes (Section 9.5.4.1)
TAS_ext-t	Global trends in surface air temperature extremes (Section 9.5.4.1)
TC-hr	Tropical cyclone tracks and intensity in high-resolution AGCMs (Section 9.5.4.3)
Droughts	Droughts (Section 9.5.4.3)

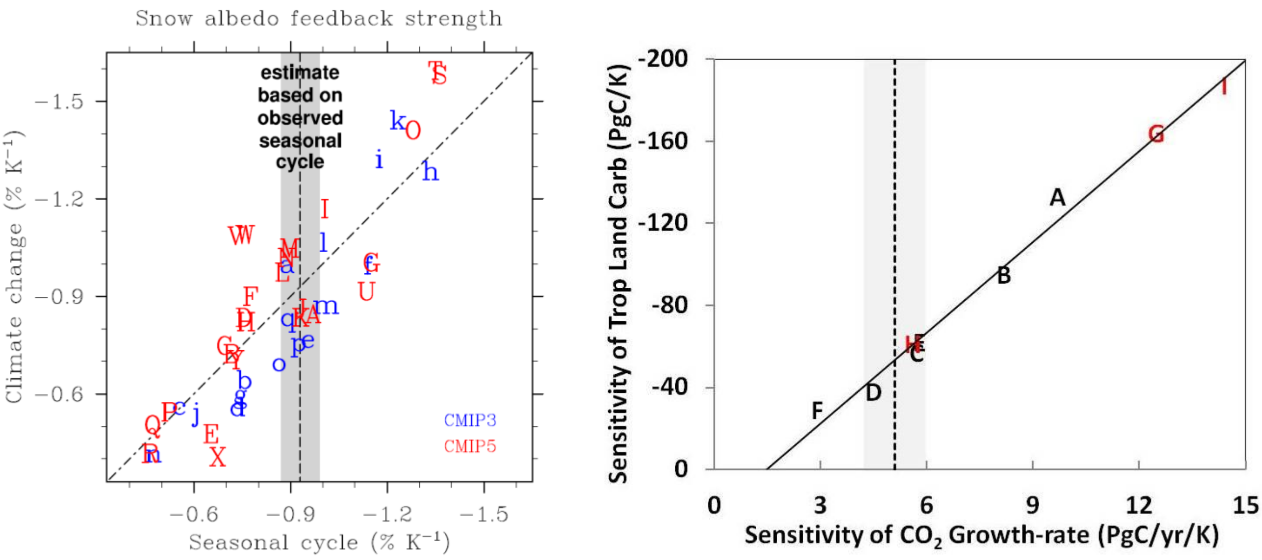
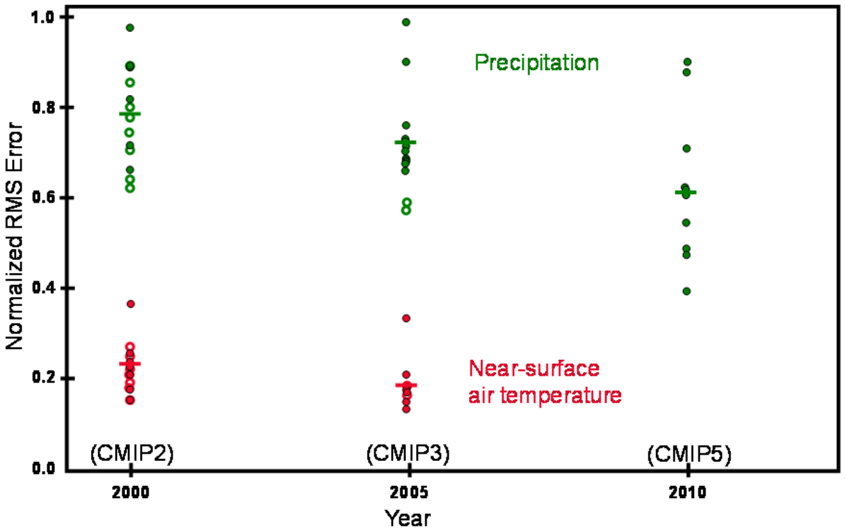


Figure 9.46: *Left:* Scatter plot of simulated springtime snow albedo feedback ($\Delta\alpha_s/\Delta T_s$) values in climate change (y-axis) versus simulated springtime $\Delta\alpha_s/\Delta T_s$ values in the seasonal cycle (x-axis) in transient climate change experiments from 17 CMIP3 (blue) and 25 CMIP5 models (α_s and T_s are surface albedo and surface air temperature, respectively). Adapted from Hall and Qu (2006) and Qu and Hall (2012). *Right:* Constraint on the climate sensitivity of land carbon in the tropics (30°N–30°S) from interannual variability in the growth-rate of global atmospheric CO_2 (Cox et al., 2012). This is based on results from ESMs with free-running CO_2; C^4MIP GCMs (black labels, (Friedlingstein et al., 2006)), and three land carbon “physics ensembles” with HadCM3 (red labels, (Booth et al., 2012)). The y-axis is calculated over the period 1960-2099 inclusive, and the y-axis is calculated over the period 1960-2010 inclusive. In both cases the temperature used is the mean (land+ocean) temperature over 30°N–30°S. The vertical grey band shows the estimated sensitivity of the observed global CO_2 growth-rate to the observed tropical mean temperature.



Preliminary **DRAFT** – will be redrawn when more complete results are available

FAQ 9.1, Figure 1: Model error in simulating annual mean temperature and precipitation as produceed in the three recent phases of the Coupled Model Intercomparison Project (CMIP2, CMIP3 and CMIP5). The upper portion of the figure shows the root mean squared error (a measure of local discrepancies between model and observation) that has been normalized by the observational standard deviation to allow comparison across variables. Each symbol represents the result from a particular model. Larger values indicate larger errors; smaller errors going from left to right indicate model improvement. Across the bottom of the figure, the sketches indicate evolution in model complexity and model resolution, both of which have improved from CMIP2 to CMIP5. [PLACEHOLDER FOR FINAL DRAFT: to be updated as more results from CMIP5 simulations are analyzed and a consistent observational reference is used in the error calculation.] Redrafted from Gleckler et al. (2008) and updated with CMIP5 results.