Astron. Nachr./AN **324**, No. 4, 383–387 (2003) / **DOI** 10.1002/asna.200310138

Interpretation of observations by inversion

J.C. DEL TORO INIESTA

Instituto de Astrofísica de Andalucía (CSIC), Apdo. de Correos 3004, E-18080 Granada, Spain

Received 19 September 2002; accepted 26 September 2002; published online 20 May 2003

Abstract. The most recent developments in inversion techniques of the radiative transfer equation are critically reviewed and some of their findings are summarized to illustrating their achievements. Two significantly different approaches are currently being used that deserve consideration, each characterized by whether or not the model solar atmospheres are changed iteratively by the algorithm. The comparison between the two may help in finding future inversion techniques that can solve many challenging problems of solar physics that still need to be properly settled. These problems themselves suggest strategies that look more suitable than others.

Key words: Sun: magnetic fields – methods: numerical – techniques: polarimetric – techniques: spectroscopic

1. Introduction

The main aim of inversion techniques (ITs) is, of course, the inference of physical quantities that characterize the atmosphere of the Sun. Since we cannot measure directly the temperature, magnetic and velocity fields, etc., we must calibrate our real measurements. We measure light; the spectrum and state of polarization of solar light (a recent, interesting review on measurements can be found in Bellot Rubio 2003). Thus, we need some transformation rules that translate these measurements into temperatures, velocities, etc. Such calibrations have been carried out traditionally by many different forms that can be classified in four great categories: the application of direct formulas, as in the determination of velocities from Doppler shifts; the use of look-up tables, as when trying to assign heights of formation to given measurements; the forward modeling techniques, that is, the spectral synthesis methods; and the inversion techniques, i.e., automated procedures that invert somehow the radiative transfer equation. These astrophysical inferences share the same main drawback: the calibration depends on the physical model assumed for the solar photosphere. I mean that all techniques depend on whether one or several components are assumed, on whether each component is characterized by constant quantities or quantities variable with height, on whether radiative transfer assumptions like LTE are valid, etc. We are never sure if our interpretation corresponds to reality. This is what I call the astrophysical main risk! The only means we have to check the realism of our interpretation is the predictions that can be made with our model and contrasted with further observations.

Correspondence to: jti@iaa.es

Time has imposed the order in the list of inference categories above. This fact stems on an intrinsic reason that can be said "economic": with the newly developed techniques we obtain more results at yet affordable costs. The computing capabilities have increased so much that we are now able to ask the numerical codes for the value of more and more atmospheric quantities at still reasonable computer times. Therefore, we seem invited to continue in that direction by progressively using more and more ITs and continuously developing more sophisticated and efficient techniques.

Since there have been quite a few recent reviews on the subject (del Toro Iniesta & Ruiz Cobo 1995,1996,1997; Socas-Navarro 2001), I assume that most of the fundamentals of each inversion technique are already familiar to the reader. Hence, instead of describing them, I will begin in Sect. 2 by citing the most important results obtained during the last four years by each of them. The astrophysical importance of the results will illustrate the relevance of the techniques. The choice for a four year review is, of course, absolutely arbitrary. Previous results and developments can be found in the reviews cited above. Sect. 3 summarizes the main characteristics and features of the various techniques that can be grouped into two main classes. Such a summary will lead to a discussion about the advantages and drawbacks of each class. In Sect. 4, I conclude by giving my personal view and recommendations for the future developments in the field of interpretational techniques.

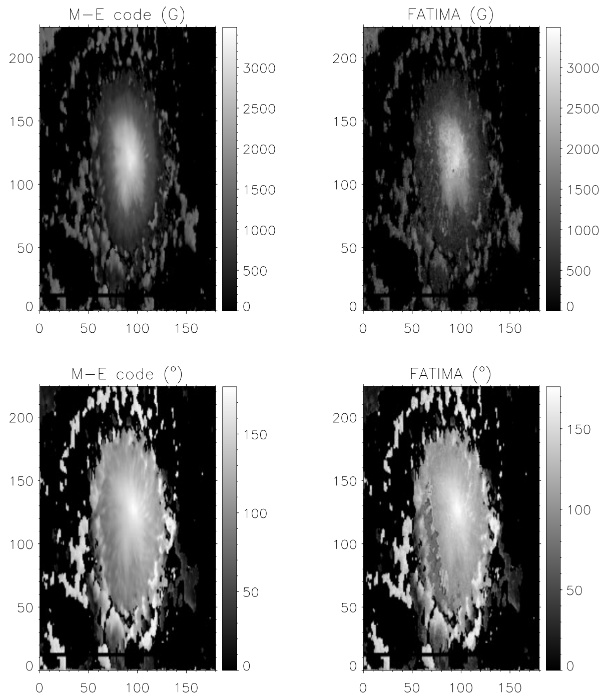

Fig. 1. Principal component analysis inversion of a sunspot. Top panels: field strength; bottom panels: inclination angle. Adapted from Socas Navarro, López Ariste, & Lites (2001) with permission of the copyright holders.

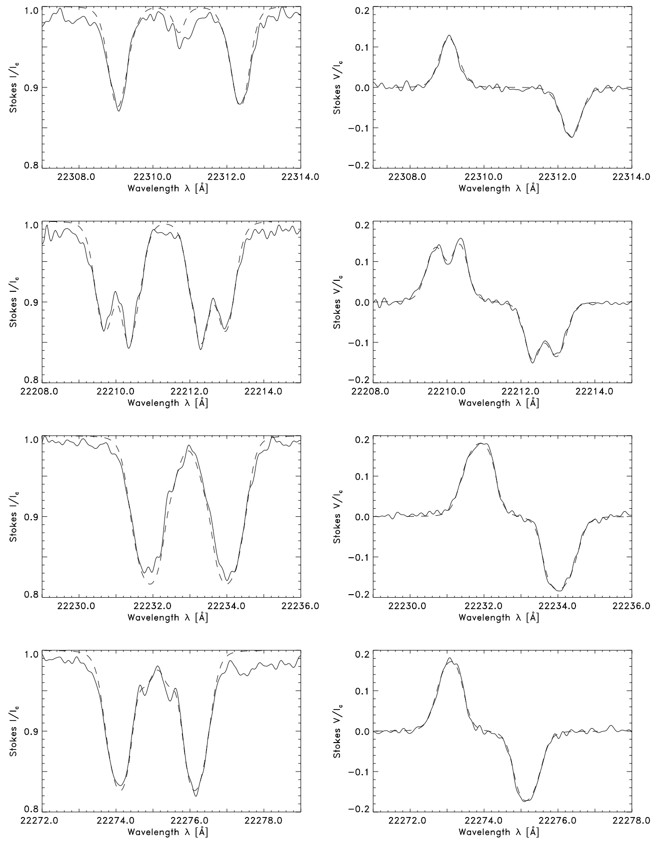

Fig. 2. Ti I lines at 2.2 μm. Observations are in solid lines and fits from the inversion in dashed lines. Adapted from Rüedi et al. (1998) with permission of the copyright holders.

2. Recent results obtained with inversion techniques

The most classical inversion technique (among those still in use) is the Milne-Eddington (M-E) technique of the HAO (Skumanich & Lites 1987). With it, Lites et al. (1998) have measured the oscillations of the vector magnetic field in a sunspot umbra for the first time, although their marginal results indicate a most probable instrumental origin for the weak oscillations detected. Lites, Skumanich, & Martínez Pillet (1998) have carried out a thorough study of the emergence of fresh magnetic flux to the photosphere. They have been able to build a very exciting, “virtual movie” of such an emergence in which weak, horizontal magnetic elements ascend to separate from the emergence zone by strengthening and becoming more vertical. Furthermore, Skumanich (2001) is currently engaged on improving the technique to include gradients of the magnetic field and of the line-of-sight velocity.

The principal component analysis (PCA) was introduced by Rees et al. (2000) and has already produced some results like those in Fig. 1 that represents the first application to a whole sunspot, and under the M-E assumption for the forward problem (Socas Navarro, López Ariste, & Lites 2001). In this comparison with the M-E technique, one can appreciate that the behavior of the new method, as compared to the standard one, is more than satisfactory, with a spectacular reduction in computing time. Only in those zones close to the neutral line, where the most strange and even bizarre Stokes profiles appear, the technique seems to fail, but this is normal because the M-E model is itself unable to reproduce such profiles. PCA implies an expansion series in terms of given *eigenprofiles*. The physical content of the leading orders of this expansion series has been studied by Skumanich & López Ariste (2002). They conclude that practically all the terms have a clear physical correspondence, like the magnetic field strength, the line-of-sight velocity, etc. Socas-Navarro & Sánchez Almeida (2002) have also used the PCA technique together with the MISMA[1] hypothesis and have found intense magnetic fields practically everywhere in network and internetwork regions. Finally, López Ariste & Casini (2002) and López Ariste, Casini, & Tomczyk (2002) are beginning to deal with the forward problem of He I D_3 line formation in prominences and of the hyperfine structure of several lines to use them within the PCA inversion scheme.

Rüedi et al. (1998; see Fig. 2) presented some amazing observations of the Stokes V profiles of the Ti I lines at 2.2 μm and their reproduction with the Zürich IT (Solanki, Rüedi, & Livingston 1992). Also with these lines and the same IT, Rüedi et al. (1999) inferred a cool Evershed channel with a magnetic field weaker than the penumbral background. Frutiger & Solanki (2001) have studied in detail the structure of the small-scale flux tubes. They and their co-workers (Frutiger et al. 2000) have also studied the stratification with

[1] MISMA is an acronym for MIcro-Structured Magnetic Atmosphere.

depth of the granulation. Finally, Mathew et al. (2003) are studying different aspects of the structure of sunspots.

SIR[2] (Ruiz Cobo & del Toro Iniesta 1992) has been adjusted to deal with stars and, in particular, with the Sun as a star, by Allende Prieto et al. (1998). Rodríguez Hidalgo et al. (2001) were able to infer the stratification with depth of the solar oscillations in both a Lagrangian and a Eulerian reference frames. The first optical tomography of a sunspot has been obtained by Westendorp Plaza and co-workers (1998, 2001a, 2001b). The discovery of Evershed downflows (Westendorp Plaza et al. 1997) has been confirmed with this tomography and with the results by del Toro Iniesta, Bellot Rubio, & Collados (2001) who detected supersonic Evershed downflows in cool penumbral tubes. Very recently, Bellot Rubio et al. (2002) and Balthasar et al. (2003a,b) are working on the structure of the penumbra by inverting spectropolarimetric data from the Tenerife Infrared Polarimeter. Bellot Rubio et al. (2000) have found convincing evidence of magnetic oscillations over the photosphere of an umbra. Borrero & Bellot Rubio (2002) have reproduced the intensity profiles of 22 Fe I lines in the quiet Sun; their model also explains the line shifts and equivalent widths of some other 800 lines.

SIR has evolved and been adapted in two ways. The first is an inversion technique including the thin flux tube model and the second is aimed at dealing with lines formed in non-LTE conditions. Recent results about the structure of flux tubes have been obtained by Bellot Rubio, Ruiz Cobo, & Collados (1999, 2000a, 2000b). They have also gathered evidence of convective collapse on the quiet Sun (Bellot Rubio et al. 2001). The introduction and the first application of the non-LTE code (Socas-Navarro, Ruiz Cobo, & Trujillo Bueno 1998; Socas-Navarro, Trujillo Bueno, & Ruiz Cobo 2000) have been very successful in reproducing Stokes profiles of infrared lines from Ca II and in obtaining a model for chromospheric umbral oscillations (Socas-Navarro, Trujillo Bueno, & Ruiz Cobo 2001).

The inversion technique developed by Sánchez Almeida (1997) to give account of the MISMA model has been applied by Sánchez Almeida & Lites (2000) and they have reproduced all type of profiles from network and internetwork regions.

The newest inversion technique is based on artificial neural networks (ANN). The first introduction was carried out by Carroll et al. (2001) and the first tests of a code were published by Carroll & Staude (2001, 2003; see also Socas-Navarro 2002). The technique looks very promising with even faster runs than those of PCA. An example of their performance can be seen in Fig. 3 where a comparison is presented between the true and the retrieved magnetic field strengths from numerical experiments of the technique.

In my opinion, the Community Inversion Codes project of the HAO is a certainly laudable initiative through which the community will have a series of general-purpose, well documented codes with different inversion techniques. The three first of the set are LILIA which is similar to SIR in philosophy but which is based on a different iterative algorithm,

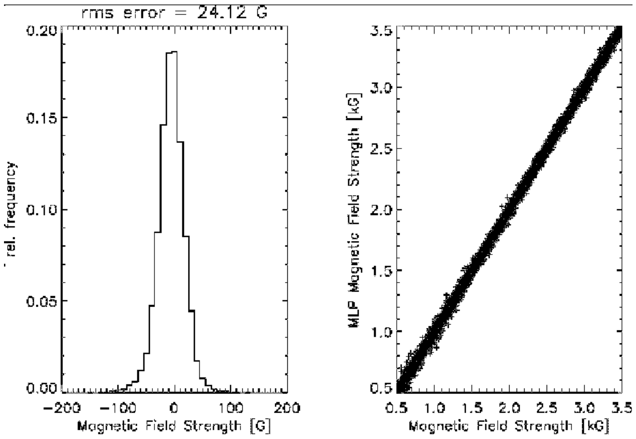

Fig. 3. Distribution of deviations from true values (left) and comparison (right) between true and retrieved values by the neural networks of Carroll & Staude (2001). Adapted with permission of the copyright holders.

FATIMA which is a code based on the principal component analysis, and NICOLE, a non-LTE inversion code based on the same algorithm as LILIA.[3] If these codes spread through the community we will be able to share common means of analysis. Will this be possible, or will everyone continue using his/her own technique? This is a question that I leave open for discussion. Is it indeed convenient at all?

3. Choosing among the available techniques

All the inversion techniques mentioned so far can be grouped in two broad classes depending on the use they make of the atmospheric model (see Table 1). The first class includes the M-E- or SIR-like techniques that use a closed cycle in which the model atmosphere is iteratively modified. The second class encompasses the PCA- or neural-network-like techniques. These techniques do not perform a closed cycle: the retrievals come from a search on a database or from a learning process with a previously created database.

Therefore, there is plenty of techniques, approaches, and flavors within each approach. Then, one should choose among all the possibilities or even embark in the development of new ones. I do not pretend to establish an unambiguous characterization of the different techniques that leads the reader to use a given code. Of course, there is a lot of variables or features to consider and also, why not?, personal preferences. Table 2 lists some features that deserve consideration when deciding which IT is more convenient for a given problem. Obviously, the second class wins in speed. There is no doubt. You may take some time in preparing the database but the analysis of the observations can be carried out quasi on line in some cases. There is a dramatic decrease in computing time. From the computational point of view, Class II

[2] SIR is an acronym for Stokes Inversion based on Response functions.

[3] LILIA is an acronym for LTE Inversion based on the Lorien Iterative Algorithm; FATIMA is an acronym for Fast Analysis Technique for the Inversion of Magnetic Atmospheres; and NICOLE is an acronym for Non-LTE Inversion COde based on the Lorien Engine.

Table 1. Currently available inversion techniques

Class I	Class II
HAO M-E (Skumanich & Lites 1987)	**PCA** (Rees et al. 2000)
IAC SIR (Ruiz Cobo & del Toro Iniesta 1992)	**HAO FATIMA** (Socas-Navarro, López Ariste, & Lites 2001)
ETH IT (Solanki, Rüedi, & Livingston 1992)	**ANN** (Carroll et al. 2001)
IAC MISMA (Sánchez Almeida 1997)	**HAO ANN** (Socas-Navarro 2002)
IAC Flux tube IT (Bellot Rubio, Ruiz Cobo, & Collados 1997)	
ETH Flux tube IT (Frutiger & Solanki 1998)	
IAC NLTE IT (Socas-Navarro, Ruiz Cobo, & Trujillo Bueno 1998; Socas-Navarro, Trujillo Bueno, & Ruiz Cobo 2000)	
HAO LILIA (Socas-Navarro 2001)	
HAO NICOLE (Socas-Navarro 2001)	

Table 2. Choosing inversion techniques

Feature	IT Class
Speed	II
Robustness	II, I
Accuracy	I, II
Realism	Model predictions?
Insight	I, II
Utility	II
Unicity	Astrophysical main risk

wins also in robustness: the iterative cycles of Class I techniques run the risk of stopping in local minima that may depend on the initialization; some codes of the first class are more robust than others, but one must realize that PCA always ends with the very same solution because the search is made through the whole database. However, it seems easier for Class I techniques to get more accurate results. The accuracy in the second class depends strongly on the density (and, hence, on the size) of the look-up table. The realism of the physical description is much more difficult to be known. It might only be contrasted if some predictions can be extracted from the retrieved model atmosphere. In any case, questions about realism might be meta-scientific. We do not measure the solar atmosphere. We measure the light that comes over and need models that reproduce the Stokes spectrum, nothing else. The insight or detail in the retrieved model atmospheres is perhaps easily achievable for Class I techniques. Again, the size of the look-up table is determinant and one may need huge databases hardly handled if the main interest is a very complicated physical model with many free parameters. The utility or versatility is on the side of Class II inversions. Techniques of the first class must be specifically designed for one problem or otherwise be extremely lengthy in terms of computer time, whilst a single Class II IT can be used with several physical models. Perhaps, unicity is the most important concern that people has against ITs and it is true that two different inversion techniques with different assumptions and different physical models can eventually get fits to the observed profiles of similar quality. Nevertheless, everybody must recognize that this is not exclusive of ITs but much on the contrary, it is a common feature of the whole of astrophysics. I have called it the astrophysical interpretation's main risk.

4. Discussion and conclusions

Two different atmospheric models may produce the same Stokes profiles (del Toro Iniesta & Ruiz Cobo 1996; Frutiger & Solanki 1998). The differences between the models can be such that even macro-structuring and micro-structuring of the solar atmosphere can be compatible with given observations as it seems to occur in observations of network and internetwork magnetic fields. While most of the magnetic structures are thought strong for the MISMA scenario (above 1 kG; see Socas-Navarro & Sánchez Almeida 2002), the fields are found to be weak with macro-structured scenarios. Other observations, both in the visible and in the IR, interpreted with different techniques conclude on the weakness of these fields (Keller et al. 1994, Lin 1995, Lin & Rimmele 1999, Collados, 2002). The comparison among the different results may shed some light into the problem; I do not want, however, to enter into discussion about this subject but about the unicity of the retrievals.

It is very important to remember that this lack of unicity is shared by practically all astrophysical inferences. Even the simplest measurements of line core Doppler shifts may have different meanings (that is, it can be interpreted differently) depending on whether one assumes that the material velocity is constant with height in the atmosphere or not. While doing astrophysics we all run the risk of believing that we measure the real things but we only measure light. It is crucial to be humble, and mine is a call to humility.

To conclude, it is my feeling that we better learn from Class II techniques and start by separating the wavelength dependence of the Stokes parameters from their dependence on the atmospheric quantities. If we are able to reproduce the Stokes profiles with a few eigenprofiles we are reducing clearly the dimensionality of the problem: instead of reproducing, say, 50 wavelength samples, we will have to reproduce, say, 10 coefficients. There is a neat gain in speed which is very relevant when dealing with huge amounts of data as those likely to come from modern instruments like GREGOR. By proceeding this way, we will be able of thinking of multiple physical scenarios and use them simply through forward modeling that is computationally easier. Perhaps we will still prefer to use Class I techniques but with less data redundancy.

The main problem that we may face is the universality of the databases. The look-up tables built so far are hardly unique and are rather specifically designed. I am still not sure, but a possible solution may be found in forgetting physics and using mathematics for a while: the Hermite functions, the product of the Gaussian by the Hermite polynomials, are a particularly useful basis for expanding the Stokes profiles. As a matter of fact, these functions are a basis of the space of square integrable functions, $\mathcal{L}^2$, to which the Stokes profiles belong. (Well, Stokes Q, U, and V plus Stokes I in depression.)

Acknowledgements. I warmly thank the organizers for the invitation to participate in this very interesting workshop and L.R. Bello Rubio for his careful reading of the manuscript. I would like to dedicate this paper to the memory of Prof. Dr. Egon Horst Schröter, a solar physics champion and a good man. This work has partly been funded by the Spanish Ministry of Science and Technology under Project No. AyA2001-1177.

References

Allende Prieto, C., Ruiz Cobo, B., García López, R.J.: 1998, ApJ 502, 951
Balthasar, H., Bellot Rubio, L.R., Collados, M.: 2003a, A&A, in preparation
Balthasar, H., Bellot Rubio, L.R., Collados, M.: 2003b, AN 324, 390
Bellot Rubio, L.R.: 2003, in: H. Sawaya-Lacoste (ed.), *Magnetic coupling of the solar atmosphere*, ESA SP-505, in press
Bellot Rubio, L.R., Collados, M., Ruiz Cobo, B., Rodríguez Hidalgo, I.: 2000, ApJ 534, 989
Bellot Rubio, L.R., Collados, M., Ruiz Cobo, B., Rodríguez Hidalgo, I.: 2002, NCimC 25, in press
Bellot Rubio, L.R., Rodríguez Hidalgo, I., Collados, M., Khomenko, E., Ruiz Cobo, B.: 2001, ApJ 560, 1010
Bellot Rubio, L.R., Ruiz Cobo, B., Collados, M.: 1997, ApJ 478, L45
Bellot Rubio, L.R., Ruiz Cobo, B., Collados, M.: 1999, A&A 341, L31
Bellot Rubio, L.R., Ruiz Cobo, B., Collados, M.: 2000a, ApJ 535, 475
Bellot Rubio, L.R., Ruiz Cobo, B., Collados, M.: 2000b, ApJ 535, 489
Borrero, J.M., Bellot Rubio, L.R.: 2002, A&A 385, 1056
Carroll, T.A., Balthasar, H., Muglach, K., Nickelt, I.: 2001, in: M. Sigwarth (ed.), *Advanced solar polarimetry. Theory, observation, and instrumentation*, ASP Conf. Ser. 236, 511
Carroll, T.A., Staude, J.: 2001, A&A 378, 316
Carroll, T.A., Staude, J.: 2003, AN 324, 392
Collados, M.: 2002, private communication
Frutiger, C., Solanki, S.K.: 1998, A&A 336, L65
Frutiger, C., Solanki, S.K.: 2001, A&A 369, 646
Frutiger, C., Solanki, S.K., Fligge, M., Bruls, J.H.M.J.: 2000, A&A 358, 1109
Keller, C.U., Deubner, F.-L., Egger, U., Fleck, B., Povel, H.P.: 1994, 286, 626
Lin, H.: 1995, ApJ 446, 421
Lin, H., Rimmele, T.: 1999, ApJ 514, 448
Lites, B.W., Skumanich, A., Martínez Pillet, V.: 1998, A&A 333, 1053
Lites, B.W., Thomas, J.H., Bogdan, T.J., Cally, P.S.: 1998, ApJ 497, 464
López Ariste, A., Casini, R.: 2002, ApJ, submitted
López Ariste, A., Casini, R., Tomczyk, S.: 2002, ApJ, submitted
Mathew, S.K., Lagg, A., Solanki, S.K., Collados, M., Berdyugina, S., Frutiger, C., Krupp, N., Woch, J.: 2003, AN 324, 388
Rees, D.E., López Ariste, A., Thatcher, J., Semel, M.: 2000, A&A 355, 759
Rodríguez Hidalgo, I., Ruiz Cobo, B., Collados, M., Bellot Rubio, L.R.: 2001, ApJ 547, 491
Rüedi, I., Solanki, S.K., Keller, C.U.: 1999, A&A 348, L37
Rüedi, I., Solanki, S.K., Keller, C.U., Frutiger, C.: 1998, A&A 338, 1089
Ruiz Cobo, B., del Toro Iniesta, J.C.: 1992, ApJ 398, 375
Sánchez Almeida, J.: 1997, ApJ 491, 993
Sánchez Almeida, J., Lites, B.W.: 2000, ApJ 532, 1215
Skumanich, A.: 2001, in: M. Sigwarth (ed.), *Advanced solar polarimetry. Theory, observation, and instrumentation*, ASP Conf. Ser. 236, 543
Skumanich, A., Lites, B.W.: 1987, ApJ 322, 473
Skumanich, A., López Ariste, A.: 2002, ApJ 570, 379
Socas-Navarro, H.: 2001, in: M. Sigwarth (ed.), *Advanced solar polarimetry. Theory, observation, and instrumentation*, ASP Conf. Ser. 236, 487
Socas-Navarro, H.: 2002, Neural Networks, submitted
Socas-Navarro, H., López Ariste, A., Lites, B.W.: 2001, ApJ 553, 949
Socas-Navarro, H., Ruiz Cobo, B., Trujillo Bueno, J.: 1998, ApJ 507, 470
Socas-Navarro, H., Sánchez Almeida, J.: 2002, ApJ 565, 1323
Socas-Navarro, H., Trujillo Bueno, J., Ruiz Cobo, B.: 2000, ApJ 530, 977
Socas-Navarro, H., Trujillo Bueno, J., Ruiz Cobo, B.: 2001, ApJ 550, 1102
Solanki, S.K., Rüedi, I., Livingston, W.: 1992, A&A 263, 339
del Toro Iniesta, J.C., Bellot Rubio, L.R., Collados, M.: 2001, ApJ 549, L139
del Toro Iniesta, J.C., Ruiz Cobo, B.: 1995, in: N. Mein, S. Sahal-Bréchot (eds.), *La polarimétrie outil pour l'étude de l'activité magnétique solaire et stellaire*, Obs. de Paris, Paris, p. 127
del Toro Iniesta, J.C., Ruiz Cobo, B.: 1996, SoPh 164, 169
del Toro Iniesta, J.C., Ruiz Cobo, B.: 1997, in: N. Mein, S. Sahal-Bréchot (eds.), *Forum THEMIS. Science with THEMIS*, Obs. de Paris, Paris, p. 93
Westendorp Plaza, C., del Toro Iniesta, J.C., Ruiz Cobo, B., Martínez Pillet, V., Lites, B.W., Skumanich, A.: 1997, Nature 389, 47
Westendorp Plaza, C., del Toro Iniesta, J.C., Ruiz Cobo, B., Martínez Pillet, V., Lites, B.W., Skumanich, A.: 1998, ApJ 494, 453
Westendorp Plaza, C., del Toro Iniesta, J.C., Ruiz Cobo, B., Martínez Pillet, V., Lites, B.W., Skumanich, A.: 2001a, ApJ 547, 1130
Westendorp Plaza, C., del Toro Iniesta, J.C., Ruiz Cobo, B., Martínez Pillet, V.: 2001b, ApJ 547, 1148