Towards using modern data assimilation and weather forecasting methods in solar physics

A.S. Brun[*]

DSM/DAPNIA/SAp, CEA-Saclay, 911191 Gif-sur-Yvette, France

Received 2006 Dec 8, accepted 2007 Jan 16
Published online 2007 Feb 28

Key words Sun: activity – Sun: magnetic fields – Sun: rotation – solar-terrestrial relation – methods: numerical – methods: data analysis

We discuss how data assimilation and forecasting methods developed in Earth's weather prediction models could be used to improve our capability to anticipate solar dynamical phenomena and assimilate the huge amount of data that new solar satellites, such as SDO or Hinode, will provide in the coming years. We illustrate with some simple examples such as the solar magnetic activity cycle, the eruption of CMEs, the real potential of such methods for solar physics. We believe that we now need to jointly develop solar forecasting models, whose purpose are to assimilate observational data in order to improve our predictability power, with "first principle" solar models, whose purpose is to understand the underpinning physical processes behind the solar dynamics. These two complementary approaches should lead to the development of a solar equivalent of Earth's general circulation model.

1 Introduction

Today, everybody is used to turn on his radio, computer or tv set to find out what today's weather is going to be. Few actually realize that in order to predict as accurately as possible the weather over a given geographical area, a sophisticated procedure, called the "analysis cycle", has to be operated in national or international weather centers such as Météo France, the European Center for Medium-Range Weather Forecast (ECMWF) or the US Climate Prediction center (NCEP). In this procedure, huge amounts of observational data points ($N > 5 \times 10^6$/day) have to be assimilated in sophisticated general circulation numerical models of the Earth's atmosphere. These models are continuously run on vector or massively parallel computers and monitored and analysed by engineers and researchers in weather prediction centers in order to deliver accurate weather forecast.

For decades researchers in the field of weather prediction have developed original technics, both fast and reliable, in order to be able to run operationally such numerical models and assimilate huge amounts of data on an intra daily basis (Talagrand 1997; Kalnay 2003). One among many effects that these models have to take into account is the external radiative forcing imposed by the Sun on the Earth's atmosphere and all the dynamical and energetic consequences that follow. Further, the Sun as a magnetically active star has also through its solar wind and energetic plasma eruptions a direct impact on the Earth's magnetosphere and ionosphere and on our technological society. It can impair satellites in orbit, interfere with high frequency radio communication and radars and create massive blackouts in high latitude countries such as in Quebec in 1989. The impact of the magnetic Sun has started to be considered seriously by the weather forecast community in the mid 70's thanks to researchers like Eddy (1975; see also Pap & Fox 2004).

It is thus important to be able to predict and anticipate solar storms, the variation of the solar luminosity (see Eddy 1982; Foukal et al. 2006 for a recent review) and other solar dynamical phenomena, such as the solar magnetic activity if one wants to progress on both the solar dynamics and the predictions of the Earth's weather on different temporal scales (short to long term trends). To succeed in that challenging task we need to develop a global approach that incorporates and link theory, observations, numerical models and modern data assimilation technics, not unlike what our geophysicist and meteorologist colleagues did for weather forecasting. We thus need to favour the development of solar general circulation models (GCM) and progressively incorporate all the known key physical processes at the origin of the fascinating solar dynamics and activity. Of course we need to adapt weather prediction methods in order to be able to anticipate and predict over periods of days, months and even years, solar dynamical phenomena as diverse as the 22-year magnetic cycle and its associated butterfly diagram, subsurface solar weather or coronal mass ejections (CMEs) (Stix 2002).

Some aspects of this integrated approach have started to be addressed in space weather with projects such as CAWSES (Climate and Weather of the Sun-Earth System), or in the several attempts done to predict solar cycle 23 and the soon starting cycle 24 (Hathaway et al. 1999, 2004). The recent launch of many solar satellites such as Hinode,

[*] Corresponding author: sacha.brun@cea.fr

STEREO or SDO by the space agencies, demonstrate that the solar dynamics and its relationship with the Earth are becoming more and more important to understand, to model and to predict. These new satellites will try to map in 3D the solar surface and corona phenomena in order to progress in for example the triggering mechanism of CMEs. Unfortunately they do not yet form the observational network that one would need to assimilate daily data into a solar GCM.

The Sentinels project submitted to NASA's international program Living With a Star, constitutes with its six satellites (4 probes within Mercury's orbit, one near the Earth and one on the far side of the Sun) a first step toward developing an observational and operational network of the Sun. However, most of these observations are mainly concerned by surface magnetic activity. But it is fundamental to realize that this surface phenomena are intimately linked to the solar internal magnetism through dynamo action (see Sect. 3).

To progress in our ability to couple internal and external magnetic and turbulent phenomena, we also need to progress in our understanding of the inner working of the Sun. The helioseismology community, has made real progress towards putting more constraints on the internal solar dynamics, by inferring the depth of the solar convection zone, its internal rotation profile and near surface weather like pattern. The launch of SDO in early 2008 with the HMI experiment on board will provide continuous and accurate data for cycle 24 and the solar surface dynamics, such as torsional oscillations or meridional flows. For the deeper solar layers the project DynaMICS (Turck-Chieze et al. 2005) will study the dynamical link between inner magnetism in the solar radiative zone and the solar cycle and dynamo generated fields by coupling global seismology and surface observations.

In the meanwhile, real progresses are being made to develop realistic numerical models of the solar convection zone, internal dynamics and dynamo. It is thus clear that most of the pieces constituting the solar puzzle, i.e. observational data of the solar surface and interior (through helioseismology), theoretical studies, numerical tools and preliminary assimilation and forecasting methods are already partly there and they just need to be put together in a more coherent and organised way. Clearly an integrated approach, using both data assimilation/forecasting methods and dedicated numerical models of key physical ingredients is needed, in order to develop a fully coupled model of the Sun, from its inner core up to its extended corona.

The paper is organised in the following manner. In Sect. 2, we briefly present recent data assimilation techniques used in weather forecast. In Sect. 3 we discuss the current status of data assimilation and forecast models in solar physics. We argue that the solar dynamical phenomena share common attributes with the Earth's atmosphere, and solar prediction could certainly benefit from more advanced data assimilation techniques providing that we adapt them to our specific problems, such as taking into consideration magnetic fields. Further, in section Sect. 4 we present recent 3-D MHD models of the solar interior and use these models to put constraints on key physical ingredients, such as differential rotation or the tachocline. Finally we conclude in Sect. 5.

2 A brief overview of data assimilation methods in weather prediction

In modern meteorology, numerical global circulation models (GCM) solving so called *primitive* equations that are simplifications of the equations of fluid mechanics (i.e continuity, Navier-Stokes, energy, Pedlosky 1987) for a stratified fluid are integrated in space and time. These models are regularly updated by assimilating a subset of more than 10^6 observational data points of the state of the atmosphere that are provided by balloon probing, satellites, ground devices, aircraft, etc... In order to update the numerical model, increasingly sophisticated data assimilation techniques in weather forecast have been developed over the last 5 decades. A useful definition of data assimilation can be found in Talagrand (1997): "using all available information, to determine as accurately as possible the state of the atmospheric (or oceanic) flow."

Modern data assimilation technics rely on statistical estimation theory, such as least squares methods. The generalisation of such statistical methods to multivariate systems, lead to what is called the optimal interpolation (OI) for data assimilation (Lorenc 1981). Optimal interpolation consists of taking into account (assimilate) the new information that the observational data provide in order to advance in time the "background" state (also called first guess or prior information) that the weather forecasting numerical code has predicted. The increment is obtained by taking the difference or innovation between the observational data and the observational operator. The new state or analysis is then the result of the assimilation/forecast procedure. More specifically, let $\boldsymbol{x}^b$ be the background vector state characterising the current state of the model, $H(\boldsymbol{x}^b)$ the observational operator and $\boldsymbol{y}^o$ the observational data to be assimilated in the model, then one can show that the analysis $\boldsymbol{x}^a$ is

$$\boldsymbol{x}^a = \boldsymbol{x}^b + W(\boldsymbol{y}^o - H(\boldsymbol{x}^b)), \quad (1)$$

with W the weights determined from the estimated statistical error covariances of the forecast and the observations (Kalnay 2003). This equation is the base of modern data assimilation. The various assimilation methods will differ in the way they combine the observations and the background state to generate the analysis.

In practice the background state, the observations and even the numerical model used to simulate the Earth's atmosphere (i.e the primitive equations), possess errors. The assimilation methods consist of predicting the evolution of the errors and of course of minimising it, i.e keeping it under control as much as possible given the very chaotic nature of the Earth's atmosphere. As the famous scientist Lorenz, recently said: "the atmosphere is chaotic: the present determines the future, but the approximate present does not

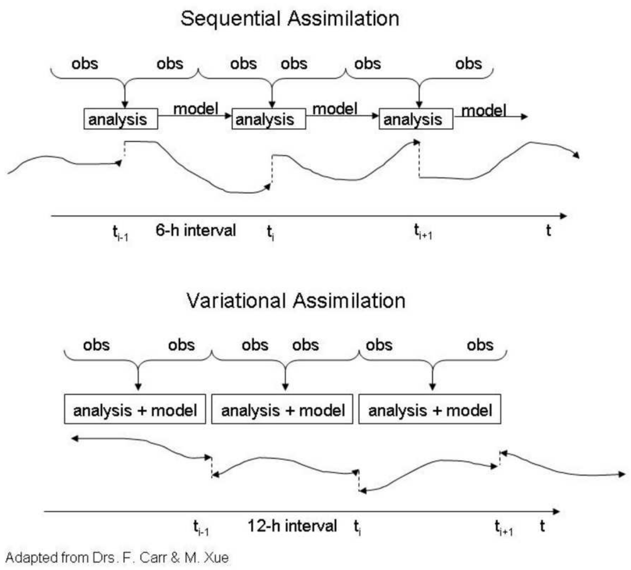

Fig. 1 Schematic representation of the sequential and 4-D variational data assimilation methods used in weather forecast. In the sequential method, the background state is updated every 6 h and the model evolves the state until the next step at which observational data are again assimilated. In the 4-D variational method within a 12 h interval the model and the observations are taken into account in the cost function J that need to be "minimised".

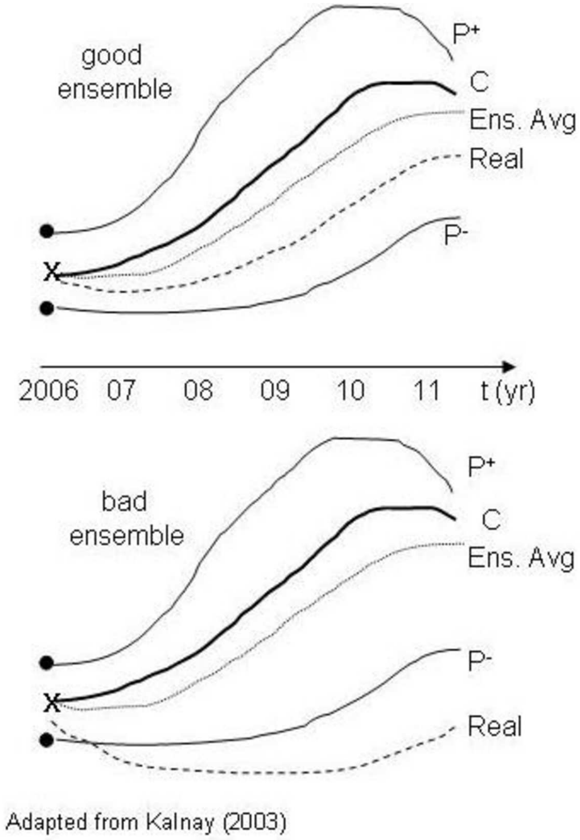

Fig. 2 Schematic representation of a "good" or "bad" ensemble in the case of the solar 11-yr sunspot cycle 24 prediction. The different curves are: C the controlled forecast, P^+ and P^- are positive and negative perturbations of case C, "ens. avg" the average of the models and "real" is reality. In the good ensemble the reality is part of the ensemble forecast.

determine the approximate future". Errors in the dynamical atmospheric system are known to double every two to three days which leads to a predictability limit for weather forecasting that Lorenz in 1963 was the first to quantify to be of the order of 15 days. This is a very strong constraint on our ability to predict weather pattern and solar equivalent predictability limits must exist. However some atmospheric properties may be easier to predict over long period than other, such as weakly average rain falls or temperature, and it is likely that for the Sun some of its characteristics can also be predicted over a longer period of time. Since the error evolution is central for making accurate predictions over the longest period, better our knowledge of the correlation between all types of uncertainties better our forecast will be.

In order to achieve such a goal of optimal control (Lions 1971), data assimilation methods are now commonly split into two categories: sequential or variational (see Fig. 1 and Talagrand 1997; Daley 1991; Kalnay 2003). In the sequential methods, such as OI or Kalman filter, observational data are assimilated in the numerical model at fixed time, say every 6 hours, and then evolved forward in time. In the so called 4-D variational technics, rather than updating the numerical weather model at fixed time, one seeks to minimise a cost function $J(\xi)$ within a certain time interval (usually 12 hours) for which data is available before making a forecast. The procedure converges when J reached its minimum which occurs for $\xi = \boldsymbol{x}^a$ (see Talagrand 2003). Then in the next 12 hours period the procedure is applied again, using as background state the numerical model of the previous 12 hours interval. 4-D-var methods provide an efficient way of improving our error control and has now been adopted operationally in weather prediction centers since the mid 90's.

However variational methods require the development and maintenance of a so-called adjoint model of the dynamical equations considered. This adjoint model computes efficiently the gradient $\partial J/\partial\xi$ necessary to the iterative minimising procedure, by evolving backward the adjoint system of equations from the forward 12 h temporal integration (Talagrand 2003; Kalnay 2003). Developing such a model and procedure is a straightforward but costly task and no such models have been yet developed for the full MHD system of equations (and in particular the induction equation for the magnetic field, see next sections) that is required to model the solar dynamics and magnetic activity. Such efforts to develop the adjoint models of the induction equation are now under way.

The most recent development in modern statistical data assimilation is called *ensemble* forecasting. It consists of running several simulations of the dynamical system with various initial conditions. The result of the different forecasts provide a statistical way of assessing the quality of the forecast and to form an ensemble average forecast. An important notion in ensemble forecasting is the definition of a “good” or ”bad” ensemble (see Fig. 2). In a “bad” ensemble none of the forecasting models correctly predicts the evolution of the dynamical system studied. Conversely, in a “good” ensemble the real evolution of the system is included in the range of predictions made by the different models (Kalnay 2003). A real advantage of these ensemble technics is that they help quantifying the intrinsic error of each individual model. They compare the prediction of one model to the other as well as the individual and ensemble errors with respect to the observations and provide as such a step towards a better control of the error evolution. For instance on weather maps, such an ensemble forecast produces “spaghetti” plots, by having all the predictions of a given isobar overplotted on the same map. These maps clearly indicate where the prediction is reliable (regions where the lines overlap) and where it is less so (large departure from one line to the other) and as such give a good sense / evaluation of what is called the “error of the day”. This can lead meteorological centers to request temporarily extra or finer sampled data points in the area of poor predictability, for example by doing specific balloon probing or aircraft flight over the region of interest in order to improve the forecast. There is thus a real feedback between the forecast model and the observation network.

In the analysis cycle, independently of the assimilation methods used, the model forecast is crucial, because it transports information from data rich regions to data poor regions and it provides a complete estimation of the four dimensional state of the atmosphere. It is thus crucial to further improve and develop the numerical model used to predict the evolution of the atmosphere. This clearly favours a joint approach with on the one hand, the development (improvement) of detailed 3-D numerical models of key physical processes and on the other hand, the operation of simplified forecast models relying both on a reduced set of the dynamical equations and on the assimilation of all the available observations. We refer to the book of E. Kalnay and R. Daley for a much more detailed explanation of the different assimilation technics used in weather forecasting.

3 Few examples of predictable solar phenomena

Being able to anticipate and predict the turbulent solar dynamics and magnetic activity is very important. It is currently thought, that the operation of the solar global dynamo rests on several processes (Moffatt 1978; Charbonneau 2005; Brun et al. 2004): the generation of fields by the intense turbulence of the deep convection zone, the transport of these fields into the tachocline region near the base of the convection zone, the storage and amplification of toroidal fields in the tachocline by differential rotation, and the destabilization and emergence of such fields due to magnetic buoyancy as active regions (Cline et al. 2003; Fisher et al. 1999). Coexisting with these large-scale ordered magnetic structures are small-scale but intense magnetic fluctuations that emerge over much of the solar surface, with little regard for the solar cycle (see Stix 2002; Charbonneau 2005). The observed large diversity of magnetic phenomena must thus be linked to two conceptually different dynamos: a large-scale/cyclic dynamo and a turbulent small-scale one (e.g., Cattaneo & Hughes 2001; Ossendrijver 2003). Unfortunately, self-consistent magnetohydrodynamic (MHD) simulations which realistically incorporate/integrate all of these dynamical processes from the deep layers up to the extended solar corona are not yet computationally feasible. Some specific elements can now be studied with reasonable fidelity (see Sect. 3 and 4) and models of key physical interfaces such as the tachocline (Spiegel & Zahn 1992; Brun & Zahn 2006) or the photosphere (Stein & Nordlund 1998; Vogler et al. 2005) are being pursued. However we need to integrate all these approaches in a more global picture. One way is to develop a solar general circulation model and to apply modern data assimilation to solar problems such as the solar 22-yr cycle and CMEs.

3.1 Solar magnetic cycle prediction

The Sun possesses a large range of magnetic and dynamical phenomena (see above). Most striking is its magnetic cycle of activity which has been observed since the early 1600’s and trace back via ^{10}Be concentration found in ice core data for at least 10 000 years (Beer et al. 1998). The full cycle is 22 years long and it is composed of two consecutive 11-year sunspot cycles (i.e. butterfly diagrams). It has been empirically determined that odd numbered cycles are usually stronger than even numbered cycles, that on average the cycle rises in 4.8 years and falls in 6.2 years, even though strong cycles rise faster to their maximum. An useful quantity to assess the intensity of a cycle is to compute the yearly averaged Wolf sunspot number

$$R = k(10g + s) , \qquad (2)$$

with g being the number of sunspot groups, s the total number of individual sunspots in all the groups, and k is a variable scaling factor (with usually $k < 1$) that accounts for instrumental or observational conditions.

The review of Hathaway et al. (1999) summarises most of the currently used methods to predict the next solar cycle using historical data. Methods such as regression or curve fitting work well near solar maximum while others such as geomagnetic precursors perform better near minimum. They suggest that a synthesis of all these methods can provide a more accurate and useful forecast of solar cycle activity levels and of the evolution of the Wolf numbers. Given the importance of the solar activity, the number of sunspots is recorded every day by more than 25 stations over the world, and the international sunspot number is produced by the SIDC center in Belgium. The study of the most recent cycles reveals that cycle 23 is actually weaker than cycles

22, 21, and 19 the strongest of all, with respectively a maximum Wolf number R of: 119.6 in 2000, 157.6 in 1989, 155.4 in 1979 and 190.2 in 1957. Since cycle 23 was predicted by the solar cycle 23 panel to be slightly stronger ($R \simeq 160$) than cycle 22 but that in reality with an observed value of about 120, it turned out to be almost as weak as the even numbered cycle 20 ($R = 105.9$ in 1968), the predictions of cycle 23 were not as accurate as one could have hoped. Further in the prediction summary of the solar cycle 23 panel only few of the many predictions (even by taking into account their error bars), were actually including the observed value of 120. One thus needs to be careful with the standard indicators used up to now. However, the fact that there is a panel prediction can be seen as an attempt to use ensemble forecasting. The relative success of these methods, in particular for cycles 21 and 22, much less so for cycle 23, could be a sign that the set of equation models used in the panel form a good ensemble as defined earlier. However most of the techniques considered by Hathaway et al. do not resolve the spatial dependence of the solar activity, they just focus on global properties such as number of sunspots or the timing of the next maximum. As such these techniques are much less sophisticated than the one used in weather forecasting. We thus need to develop more physically based forecast models of the solar cycle.

Historically two types of physical models have been developed in order to understand the solar global dynamo: 2-D mean field models and 3-D MHD simulations (Ossendrijver 2003). However none of these models were used, up to very recently, to predict the evolution of the solar cycle or observed butterfly diagram. In order to take into account the spatial dependency of the solar activity, more recent approaches solve numerically the induction equation in a meridional plane and impose through a surface term the latitudinal band of activity (Dikpati & Gilman 2006; Cameron & Schuessler 2007). By assimilating sunspot or meridional flow data they try to predict the next cycle 24. Today the predictions for the next solar cycle recently summarized by the cycle 24 prediction panel, differ quite significantly from one model to another. Some techniques, such as the ones based on geomagnetic precursors, predict a weak cycle 24 ($R < 100$, Svalgaard, Cliver, Kamide 2005; Duhau 2003) other based on dynamo models or meridional flow speed predict a stronger cycle ($R > 140$, Dikpati, Toma & Gilman 2006; Hathaway & Wilson 2004).

It is worth noting that all the predictions for a weak cycle 24 relies on cycle 23, i.e cycle n is mostly correlated with cycle $n-1$, whereas those predicting a strong cycle 24 (i.e stronger than cycle 23) favour a correlation with cycle 22, i.e cycle n is well correlated with cycle $n-2$. The predictions of the cycle 24 panel also differ on the timing of the next maximum, with predictions ranging between 2010 and 2012, depending on how fast the next cycle will rise to reach its maximum (fast if strong, slow if weak). The idea of using a mean field dynamo model to assimilate some of the solar distinct magnetic features to predict the solar activity is interesting since it is one more step toward using the more sophisticated data assimilation used in weather forecast methods. We will see in 4 to 5 years if their forecast capability is indeed better.

3.2 Coronal mass ejection and space weather forecast

At the surface the Sun is turbulent and very active with magnetic phenomena such as flares, prominences, CMEs. This intense activity is known to have a direct impact on Earth's upper atmosphere and on our technological society. This has lead to the development of space weather studies and the beginning of space weather forecast. Answering key questions such as which physical processes lead to eruptive phenomena, what is the associated spectrum of solar energetic particles (SEP) and what leads to geoeffective interplanetary coronal mass ejections (ICMEs) constitute the main purpose of space weather. CMEs take between 1 to 3 days to reach the Earth but relativistic particles can travel 1 AU within few minutes after the CMEs leaving not enough time to anticipate and protect our satellites or astronauts on board the international space station. It is thus crucial to be able to model and predict the onset of CMEs days before they should occur. This implies to develop a deep physical understanding of CMEs and eruptive flares phenomena. Currently it is believed that CMEs are structured in three parts: a bright leading edge, dark cavity and bright dots. The complete physical description of the triggering mechanism of CMEs is still under construction, but it is the most accepted scenario that when a CME takes place, the magnetic field in the corona above the active region undergoes restructuring due to both the continuous evolution of the magnetic field lines that are rooted in the turbulent convection zone and the emergence of new small scale magnetic flux. Multi-wavelength observations have been crucial in this recent progress. Observational programs such as the Whole Month Sun (Gibson et al. 1999; Mikic et al. 2006) devoted to coordinate co-temporal / multi-wavelength observations of eruptive phenomena such as flares or CMEs lead to interesting result. The main results is that reconnection plays a pivotal role in the triggering mechanism of the CMEs.

Two models for explaining the CME onset are currently proposed: In a first model the amount of twist within the filament is key to triggering the eruption by reaching a kink type instability threshold in the highly twisted structure (Amari et al 2003; Fan & Gibson 2004). The second model, rests on the build up of magnetic energy in the sheared magnetic arcade above the active region and do not rely on the magnetic helicity to explain the eruption (Devore & Antiochos 2000). Observations have still difficulties to distinguish between the two models, since the magnetic field topology is hard to be assessed correctly above active regions from far away observations. For example the observations of sigmoid shape in X-rays imaging of pre-CMEs configuration, could favour the first model but CMEs have been observed to erupt without such features (Pevtsov 2002; Rust & Labonte 2005; Leamon et al. 2003; Inobe &

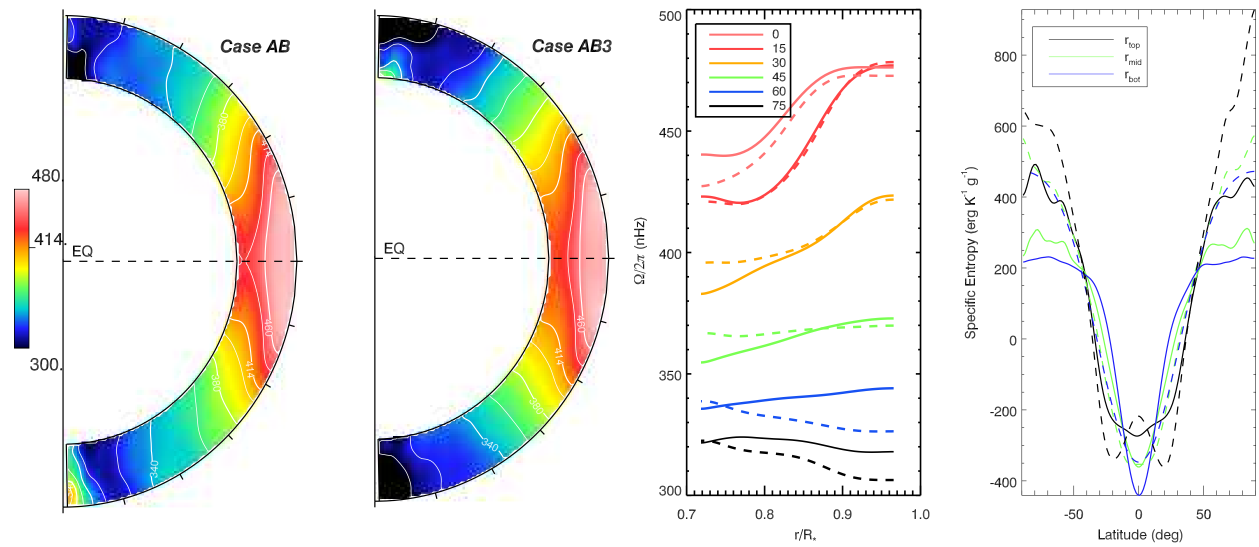

Fig. 3 *Left*: Angular velocity contours for case AB and AB3 along with radial cuts at indicated latitudes (case AB in solid). *Right*: Entropy variations for cases AB & AB3 as a function of latitude for three radial positions at respectively top, middle and bottom of the shell (with again case AB in solid).

Kusano 2006). Recent solar space projects, Hinode, SDO, STEREO, intend to map in 3-D the field topology in order to improve our ability to characterize the field topology near the surface, in the filament and the surrounding magnetic arcades. Given the important role played by the reconfiguration of the coronal field in CMEs, understanding how the corona can sustain the stresses that lead to the increase of magnetic energy and the field reorganisation within pre-CME magnetic configuration is very important. Accurate extrapolation of the magnetic field in the corona is crucial to assess the energy and magnetic helicity budget of pre-CMEs configuration (Carcedo et al 2003). It has been theoretically established that the magnetic energy contained in coronal fields above active regions is bounded within two field topologies: the potential topology of minimal energy and the open magnetic field configuration (also called coronal holes) of maximal energy. In between the field can take a complex topology such as a force free magnetic field (i.e $\boldsymbol{\nabla} \times \boldsymbol{B} = \alpha \boldsymbol{B}$) with non constant α (Regnier & Amari 2003; Shrijver et al. 2006).

Solar eruptive phenomena are associated to active regions, i.e complexes of sunspots, that possess intricate magnetic field topology. There is a direct link between internal magnetism and surface magnetic phenomena, since active regions are related to the emergence of strong toroidal structures most likely generated in the deep solar tachocline of intense latitudinal and radial shear at the base of the CZ (Cline et al. 2003). These toroidal structures become unstable, subsequently rise through the solar convection zone to appear at the surface as active regions (Magara & Longcope 2003; Fisher et al. 1999; Archotis et al. 2005; Jouve & Brun 2006). As we have seen in the previous section, the number of sunspots or group of sunspots is directly connected to the phase of the solar cycle, with the magnetic solar poloidal field (mostly its dipolar component) reversing sign about the time when the solar activity reaches its maximum in the active latitudinal belt. The link between the solar cycle, CMEs and the geoeffectiveness of solar events is not straightforward to assess (Pevtsov & Canfield 2001), but it is clear that one important goal of space weather is to characterize the configurations (strength, location, field topology, etc...) that lead to geoeffective events. For instance, at the maximum of solar activity the heliospheric magnetic field is highly sensitive to the surface solar activity and its interaction with the magnetosphere is quite unpredictable. The alignment (or anti alignment) of the Earth's magnetosphere with the solar dipole during odd (even) numbered cycles has also direct consequence for stronger geomagnetic activity when the fields possess opposite polarity.

Field extrapolation of observed vector magnetograms of given active regions are already used to model CMEs. The field configuration is then followed in time to compute its evolution. Some models can also follow the development of ICMEs that will impact the Earth and lead to intense geoactivity. But no operational CME numerical models are yet assimilating sequentially observations of a given active region as it evolves by moving across the solar surface. It is also possible to compute from extrapolating the photospheric field into the solar corona the polarization brightness, to compare it to intensity observations of the coronal loops, and even for example to predict the coronal emission during a total eclipse from earlier observations (Mikic et al. 1999, 2006). Assimilation of solar data in numerical models has thus already started. However we do not yet run continuously and operationally the solar equivalent to a global circulation model. Intrinsic difficulties in the solar data assimilation and forecasting problems are linked to the fact that we do not have yet a complete comprehension of the solar magnetic dynamo, cycle and surface activity. For every "piece" constituting the full solar puzzle, theoretical developments are still underway.

What could be the next step in solar prediction and data assimilation? Clearly, the solar community should develop more sophisticated data assimilation methods, such as sequential or 4-D var, using as dynamical equation the induction equation (before the full MHD equations). One first easy step could be to introduce by hand an extra term (or "nudging" term, Kalnay 2003) in the right hand side of the induction equation, i.e:

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\boldsymbol{v} \times \boldsymbol{B}) - \boldsymbol{\nabla} \times (\eta \boldsymbol{\nabla} \times \boldsymbol{B}) + \frac{\boldsymbol{B} - \boldsymbol{B}_{obs}}{\tau_B}, \quad (3)$$

with τ_B the correlation time to "nudge" the temporal evolution of $\boldsymbol{B}$ toward observations. The observations could be for example vector magnetograms and the model could assume a force free coronal field. The question then is what to use for the velocity $\boldsymbol{v}$. One possibility is to consider the mean field variant of the induction equation and to run a mean solar dynamo model in the meridional plane, this is the spirit of Dikpati's work. Another, is to assimilate surface phenomena, by having a regularly updated state of the temporal and spatial evolution of a given observed active region in order to predict as accurately as possible if it could erupt, rather than as it is done today, only initiating the 3-D MHD simulations with the observations. Indeed taking into account the evolution of the active region through updated observations could certainly improve our prediction capability (Bélanger et al. 2005).

4 Towards an integrated 3-D MHD model of the turbulent Sun

Another successful approach to improve our understanding of the solar dynamics and activity is to actually compute from first principles 3-D global time dependent models of the solar plasma rather than simplified or reduced version of it as in prediction models. Indeed, 3-D time dependent models allow the detailed study of nonlinear processes and caution must be taken to use simple indicators, as reviewed in Sect. 3.1, to predict the behaviour of chaotic systems like the Sun (Tobias et al. 2006). These 3-D nonlinear numerical models integrate more or less sophisticated variants (such as ideal MHD or two fluids equations) of the following equations:

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0, \quad (4)$$

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) = 0, \quad (5)$$

$$\rho \left(\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \boldsymbol{\nabla}) \boldsymbol{v} \right) = -\boldsymbol{\nabla} P + \rho \boldsymbol{g} - 2\rho \boldsymbol{\Omega_0} \times \boldsymbol{v} + \frac{1}{4\pi} (\boldsymbol{\nabla} \times \boldsymbol{B}) \times \boldsymbol{B} - \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{D}}, \quad (6)$$

$$\rho T \frac{\partial S}{\partial t} + \rho T \boldsymbol{v} \cdot \boldsymbol{\nabla} S = \boldsymbol{\nabla} \cdot (\kappa_{\rm rad} \rho c_p \boldsymbol{\nabla} T) + \frac{4\pi\eta}{c^2} \boldsymbol{J}^2 + 2\rho\nu \left[e_{ij} e_{ij} - 1/3 (\boldsymbol{\nabla} \cdot \boldsymbol{v})^2 \right], \quad (7)$$

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\boldsymbol{v} \times \boldsymbol{B}) - \boldsymbol{\nabla} \times (\eta \boldsymbol{\nabla} \times \boldsymbol{B}), \quad (8)$$

with ρ the density, T the temperature, P the pressure, $\boldsymbol{v}$ the fluid velocity, $\boldsymbol{J} = c/4\pi\,(\boldsymbol{\nabla} \times \boldsymbol{B})$ the current density, g the gravity, Ω_0 the solar rotation rate, S the specific entropy, e_{ij} the strain rate tensor, $\boldsymbol{\mathcal{D}}$ the viscous tensor, η the magnetic diffusivity, ν the kinematic viscosity, $\kappa_{\rm rad}$ the radiative diffusitivity and c_p the specific heat at constant pressure. It then suffices to know two thermodynamic variables, say density and temperature, the fluid velocity, the magnetic field $\boldsymbol{B}$ and an initial time instant, and the thermodynamic relations giving the pressure (equation of state) and the entropy as function of ρ and T in order to be able to evolve in time the MHD system. This system of equations is similar (put aside the magnetic field) to that solved by the general circulation models of dynamical meteorology models. One way to get solar weather like forecasting models, is to simplify compressible MHD equation to deduce the equivalent of weather's primitive equation. Today the models that are closer to GCM models are models solving the MHD equation in full geometry, such as ASH (Clune et al. 1999; Miesch et al. 2000; Brun et al. 2004). Since these models can compute global properties such as the solar differential rotation or its meridional circulation, they are very useful to pin down the physical processes at the origin of the observed global (weather-like) solar phenomena and to eventually put constraints on them.

We wish here to discuss some recent insights that such 3-D MHD simulations of the solar internal dynamic has lead to. The ASH code has been used to compute several purely hydrodynamical or MHD models of the solar convection zone in which we have varied some of the control parameters (such as the Rayleigh, Reynolds and Prandtl numbers) or boundary conditions (BCs) (see Brun & Toomre 2002; Brun et al. 2004; Miesch, Brun & Toomre 2006; Browning et al. 2006 for further details). Let's consider a typical numerical model of the solar convection computed with the ASH code. This model is a simplified description of the solar convection zone: solar values are taken for the heat flux, rotation rate, mass and radius, and a perfect gas is assumed. The computational domain extends from 0.72 to 0.97 $R_\odot$ (or $L = 1.72 \times 10^{10}$ cm), thereby concentrating on the bulk of the unstable zone and here not dealing with neither penetration into the radiative interior nor convective motions in the near surface shear layer. In Miesch et al. 2006, we have computed several models that possess an imposed entropy latitudinal variation at the bottom of the CZ to mimic the thermal influence of the tachocline on the mean global flows (see also Rempel 2005). When comparing the differential rotation profile of the best solar like case, namely AB3, to a case where such thermal forcing was not imposed, namely case AB of Brun & Toomre (2002), we can deduce the ex-

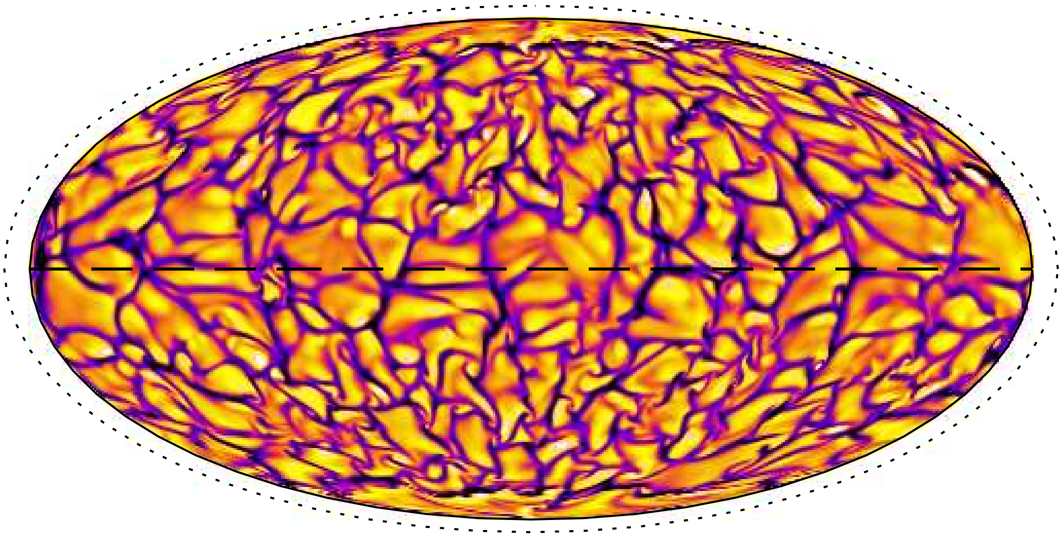

Fig. 4 Mollweide projection of the radial component of velocity near the surface of a simulation of turbulent convection. Typical flow speed is 100 m/s. In dark tone we represent downflows. Note in this highly resolved global simulation (equivalent horizontal resolution of 3000 km), how the fine structures are embedded in large scale (giant cell) convective patterns.

tra entropy contribution needed from the tachocline to make the differential rotation profile more conical.

In Fig. 3 we represent the azimuthal average of the angular velocities of case AB and AB3 as well as radial cuts of the entropy profiles at three different depths. We clearly see that the angular velocity of case AB3 is more constant along radial lines at mid latitudes than the angular velocity of case AB, thus being closer to helioseismic inversions (Thompson et al. 2003). The difference $\Delta S(\theta)$ between the pole and the equator in case AB3 will quantify the amount of entropy variation coming from the radial shear present in a tachocline assumed in exact thermal wind balance. We can thus use our 3-D models to put constraints on the properties of the solar tachocline.

More specifically, starting from the Navier-Stokes equation we can derive the thermal wind balance equation (Pedlosky 1987; Miesch et al. 2006)

$$\frac{\mathrm{d}v_\phi}{\mathrm{d}z} = \frac{g}{2\Omega_0 r c_p}\frac{\partial S}{\partial \theta}\ , \tag{9}$$

with, $v_\phi = r\sin\theta\,\Omega$ the longitudinal velocity, and

$$\frac{\mathrm{d}}{\mathrm{d}z} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}.$$

Let us assume that in the tachocline, variations in the radial direction dominate over that of the latitudinal direction and that Ω varies on a scale much smaller than the local pressure scale height or the tachocline radius $r_{\rm tacho}$ (i.e short wavelength approximation), the thermal wind balance becomes

$$\cos\theta\frac{\partial v_\phi}{\partial r} \sim r\cos\theta\sin\theta\frac{\partial\Omega}{\partial r} = \frac{g}{2\Omega_0 r c_p}\frac{\partial S}{\partial\theta}\ . \tag{10}$$

From this equation we can deduce after some algebra and proper latitudinal integration an expression for the thickness of the tachocline for a given latitudinal variation of the entropy:

$$h = \frac{\Delta\Omega(r)\ \Omega_0\ r^2_{\rm tacho}\ c_p}{g\ \Delta S(\theta)}\ , \tag{11}$$

with $\Delta\Omega(r)$ a representative measure of the radial velocity jump in the tachocline. From Fig. 3, we see that the entropy variation from the equator to the north pole at the bottom of the domain is about 650 erg/g/K for case AB, and about 820 erg/g/K for case AB3. We can thus deduce that a latitudinal entropy difference $\Delta S \sim 820$ erg/g/K is needed in order to get a conical differential rotation profile. Assuming for the solar rotational rate $\Omega_0 = 2.6\times10^{-6}$ rad/s, and the following typical values for the other parameters in the tachocline, $r_{\rm tacho} = 0.7$ R$_\odot$, $c_p = 3.5\times10^8$ erg/g/K, $g = 5.5\times10^4$ cm/s^2, and a radial Ω jump $\Delta\Omega = 10^{-7}$ rad/s, we deduce a tachocline thickness $h \sim 0.08$ R$_\odot$. This tachocline thickness is in reasonable agreement with constraints obtained via helioseismic inversion (Corbard et al. 1999) or 1-D solar models (Brun et al. 2002), given the simplification made in deriving Eq. (11).

Another important solar property that 3-D models are self-consistently computing is the shape and amplitude of the meridional circulation down to the base of the convection zone. Currently these models predict a multi cellular flow with at least two cells in radius and in latitude (Brun & Toomre 2002). Local helioseismology techniques are able to probe accurately the meridional flow down to about 0.95 R$_\odot$ (Haber et al. 2002; Giles et al. 1997) and with less accuracy down to 0.85 R$_\odot$ (Braun & Fan 1998). They indicate that the meridional flows are mostly poleward in each hemisphere with a typical flow speed of 20 m/s and that near the surface there are no persistent counter cells. However there are some observational evidence that during maximum of activity the meridional flow could possess more cells (Haber et al. 2002). Given the potential role of meridional flows for setting the solar cycle and butterfly diagram and the consequence that this flow (by transporting magnetic flux) can have on the solar activity and indirectly on geoeffective phenomena such as ICMEs, it is important to constraint this solar physical process. Since direct observations are not possible and helioseismic inversions are presently limited to 0.9 $R_\odot$, 3-D numerical models are currently the only tool that predict the structure of the meridional flows deep inside the Sun.

Another interesting dynamical property that 3-D high resolution global models of the solar convection zone predict is the presence of giant convective cells in the Sun (Brun & Toomre 2002; Thompson et al. 2003). As we can see in Fig. 4 where we display the radial convective velocity near the surface of the numerical domain ($r \sim 0.96\ R_\odot$), turbulent convection naturally develops multi-scale structures, with intermediate to large-scale convective cells and narrow fast downflows. Since in the photosphere, granulation and supergranulation are the dominant convective patterns (Derosa & Toomre 2004), it is not easy to detect the weak signal of such deep large-scale convective cells. They are worth continuing to look for since they play a significant dynamical role in our numerical models by transporting angular momentum and heat and contributing to the maintenance of the solar differential rotation.

Finally, the introduction of magnetic field in spherical 3-D MHD simulations of the solar convection zone, such as the ones performed by Brun et al. (2004), revealed that the magnetic stresses reduce the amplitude of the differential rotation by about 30%. An active Sun should thus possess a weaker differential rotation than a quiet Sun. We have studied what would be the differential rotation contrast of the Sun during the Maunder minimum (Brun 2004). By following a similar procedure than the one discussed above, i.e comparing the surface differential rotation of a magnetized case with that of purely hydrodynamical case, we have been able to deduce the latitudinal variation imposed by the magnetic field on Ω. We have found that the equator should rotate faster and the poles slower during the quiet magnetic phase. We again see, that thanks to 3-D models, one can gain some insights on the solar internal dynamics and even make quantitative predictions.

5 Conclusion

In this paper we have discussed the current status of data assimilation and forecasting methods in meteorology and solar physics, and illustrated with few examples their advantages and disadvantages. We also showed how current solar models can be used to put constraints on the dynamical processes at the origin of the solar complex dynamics, such as its differential rotation or the thickness of its tachocline. However given the complexity of the solar dynamical phenomena, we are not yet able to compute a fully integrated model of the Sun from its inner core to its surface and extended atmosphere. It is clear that simpler versions of the full MHD equations must be developed, like the meteorologist did by developing the primitive equations for the Earth's atmosphere. These simpler time dependent 3-D models should be designed with the purpose of being run operationally, in order to start developing a solar general circulation model. Clearly, parallel approaches involving high quality and continuous observational data, modern data assimilation methods, realistic numerical models of detailed physical processes and simpler 3-D forecast models are needed if one wants to progress in understanding the Sun and being able to predict and anticipate its future behaviour. Ideally the creation of an international solar prediction and modeling center where all these complementary approaches would be jointly developed and integrated in a general solar framework would be a natural and desirable outcome.

Acknowledgements. We are thankful to A. Vincent and O. Talagrand for interesting discussions regarding data assimilation and forecasting methods, to S. Turck-Chièze for fruitful discussions on space weather and climate and to the organisers of the HELAS meeting for their invitation.

References

Amari, T., et al: 2003, ApJ 595, 1231
Archotis, V., Moreno-Insertis, F., Galsgaard, K., Hood, A.W.: 2005, ApJ 635, 1299
Beer, J., Tobias, S.M., Weiss, N.O.: 1998, SoPh 181, 237
Bélanger, E., Charbonneau, P., Vincent, A.: 2005, JRASC 99, 133
Braun, D.C., Fan, Y.: 1998, ApJ 508, L105
Browning, M., Miesch, M.S., Brun, A.S., Toomre, J.: 2006, ApJ 648, L157
Brun, A.S.: 2004, SoPh 220, 333
Brun, A.S., Toomre, J.: 2002, ApJ 570, 865
Brun, A.S., Zahn, J.-P.: 2006, A&A 457, 665
Brun, A.S., Miesch, M.S., Toomre, J.: 2004, ApJ 614, 1073
Brun, A.S., Antia, H.M., Chitre, S.M., Zahn, J.-P.: 2002, A&A 391, 725
Cameron, R., Schuessler, M.: 2007, ApJ in press (astroph/0612693)
Carcedo, L., et al.: 2003, SoPh 218, 29
Cattaneo, F., Hughes, D. W.: 2001, A&G 42, 3, 18
Charbonneau, P.: 2005, http://www.solar-living-review.org
Cline, K., Brummell, N.H., Cattaneo, F.: 2003, ApJ 588, 630
Clune, T.L., et al.: 1999, ParC 25, 361
Corbard, T., Blanc-Féraud, L., Berthomieu, G., Provost, J.: 1999, A&A 344, 696
Daley, R.: 1991, *Atmospheric Data Analysis*, Cambridge Univ. Press, Cambridge
Derosa, M. L., Toomre, J.: 2004, ApJ 616, 1242
Devore,, C., Antiochos, S.: 2000, ApJ 539, 954
Dikpati, M., Gilman, P.: 2006, ApJ 649, 498
Dikpati, M., Toma, Gilman, P.: 2006, GeoRL 33, 5, L05102
Duhau, S.: 2003, SoPh 213, 203
Eddy, J.: 1975, BAAS 7, 410
Eddy, J., Gilliland, R., Hoyt, D.: 1982, Nature 300, 689
Fan, Y., Gibson, S.: 2004, ApJ 589, L105
Fisher, G.H., et al.: 2000, SoPh 192, 119
Foukal, P., et al.: 2006, Nature 443, 161
Gibson, S.E., et al.: 1999, ApJ 520, 871
Giles, P., et al.: 1997, Nature 390, 52

Haber, D.A., et al.: 2002, ApJ 570, 855
Hathaway, D., Wilson, R.M.: 2004, SoPh 224, 5
Hathaway, D., Wilson, R.M., Reichmann, E.: 1999, JGR 104, 22, 375
Inobe, S., Kusano, K.: 2006, ApJ 645, 742
Jouve, L., Brun, A.S.: 2006, proceedings of french astrophysics week (SF2A), held in Paris, June 2006
Kalnay, E.: 2003, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge Univ. Press, Cambridge
Leamon, R.J., Canfield, R.C., Blehm, Z., Pevtsov, A.A.: 2003, ApJ 596, L255
Lions, J.L.: 1971, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin
Lorenc, A.C.: 1981, MWRv 109, 701
Magara, T., Longcope, D.W.: 2003, ApJ 586, 630
Miesch, M.S., Brun, A.S., Toomre, J.: 2006, ApJ 641, 618
Miesch, M.S., et al.: 2000, ApJ 532, 593
Mikic, Z., et al.: 1999, PhPl 6, 2217
Mikic, Z., et al.: 2006, AAS SPD meeting 37, 14.05
Ossendrijver, M.: 2003, A&ARv 11, 287
Pap, J.M., Fox, P.: 2004, *Solar Variability and its Effects on Climate*, Geophysical Monograph 141, AGU press, Washington
Pedlosky, J.: 1987, *Geophysical Fluid Dynamics*, Springer-Verlag, New York
Pevtsov, A.A.: 2002, SoPh 207, 111
Pevtsov, A.A., Canfield, R.C.: 2001, JGR 106, 25191
Rempel, M.: 2005, ApJ 622, 1320
Regnier, S., Amari, T.: 2003, A&A 425, 345
Rust, D., Labonte, B.J.: 2005, ApJ 622, L69
Schrijver, C., et al.: 2006, SoPh 235, 111
Spiegel, E.A., Zahn, J.-P.: 1992, A&A 265, 106
Stein, R.F., Nordlund, Å.: 1998, ApJ 499, 914
Stix, M.: 2002, *The Sun: an Introduction*, 2nd ed, Springer-Verlag, Berlin
Svalgaard, L., Cliver, E.W., Kamide, Y.: 2005, GRL 32, L01104
Tobias, S., Hughes, D., Weiss, N.: 2006, Nature 442, 26
Talagrand, O.: 1997, J. Met. Soc. japan 75, 191
Talagrand, O.: 2003, in: R. Swinbank et al. (eds.), *Data Assimilation for the Earth System*, p. 37
Turck-Chièze et al.: 2005, in: F. Favata, A. Gimenez (eds.), *Trends in Space Science and Cosmic Vision 2020*, 39thESLAB Symp., ESA-SP 588, p. 193
Vögler, A., et al.: 2005, A&A 429, 335

