



The solar cycle and solar dynamo models: past accomplishments, present status and a strategy for the 21st century

Mausumi Dikpati

High Altitude Observatory, NCAR, 3280 Center Green Dr., Boulder, CO 80301, USA

We describe the primary observational features of solar cycles, as seen in the photosphere, and review progress made over the past sixty years to simulate and predict these features using magneto-hydrodynamic dynamo models. The focus is on the so-called Babcock-Leighton flux-transport (BLFT) dynamo models, calibrated for the Sun, which so far have been the most successful in simulation, and the only ones tested for prediction. The proposed 21st century strategy for progress emphasizes the need (a) to use modern data assimilation techniques, so successful for Earth's atmosphere simulation and prediction, to exploit all available solar observations, and (b) to generalize BLFT dynamo models to 3D to simulate and predict longitude-dependent cycle features. The 3D models must include (a) global HD and MHD instabilities in the solar tachocline, which probably create spatial patterns and time dependence that is reflected in surface observations, such as active longitudes, and (b) processes that capture the statistics and effects of emerging active regions that are tilted with respect to latitude circles, in order to accurately represent the surface source of poloidal fields, whose transport to the poles is responsible for polar field reversals. © Anita Publications. All rights reserved.

Keywords: Sun; Solar cycle; Solar dynamo

1 Introduction

1.1 Overall context

There is a general consensus that the solar cycle is produced by a magnetohydrodynamic dynamo operating in the solar convection zone and the solar tachocline just below. General dynamo theory began in the 1930's, motivated mainly by the existence of the Earth's magnetic field, thought to be generated by a dynamo operating in the liquid interior. Dynamo models with specific application to the Sun first appeared in [1].

Solar dynamo models have advanced greatly since then, but there are still many challenges to overcome in the quest to build a dynamo model that simulates the most prominent features of a solar cycle, and allows prediction of amplitude and other properties of future cycles. This review will describe the current state of solar dynamo modeling, including various classes of models, important mathematical tools used to build them, what has been accomplished to date, what the greatest challenges to be answered to make further progress, and how to go about overcoming these challenges. It will be from a perspective of what needs to be done in the 21st century to go forward, rather than a comprehensive highly referenced review of all that has been done. There are several excellent, recent reviews, such as [2, 3] for models, and [4, 5] for solar cycle observations, which the reader is encouraged to consult. We have here relied especially on [4, 5] for visualizations of observed solar cycle properties.

Many stars have dynamo-maintained magnetic fields, some similar to the Sun, others quite different, some with magnetic cycles, some not. Solving the solar dynamo problem has obvious value for understanding

Corresponding author :

e-mail: dikpati@ucar.edu (Mausumi Dikpati)

stellar dynamos generally; conversely, successfully modeling the properties of stellar dynamos will have great value for understanding the solar cycle problem.

1.2 Philosophical approach

Successful modeling of the global fluid flow and magnetic fields in a planetary atmosphere or ocean, or the convection zone of a rotating star, has generally begun with relatively simple models that make no attempt to model the all small-scale processes of the whole system, but rather focus on those physical processes thought to be important for that system. For example, models for 'baroclinic instability' in the Earth's atmosphere began to be developed in the 1940's [6,7]. Baroclinic instability is the fundamental energy conversion process in the atmosphere responsible for driving winds that transport heat from the hot tropical latitudes to the much colder polar regions. The consequences of this energy conversion, together with other global processes and effects, are responsible for driving the global circulation of the atmosphere which organizes terrestrial weather.

Building on early theories of baroclinic instability, it took another twenty years to develop models that had enough physical processes and enough realism and spatial resolution to plausibly model the atmospheric circulation [8]. Since then, such models have become vastly more detailed and realistic, simulating and predicting ever greater detail of the global circulation. In effect, the circulation models have been increasingly well 'calibrated' to the main features of the circulation, allowing them to have value for simulation and prediction.

Dynamo models for solar cycles are developing in an analogous way, but are at a much earlier stage. Parker's model [1] contained very relevant physical processes (in particular, shearing of magnetic field lines by differential rotation and twisting and lifting of field lines by helical turbulence), but certainly was not realistic enough to simulate the primary features of a particular solar cycle. But now such models do exist, and they can and have been calibrated to major features of solar cycles. This calibration process starts by defining what a key feature of a solar cycle is, which we discuss in section 2. Section 3 describes current models and results they produce.

Despite the complexity and realism of current global models for the general circulation of the Earth's atmosphere, such models still must parameterize important physical processes that are too small in spatial scale to be calculated explicitly everywhere on the globe. As computing power increases, still smaller spatial scales are included, but the gains associated with each computing power increase are modest, because, with three space dimensions and time, even a factor of two increase in resolution requires an increase of a factor of sixteen in compute power! The situation is the same for modelling solar cycles with a dynamo model. Currently calibrated solar cycle simulation models are virtually always in two dimensions, latitude and radius, and can simulate only axisymmetric solar cycle features – which, fortunately, describe a large part of what constitutes a cycle. In these models, all non-axisymmetric processes are parameterized with axisymmetric representations. As we discuss in later sections, the most successful of these models in simulating solar cycles have been the so-called Babcock-Leighton flux-transport (BLFT) dynamos.

So-called 'full 3D' solar dynamo models exist [9-11], which do include departures from axisymmetry; these models have recently produced cyclic dynamo solutions, but they do not yet give results close enough to the observed solar cycles, so they have yet to be calibrated for the Sun. Furthermore, they too do not spatially resolve energetically important scales of motion and magnetic fields in the Sun, and so still rely heavily on averages of and parameterizations of unresolved processes. This limitation on spatial resolution will not be overcome for many generations of computing power growth. BLFT dynamos have their origin in so-called 'mean-field' electrodynamics. But despite their higher spatial resolution, including variations in longitude, the 'full 3D' dynamos are themselves 'mean field' dynamos, since they still must average

over processes occurring on smaller spatial scales, and will for the foreseeable future. Thus, in a very real sense, all solar dynamo models are mean-field models.

1.3 Strategies to face new challenges

The previous section described in broad terms where solar cycle modeling using dynamo models currently stand. Here we outline what the major challenges are to making meaningful progress toward realistic simulation of actual solar cycles.

The most obvious, and perhaps most important, challenge is to move from 2D to 3D solar cycle simulation models, in order to correctly simulate global but longitude-dependent solar cycle features such as magnetic flux emergence, active longitudes, and solar sector structure, which has so much influence in organizing and modulating the solar wind and interplanetary magnetic field that transmits the effects of solar activity to the Earth, particularly to its upper atmosphere. Since the non-axisymmetric solar cycle features to be simulated are themselves global in scale, it is not clear the 'full 3D' [10,11] approach, which would spatially resolve a whole spectrum of convection, is either warranted or practical. Instead, it will be more productive to generalize the BLFT dynamo models to include the most important global 3D effects. We will describe in later sections how this can be and is being done [12].

There is incomplete information available about key components of solar dynamo models, particularly meridional circulation. Observations of meridional circulation are good near the surface, but much more uncertain for deeper layers. Helioseismic methods are generating some possible profiles, but they show large variations for different analysis techniques and time periods. It will be necessary to develop, in parallel with the 3D dynamo model itself, better theories for meridional circulation, which take account of what we do know about flow in the convection zone, such as the differential rotation there. This knowledge can be used to constrain meridional circulation profiles that are possible.

Modern data assimilation (DA) techniques, in the early stages of being included in solar dynamo models, will provide a powerful tool for inferring the form and amplitude of meridional circulation with depth and latitude [13,14], using so-called observing system simulation experiments (OSSE's). Even more important, DA will provide the computational framework for optimizing the use of observations of solar velocities and magnetic fields to initialize and generate the best possible simulations of actual solar cycles, including global 3D effects. More advanced use of such methods has been enormously beneficial to improving simulations and predictions of global atmospheric flows. Even though it is true that the solar data is much more at and above the surface than in the solar interior, it is far more important to use the available solar data for assimilation by implementing modern DA methods for estimating spatio-temporal profiles of unknown ingredients in the region where we lack observational data. In this context we recall the statement of Kalnay – the use of the model forecast is essential in achieving four-dimensional data assimilation. The model transports information from data-rich to data-poor regions, and it provides a complete estimation of the four-dimensional state of the atmosphere. Including the data-rich Northern hemisphere in the assimilation scheme makes much bigger improvements in the forecasts in the Southern hemisphere [15].

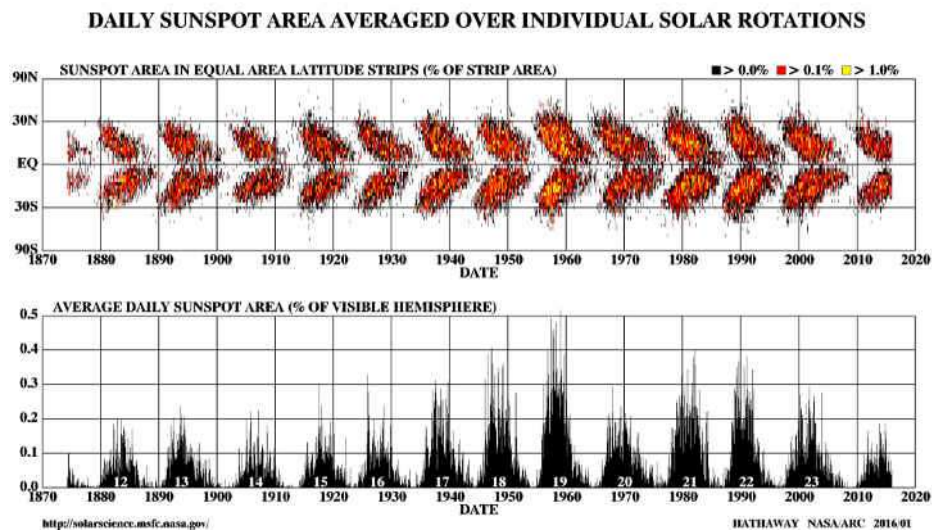
A large part of the goals for solar cycle simulation described above can be carried out within the realm of so-called 'kinematic' dynamo theory, in which the principal equation solved is the MHD induction equation. The kinematic system is certainly the place to start, because we do not need a model to simulate differential rotation, since it is nearly constant, and since simulation of 3D convection and magnetic fields adds great computational expense. In addition, this class of model has yet to be successfully calibrated to observed solar cycles.

Crucial 3D dynamical processes, such as the role of global tachocline HD and MHD instabilities, can be included as part of input of solar flow fields, without solving the full equations of motion and induction

and thermodynamic equation as a coupled system. Ultimately, it will be necessary to include nonlinear coupling, but it should be possible to do that without going to another 'full 3D' system, which would be less practical, as well as, perhaps, less able to be calibrated to solar observations.

2 Key components of a solar cycle to be modeled and predicted

The existence of a 'solar cycle' has been known since the middle of the 19th century. [Figures 1a,b](#) show the classical representations of this cycle. [Figure 1a](#) is the so-called 'butterfly diagram' which is a time latitude plot of the fractional area covered by sunspots, with North and South hemispheres averaged together. The pattern is rather regular, with each new cycle starting in the neighborhood of 30° and ending close to the equator. This is the first and probably most important cycle property that a successful solar cycle dynamo model must reproduce. We can see that the cycle period averages about 11 years, but there is a variance of more than one year, shorter and longer. As seen in [Fig 1b](#), solar cycle amplitudes as measured by sunspot area vary by a factor of 2-3 from cycle to cycle, not randomly, but with an envelope of cycle amplitudes that varies over longer time scales than a single cycle.



[Fig 1](#). Upper: Time-latitude diagram of spot area; Lower: total spot area as function of time. Credit: Hathaway/NASA/ARC

Even though cycles vary in amplitude by a factor of up to three, the statistics of the butterfly diagram are remarkably similar for different cycle strengths. This is illustrated in [Fig 2](#), which shows the centroid latitude of the butterfly diagram for small, medium and large cycles, superimposed and plotted as a function of years from cycle beginning. The separate fits for the three cycle strengths coincide within the noise limits. Thus very different amplitude cycles are largely self-similar in this property.

[Figures 1](#) and [2](#) characterize only sunspot patterns and statistics. But there is a very important relationship between sunspot fields and other surface magnetic fields, so called 'poloidal' fields, especially those near the poles. [Figure 3](#) shows this relationship. Blue and yellow shading represent opposite polarity fields. We see the low latitude butterfly diagram, but we also see rather rapid migration of poloidal fields toward the poles (the steeply slanted streaks). One sign of field predominates in this migration, opposite in North and South hemispheres, and opposite in each succeeding cycle. We can clearly see that this migration reverses the sign of the polar field near the maximum of the sunspot cycle, the polar field then holds its sign in each hemisphere until the next maximum.

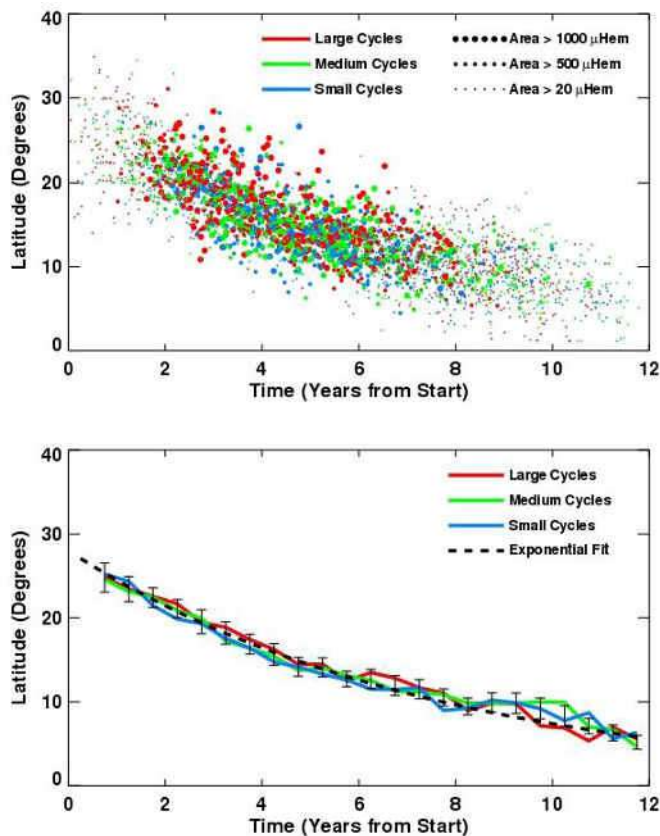


Fig 2. Drift of centroid of active regions as function of cycle phase (credit: Hathaway NASA/ARC). Note that the fits to the centroid data (lower frame) show no significant difference among small, medium and large cycles

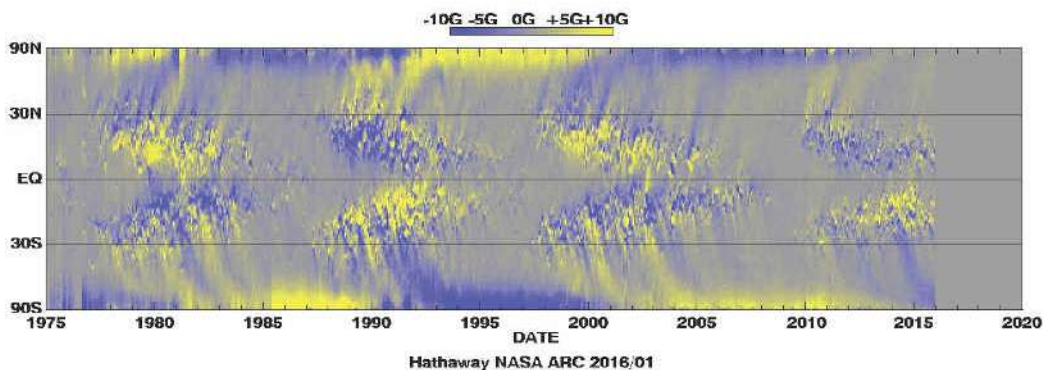


Fig 3. Magnetic butterfly diagram, constructed from longitudinally averaged radial magnetic fields obtained from Kitt Peak and SOHO measurements. Blue and yellow shading denote opposite polarities. Note the clear presence of polar field reversals and migration of flux to the poles (the nearly vertical streaks). (Credit: Hathaway NASA/ARC)

Figure 4 shows a trace of the polar field amplitude for cycles 21 to the present (see also [5]). We see that the strength of this polar field itself varies from cycle to cycle, and for the past few cycles has shown a downward trend. The transition from one polarity to the other takes place over a fairly small fraction of a cycle, followed by a longer period over which the field is often (but not always) nearly constant. There are differences in timing of polar field reversal between North and South, and occasionally there can be multiple reversals in field – but always an odd number, so that the basic pattern of field reversal is sustained. It is clear from Fig 3 that a mechanism for producing different polar field amplitudes in succeeding cycles is variations in the amount of surface magnetic flux that migrates to the poles in each cycle. If the new cycle is weaker than the previous one, then there is less new flux available to reverse the polar field, so when it does reverse the new polar field peak is likely to be weaker than the previous one (and vice versa).

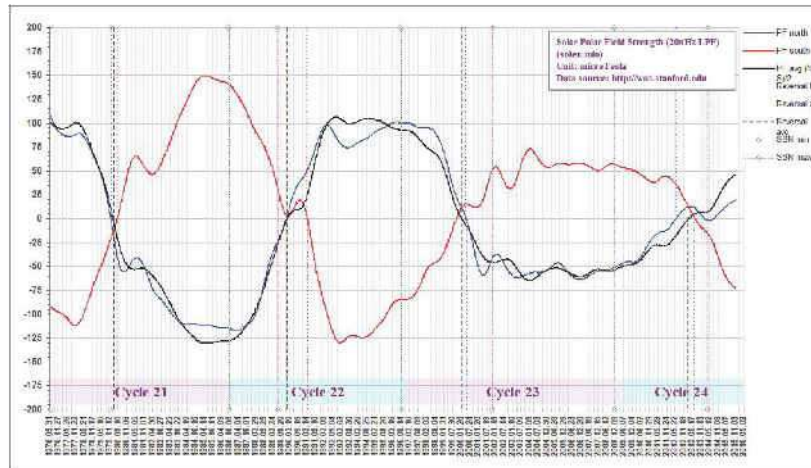


Fig 4. Polar field amplitudes for cycles 21-24 from the Wilcox Solar Observatory (WSO). Shown are North and South poles separately and together. Credit: Hoeksema and WSO website

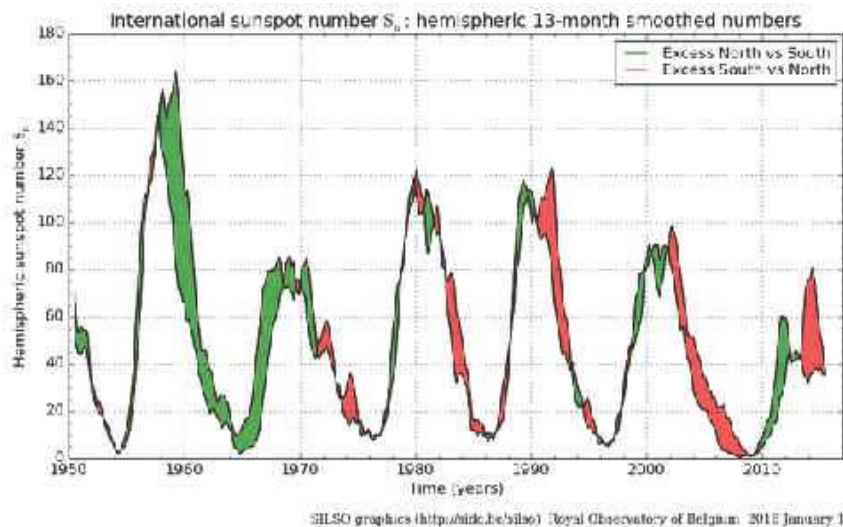


Fig 5. Excess hemispheric sunspot number, North (green) over South (red), showing that differences in phase between hemispheres can persist for multiple cycles. (Source: Royal Belgian Observatory SILSO website)

Some properties of sunspot cycles, such as amplitude and phase, can vary substantially between North and South hemispheres. If one correlates cycle peak amplitude between North and South, the correlation coefficient is about 0.5 for 3 month running averages, rising to over 0.8 for running averages exceeding 1.5 years (20 solar rotations). Thus, on short time scales the amount of activity can be very different between hemispheres. But even over a whole cycle there are substantial differences. Figure 5 shows hemispheric sunspot number, with the excess of one hemisphere over the other shaded in green or red, depending on which hemisphere has more activity. We see that for several successive years one hemisphere has consistently more flux than the other.

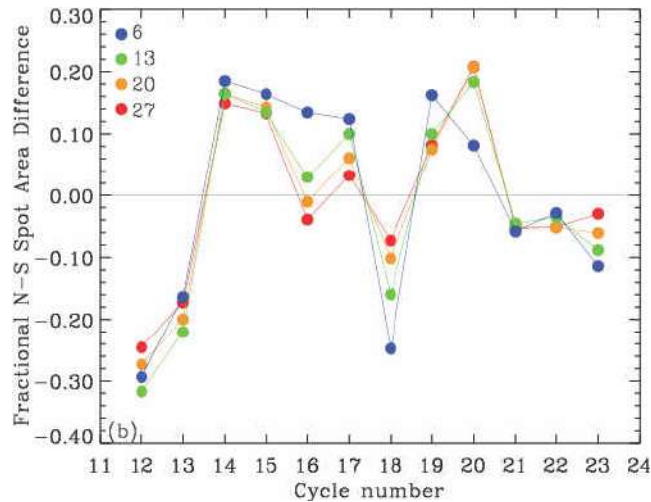


Fig 6. Fractional difference between spot area in North and South, adapted from [16] shows that the difference can be up to 30% of the average between hemispheres, and persist for multiple cycles. Color key with numbers denotes the length of the running average used, in solar rotations, showing that the basic result is evident for all averaging intervals.

The same sign of the difference often persists for more than one cycle. Figure 6 shows the fractional difference in sunspot area between hemispheres for whole sunspot cycles, starting with cycle 12. We see that one hemisphere can be consistently stronger than the other for more than one cycle, but followed by an abrupt switch to the opposite hemisphere. It appears that each hemisphere has some memory of the strength of the previous cycle in that hemisphere, but there must be enough connection between hemispheres to cause a reversal in amplitude difference.

The timing of the peaks also can differ between hemispheres, by up to more than 2 years. Left panel of Fig 7 shows this clearly. In fact, it can be argued that the North and South hemispheres avoid having synchronized activity peaks. By contrast, the timing of minimum in the two hemispheres varies by no more than a year, and usually less. So some interaction between hemispheres tends to resynchronize the cycles by the end of a cycle. This has to occur during the declining phase, because, from the right panel of Fig 7, the ascending phase is always much shorter in one hemisphere than the other, commonly by substantially more than a year.

Which hemisphere ascends faster typically persists for more than one cycle. These statistics also indicate that there is a relatively weak but not insignificant interaction between the hemispheres that keeps the timing of cycle minimum fairly well in phase, but with each hemisphere pursuing its own pace within the cycle. Thus the 'shape' of a sunspot cycle can be rather different between North and South. Correctly

simulating differences between hemispheres is a significant challenge for dynamo models; succeeding in simulating them would be a strong indication of the validity of the dynamo model used.

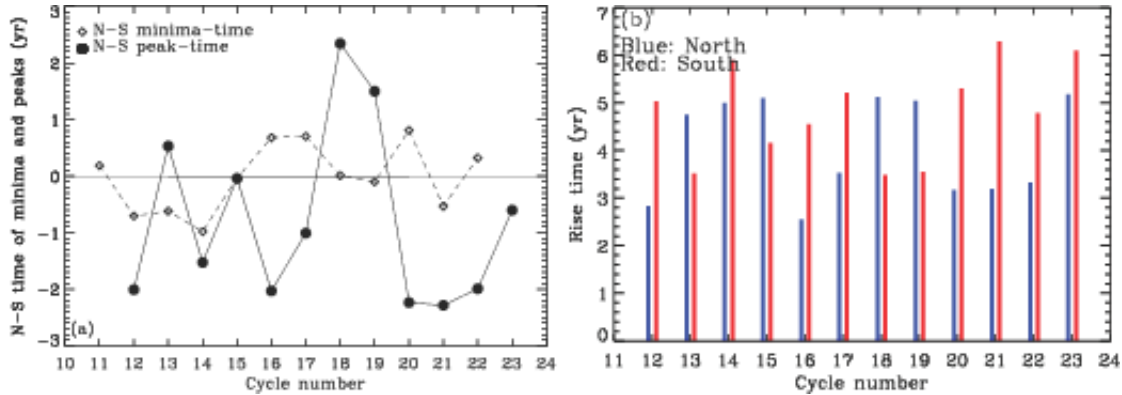


Fig 7. Difference between North and South in timing of peaks and minima, from [16]. Left panel shows that the difference in timing of cycle peaks is usually more than one year, often more than two years. By contrast the timing of minima differs by less than one year. The difference in these timing differences comes from the differences in rise time of a cycle between North and South (right panel), which is virtually always greater than one year, and sometimes more than two years. One hemisphere always jumps ahead of the other at the beginning of a new cycle.

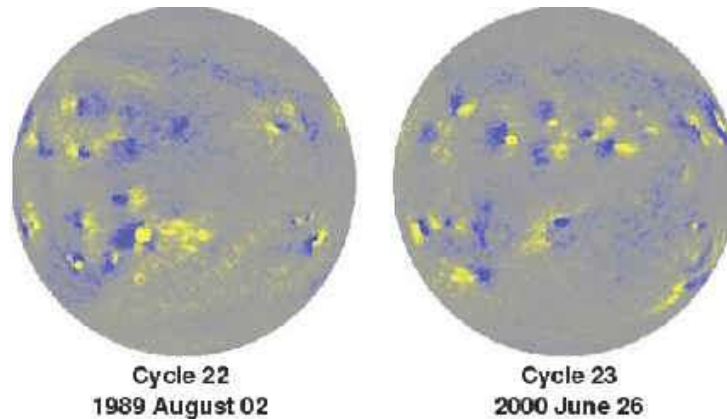


Fig 8. Typical synoptic magnetograms from cycles 22 and 23 (credit: Hathaway NASA/ARC). Hale's polarity laws and Joy's laws are very evident.

The properties of sunspot cycles described above are all axisymmetric, but we know that the emergence of solar activity in a cycle is fundamentally longitude-dependent, and this longitude dependence must be taken into account in solar dynamo models. **Figure 8** shows typical examples of solar magnetic fields on the visible disk near the maximum of cycles 22 and 23. Again, blue and yellow shading denote opposite field polarities. In **Fig 6** we see clearly Hales polarity law, which says the leading and following parts of each active region have opposite polarities in North and South hemispheres, which polarities reverse from one cycle to the next. We also see that all active regions by eye are tilted with respect to latitude circles, such that the follower polarities are closer to the poles, on average, than are the leader polarities. This tilt, called Joy's law, is crucial for determining the sign of the net magnetic flux that migrates to the

poles to reverse the polar fields, as seen in Fig 3. All dynamo models must take this tilt into account, either by including it in parameterizations of surface poloidal flux emergence (2D models) or actually calculating it, in a way that is calibrated to observed tilts.

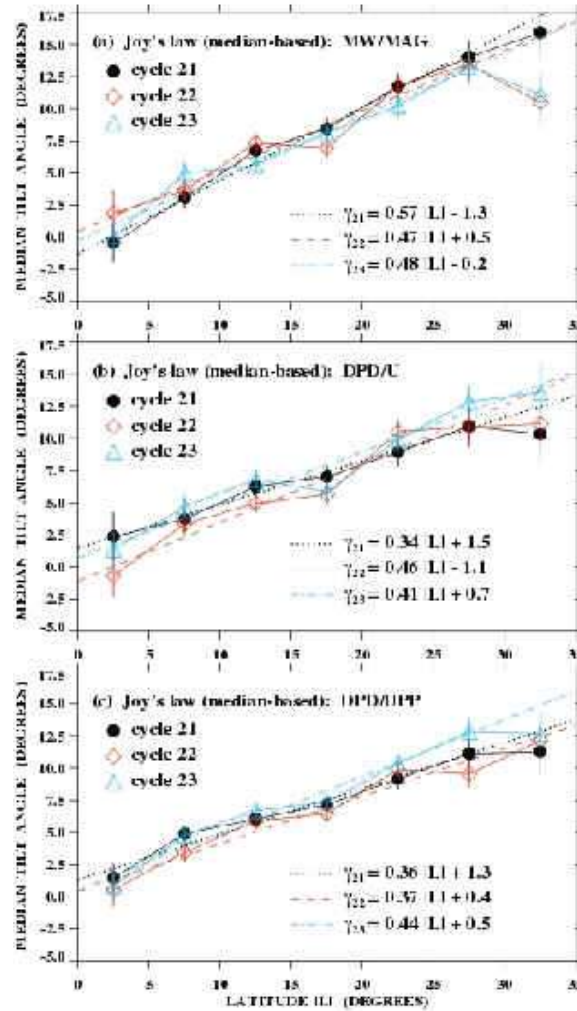


Fig 9. Active region tilts as function of cycle phase, for cycles 21,22 and 23, with linear fits (adapted from [17]).

The average amount of this tilt is now well known from observations of active regions. Figure 9 shows tilt angles as functions of latitude for cycles 21, 22 and 23. The angle is greatest at the start of the cycle, declining nearly linearly to the end. There are some differences between cycles, but in the averages these are relatively small. Different data sets give similar results. Particularly when coupled with statistics for the latitude centroid of sunspot occurrence in Fig 2, the combination of tilt angle, centroid and cycle amplitude can give us good estimates of emerged surface poloidal flux available for migration to the poles to reverse the polar fields.

There are other measures of longitude dependence in surface magnetic fields that are important for dynamo models to simulate. These involve particularly the tendency for new magnetic flux to emerge

in the same longitude band for many months, the so-called active longitudes. These persistent features are known to lead to global longitude dependent magnetic structures in the corona and interplanetary medium, the so-called magnetic sectors.

Finally, it is known that solar cycles can seem to virtually disappear for several decades, an extreme variation in the envelope of cycle amplitudes. The so-called Maunder minimum is the most recent solar manifestation of this phenomenon, for which there were very few spots seen on the Sun from about 1645-1715. There is some evidence that the cycle continued, but at a very low amplitude. Simulating these multicycle minima is a particularly difficult challenge for dynamo models, particularly from actual solar data.

3 Current models and results

3.1 Brief history of solar dynamo model developments

A brief history of developments in solar dynamo theory is diagrammed in Fig 10. The modern era of solar dynamos really began with Parker's [1] 'lifting and twisting' dynamo, which showed how a combination of helical flow and differential rotation could produce a 'dynamo wave' that propagated in latitude with time, with the right choice of helical flow and differential rotation, could roughly simulate the 'butterfly diagram' seen in Fig 2. This model was followed by more heuristic models of Babcock [18] and Leighton [19, 20] that used observed patterns and migration of fields to create magnetic cycles. The rise of 'mean-field' dynamo theory from the German school [21] gave solar dynamo theory much more rigor and connection to MHD turbulence concepts.

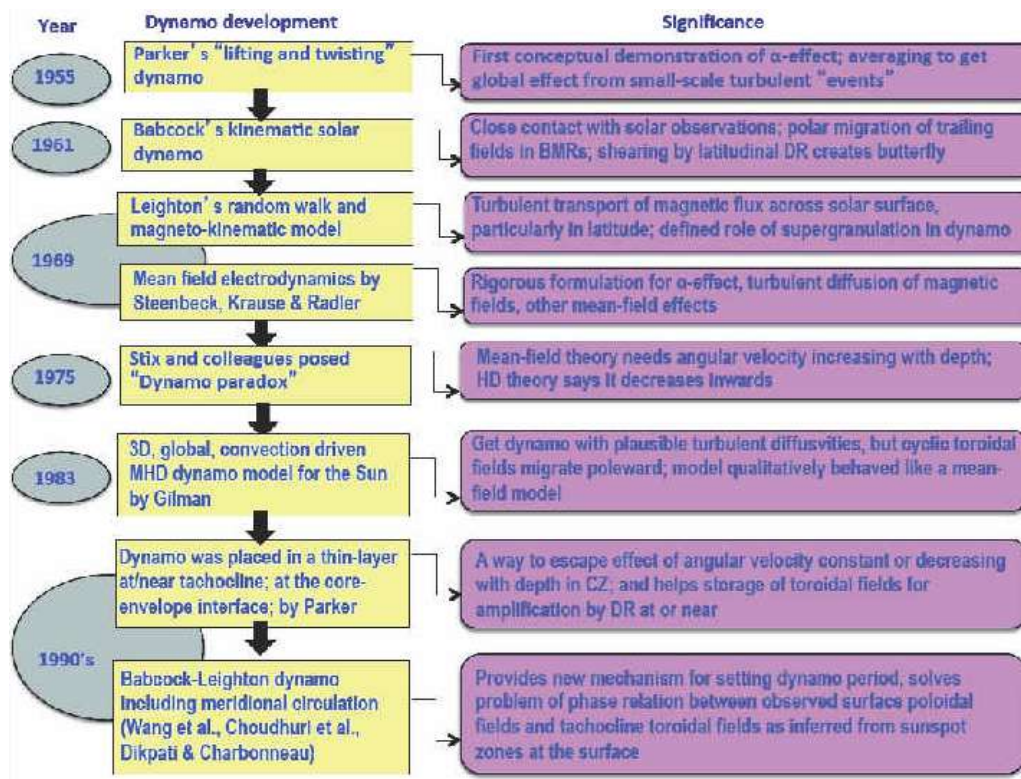


Fig 10. Flow-chart of the history of development of solar cycle dynamo models, from Parker [1] through the 1990's.

Mean-field dynamos gave plausible cycles and butterfly diagrams provided the angular velocity increased inwards in the bulk of the convection zone, but this picture was overturned by the discoveries of helioseismology. And the first full 3D convectively driven solar dynamo of Gilman [9] gave plausible differential rotation but an 'anti-solar' butterfly diagram in which toroidal fields migrated toward the poles. These developments lead to shifting the focus of the dynamo from the bulk of the convection zone to its base. The problem of reversed butterfly diagrams was overcome by the demonstration of a new mechanism for toroidal field migration, provided by meridional circulation [22, 23]. Thus was born the so-called Babcock-Leighton flux transport dynamos, which have proved so far to be the most successful dynamos for simulating solar cycles. These are 2D (latitude-radius) models, but efforts have begun to generalize them to 3D.

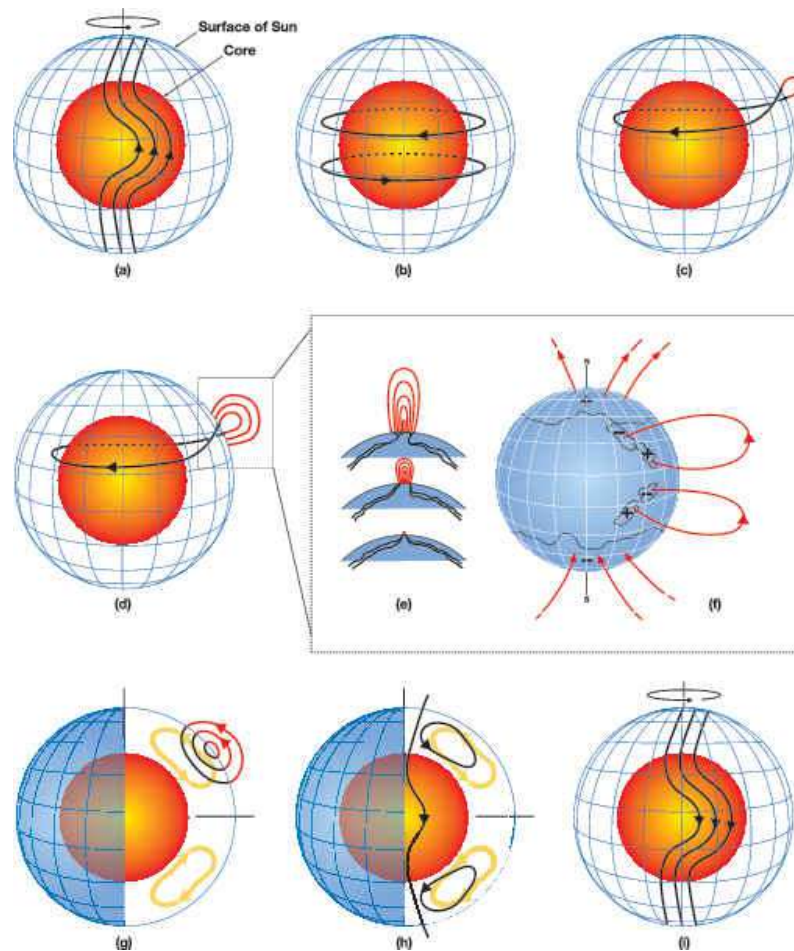


Fig 11. Schematic sequence of induction and transport processes contained in Babcock-Leighton flux-transport dynamo models, adapted from [24].

Figure 11 provides a schematic diagram that shows how this dynamo works, from [24]. It begins with the shearing of an antisymmetric poloidal field by the latitudinal differential rotation (frames a, b) to produce a toroidal field near the base of the convection zone where the turbulent magnetic diffusivity is relatively low. This is followed (frames c, d) by toroidal flux loops rising to the surface while twisting,

creating new poloidal loops there (frames e, f), which are tilted with respect to a latitude circle. These then contribute to a new axisymmetric poloidal field (frame g) that is carried toward the poles by meridional circulation and then down into the convection zone to its base (frame h), which has the opposite sign from the original poloidal field (frame a), from which a new toroidal field, also of opposite sign, is generated, leading to a repeat of the whole process, with all magnetic field signs reversed.

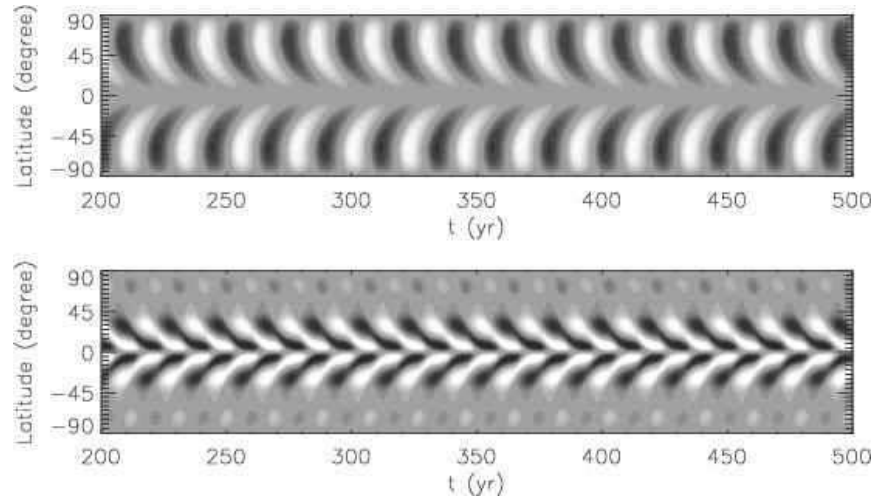


Fig 12. Upper frame: butterfly diagram for solution of Babcock-Leighton flux-transport dynamo model solution with only a surface poloidal field source, showing that in 300 years, the solution switches from antisymmetric (dipole) parity to symmetric (quadrupole) parity. Lower frame: similar simulation for which kinetic helicity or α -effect from tachocline global HD unstable modes has been added, which retains its antisymmetric (dipole) parity. Adapted from [25].

In actual dynamo simulations with the flux-transport dynamo equations, if they are done with a full spherical shell with no symmetry conditions imposed at the equator, there is no guarantee that the model will select the correct field symmetry for the Sun. In fact, various calculations showed a tendency to pick the opposite, or 'quadrupole' symmetry. Dikpati and Gilman [24] looked into this question in detail, and found that even if the simulation was started with the correct solar symmetry, if only the surface Babcock-Leighton mechanism for forming surface poloidal flux from emerging toroidal fields (Fig 11d-f) was present, within a few hundred years of simulation the symmetry switched over to quadrupole type. This is illustrated in the upper frame of Fig 12. The butterfly diagram remains plausible, but Hales polarity law is totally violated. Dikpati and Gilman [24] showed that if there was a second mechanism for generating poloidal from toroidal field, such as the lifting and twisting (α -effect) due to global hydrodynamic instability of differential rotation in the tachocline, the original dipole type symmetry was retained, as seen in the lower frame of Fig 12. Dikpati and Gilman [25] showed that the effect of the bottom α -effect preserves the original symmetry to provide a booster to the poloidal field brought down from surface polar regions, which is then swept toward the equator before it can dissipate, getting in synch with the other hemisphere to create poloidal field lines that cross the equator rather than closing there within each hemisphere. The equatorward flowing meridional circulation is a key factor, since it brings stronger poloidal field from high latitudes to merge near the equator with its opposite-hemisphere counterpart. Without the bottom α -effect, there is little poloidal field available to merge with its opposite-hemisphere counterpart through the equator to form a dipole structure, which can produce antisymmetric toroidal field after being sheared by the differential rotation.

Later calculations by others [26 - 29] have confirmed this symmetry selection process, and proposed and tested additional possible mechanisms to ensure the correct symmetry is selected.

3.2 Calibrated and benchmarked models

It makes sense to calibrate a dynamo model to solar cycle properties only if the model at least gives the correct dominant symmetry of magnetic fields about the equator. Dikpati *et al* [30] were the first to truly calibrate a solar dynamo model to solar observations, in particular the magnetic butterfly diagram shown in Fig 3. Figure 13 shows the result. The tachocline toroidal field is shown as contours, the surface poloidal field as light and dark shading. We see that there is a strong resemblance between Fig 13 and Fig 3. The shape of both field patterns is about right, the polar fields reverse near the maximum in toroidal fields, and reach maximum in between sunspot cycles. Eight versions of this class of model, using different numerical algorithms, have now been benchmarked against each other for certain parameter choices in common, and have been found to agree very closely [31].

3.3 Solar cycle simulation and prediction

The quality of the calibration suggested that the flux-transport dynamo model was ready to use to attempt the dynamo-based simulations of particular solar cycles, and even the first dynamo-based prediction of a solar cycle. Dikpati and colleagues [24, 32] used the same calibrated flux-transport dynamo model that generated Fig 13 to simulate the peaks of cycles 12-23, and predict the peak amplitude of the then future cycle, cycle 24. For cycles 12-23, they achieved a correlation between the observed and simulated peaks in excess of 0.94 for all turbulent magnetic diffusivities chosen, up to $3 \times 10^{11} \text{ cm}^2 \text{ s}^{-1}$. To do these calculations, data 'nudging' was used to drive the model with the observed time history of surface poloidal flux. Thus, magnetic surface magnetic fields from previous cycles led to the amplitude of the next cycle. When North and South hemispheres were simulated separately [16], the model produced the larger observed differences between hemispheres shown in Fig 6. The correlation between simulated and observed peaks was somewhat lower, but still in excess of 0.8. Simulations of this length were possible because for the turbulent magnetic diffusivities chosen, the 'memory' of the model was about two sunspot cycles. In the diffusivity range chosen, the dynamo was definitely in the advection-dominated regime. However, the duration of the model's memory may not be settled yet, as Yeates *et al* [33] produced only one sunspot cycle.

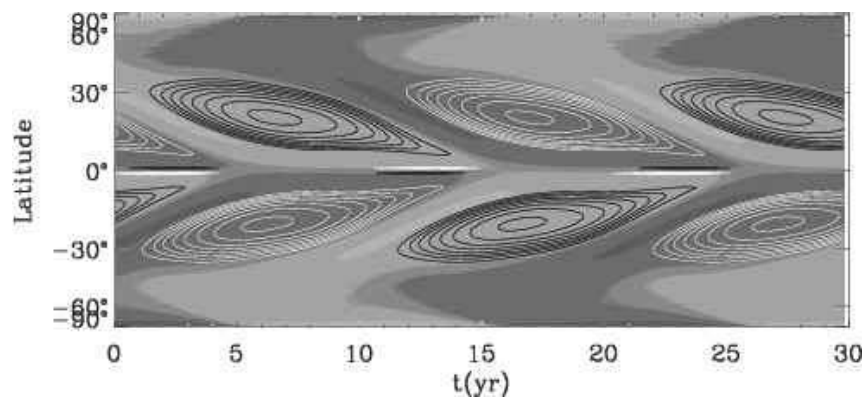


Fig 13. Calibrated butterfly diagram from the Babcock-Leighton flux-transport dynamo shown in [30]. Shading is surface poloidal field (light and dark shading opposite polarities); contours are for toroidal field amplitude at the bottom of the dynamo domain. This diagram closely resembles the observed diagram shown in Fig 3. The narrow black/white streaks at the equator near the end of each cycle are the weak remains of the toroidal field from the previous cycle, which are slow to cancel across the equator because of the low turbulent magnetic diffusivity at the bottom.

Unfortunately, the observed peak in cycle 24 has turned out to be substantially lower than predicted by Dikpati and colleagues. Also using a flux-transport dynamo model, but higher turbulent magnetic diffusivities and input of surface poloidal fields only near the poles and only near solar minimum, Choudhuri *et al* [34] predicted a much lower peak for cycle 24, actually lower than observed. Their model is in the diffusion-dominated regime, with a much shorter memory.

There is currently no consensus as to what the turbulent magnetic diffusivity is in the convection zone, so it is hard to choose among models that use different values. But there is a problem with high diffusivity models as true self-contained dynamo models, in that it can be shown by scale analysis that for diffusivities in excess of $10^{12} \text{ cm}^2 \text{ s}^{-1}$, such as suggested by mixing length applied to the Sun, and as used by Choudhuri *et al* [34], it would take a very unrealistically large surface poloidal source or α -effect to sustain a dynamo at all.

One way to see this is by calculating the critical dynamo number P for the model, which must be exceeded for a dynamo to be sustained, following, for example, [35]. The original argument is for α - ω dynamo models, but it holds equally as well for flux-transport dynamos. $P = \alpha_0 \Delta\omega \lambda^3 r_{\text{SUN}}^3 / \eta_T^2$, in which α_0 is the α -effect, $\Delta\omega$ a measure of the difference in rotation across the shell, λ the shell thickness in fractions of its radius, r_{SUN} the solar radius, and η_T the turbulent magnetic diffusivity. It is well established that P should exceed 3×10^3 for sustained dynamo action. For a mixing-length amplitude diffusivity $\sim 3 \times 10^{12} \text{ cm}^2 \text{ s}^{-1}$, this requires $\alpha_0 \sim 1.5 \times 10^3 \text{ ms}^{-1}$, which is completely unrealistic for any scale of solar convection anywhere in the convection zone, or for any surface poloidal source. By contrast, $\eta_T \sim 10^{11} \text{ cm}^2 \text{ s}^{-1}$ would require $\alpha_0 \sim 1.5 \text{ ms}^{-1}$, a much more plausible value. In fact, this is what was argued by Choudhuri *et al* [36] in their very first Babcock-Leighton flux-transport dynamo model, and therefore a value of 3 ms^{-1} was carefully considered. However, by Choudhuri and colleagues, the so-called high-diffusivity dynamo models were developed by implementing a high diffusivity to poloidal fields, but a low diffusivity to toroidal field, along with the use of an unrealistically high α -effect of about 20 ms^{-1} , to obtain a sustained dynamo solution. It remains a confusion whether a dynamo model with high and low diffusivities applied to respectively the poloidal and toroidal fields can be called a diffusion-dominated dynamo model or a dynamo model with diffusion-dominated poloidal-field and advection-dominated toroidal-field.

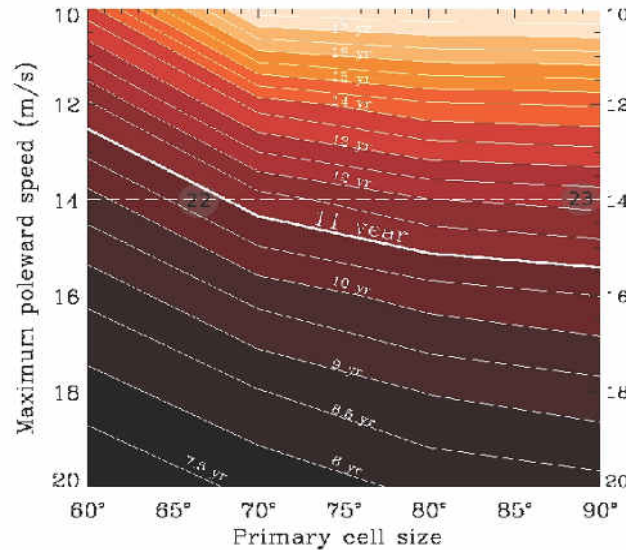


Fig 14. Relationship between latitude of poleward boundary of primary meridional circulation and period of magnetic cycles produced by Babcock-Leighton flux-transport dynamo. Adapted from [37]

If one uses the unrealistic high surface poloidal source, with high diffusivity, the calibration is negated because the resulting dynamo period is much too short. The Choudhuri *et al* [34] prediction model gets around these problems because, it does not operate as a self-contained dynamo, instead it is being forced at the top using magnetic data from previous cycles. The Sun does not have that freedom; it must be a self-contained calibrated dynamo to begin with.

In any case, it is more important to analyze reasons why the advection-dominated flux-transport dynamo-based predictions of cycle 24 amplitude have not succeeded. These have to do with the simplifying assumptions that went into the early models. In particular, the early prediction models used a steady single celled meridional circulation, but we now know that the circulation amplitude and profile both vary with time within a solar cycle. Even allowing just for a change in the light latitude boundary of the primary meridional circulation flow toward the poles is enough to explain the 2-year longer duration for cycle 23 compared to cycles 21 and 22 [37]. In the earlier cycles this cell reached to about 65° latitude, while in cycle 23 most of the time it reached all the way to the poles. The longer conveyor belt in cycle 23 led to a longer cycle, because it took longer for the surface poloidal fields in active latitudes to be transported to the poles and down to the bottom of the convection zone. Figure 14 shows a plot of the relationship between cycle period, peak circulation speed, and the latitude of the poleward boundary of the circulation.

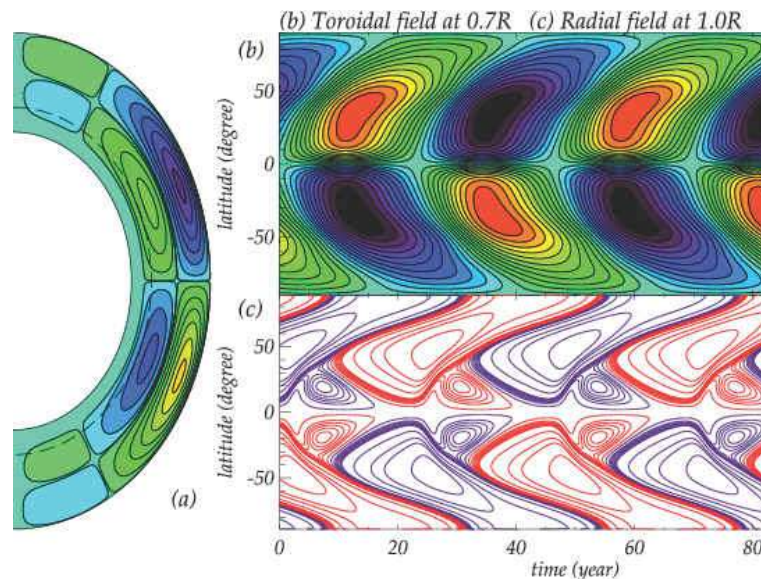


Fig 15. Butterfly diagrams (frames b,c) from flux-transport dynamo containing two primary meridional circulation cells stacked in depth (frame a). Adapted from [40]. It is evident that two cells in depth leads to a non-solar (reversed) butterfly diagram, particularly for the toroidal (sunspot) fields. This is because the toroidal field at the bottom where it is strongest is being carried toward the poles rather than the equator.

Recently there has developed evidence of a possible second, reversed meridional circulation cell below the primary cell that flows poleward, carrying the surface poloidal flux to the poles to reverse the polar field [38]. Although there exists some debates about the existence of the weak reverse cell [39], if such a two-celled pattern persists in the Sun, then the calibration of flux transport dynamo models to the Sun is completely destroyed. An example is given in Fig 15, from [40]. Because in this case the bottom circulation is toward the poles rather than the equator, as it would be with a single circulation cell with

depth, the butterfly diagram constructed from the bottom toroidal fields is completely reversed from the real Sun. If this is what is happening in the Sun, then the Sun must not be a Babcock-Leighton flux transport dynamo at all, and there must be another paradigm shift in solar dynamo theory. But observations of the second cell with depth are quite uncertain, and so it is premature to conclude such a shift is needed.

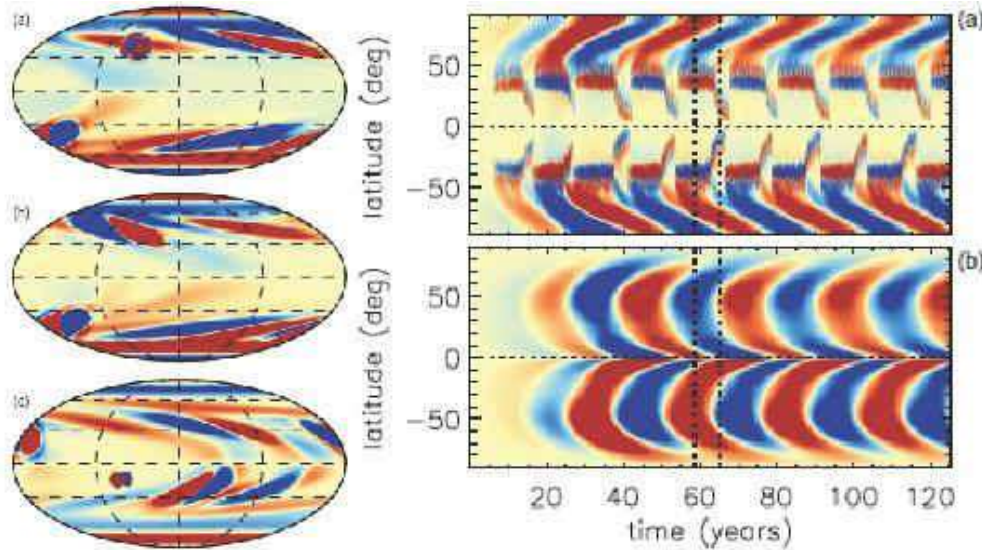


Fig 16. Early results from a 3D Babcock-Leighton flux transport dynamo model (BASH model), adapted from [12]. Frames (a)-(c) show typical evolution of the surface radial field originating from newly emerged flux. Frames (d), (e) respectively show time-latitude butterfly diagrams of the surface poloidal and bottom toroidal fields for a 120 year model run.

Beyond changes in meridional circulation, several other factors not previously accounted for in solar dynamo models may also contribute to the lack of success in solar cycle prediction models. Asymmetries between North and South hemispheres are substantial and virtually always present; the early models treated a single hemisphere. There is still uncertainty about the profile of turbulent magnetic diffusivity with depth. The data nudging scheme used in early prediction models did not make use of much of the available solar observations, which could change the predictions substantially. As discussed in section 1, modern data assimilation methods need to be incorporated into the models to use all available data. Finally, a substantial part of observed magnetic fields and patterns of flux emergence are longitude-dependent, and all prediction models so far are axisymmetric. For example, creating active region tilts is strongly longitude-dependent as well as highly variable in time. Active longitudes, while varying on longer time scales are also obviously functions of longitude.

3.4 Beginnings of application of data assimilation

Application of data assimilation methods to the solar dynamo problem is in its very early stages. As a precursor, Dikpati and Anderson [41] have found from numerical experiments that the 'response time' of a Babcock-Leighton flux-transport dynamo model to a change in meridional circulation amplitude is about six months. This is a key time scale for designing an optimum DA scheme, since the time interval for updating data needs to be long enough that the model has time to start responding to the change. In a more recent study, Dikpati, Anderson and Mitra [42] have carried out so-called Observing System Simulation Experiments (OSSE's) that show it is possible to reconstruct time changes in meridional circulation from the magnetic fields generated by the dynamo, using a relatively limited number of 'observations', synthetically

generated from the model. This kind of reconstruction will be necessary (see also Hung *et al* [14]) because there do not now exist consistent observations of meridional circulation below the solar surface, including near the bottom of the convection zone and in the tachocline, where the circulation plays a key role in moving the dynamo-generated toroidal field.

3.5 Flux emergence and active longitudes

Work has begun on incorporating longitude dependent flux emergence into flux transport dynamos. For example, using a 'spotmaker' recipe that generates statistically appropriate active region tilts, Miesch and Dikpati [12] have created a 3D Babcock-Leighton type dynamo which generates a time sequence of longitude-dependent domains of surface flux similar to those in the original Leighton model, as seen in Fig 16a. Averaging in longitude and time over this sequence creates a surface poloidal flux source for the 2D Babcock-Leighton flux-transport dynamo. Solutions from this dynamo contain plausible butterfly diagrams for surface radial field and toroidal field near the base, seen respectively in Fig 16a, b.

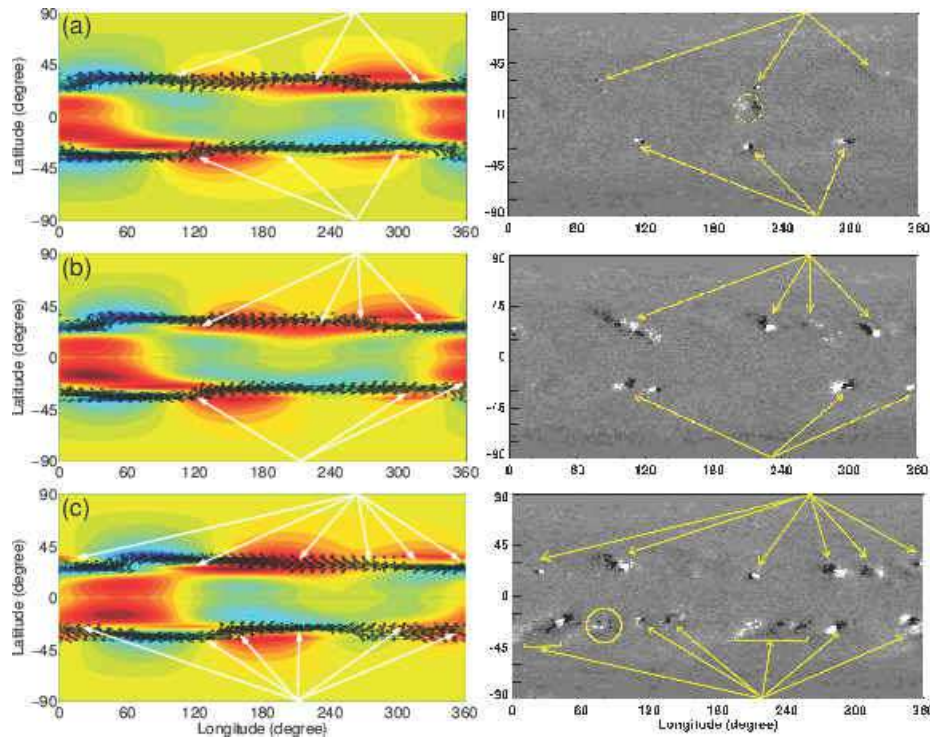


Fig 17. Theory of active longitudes and observed newly emerging magnetic flux at active longitudes, taken from [43]. Left column shows a sequence of theoretical synoptic maps of the location of tachocline upward bulges (red) from unstable tachocline MHD modes $m = 1, 2$, for a narrow band of toroidal field placed near 30° latitude. These bulges are the most likely longitude for emergence of new magnetic flux at the surface. Frame (a) is determined by adjusting the longitude phase of $m = 1, 2$ modes of equal amplitude to place the bulges of toroidal flux at similar longitudes as the observations shown in the right hand column (top frame) for Carrington rotation 1921. Then the MHD modes are allowed to propagate at their eigenvalue-determined phase speed to Carrington rotations 1927 and 1936. The predicted longitudes of the bulges are then compared to the additional emerged flux. The yellow arrows show where the bulges would fall on the observed synoptic map. We can see the correspondence remains quite good for the 15 rotations, even though new flux has emerged at different longitudes.

The longitude locations where new magnetic flux emergence appears are not to be random; certain locations are repeated sources for new flux, which can persist for many months, even years. Dikpati and Gilman [43] proposed a theory for these 'active longitudes' in terms of global HD and MHD instabilities occurring in the solar tachocline. Because the upper boundary of the tachocline can deform, instability of the combined latitudinal differential rotation and toroidal field there creates 'bulges' of the tachocline into the convection zone above that could be favored locations from which toroidal flux can rise through the convection zone to the photosphere. They showed that using just two unstable modes, of longitude wave number $m = 1, 2$, it is possible to match quite well the location of observed active regions with the bulges in the tachocline over a time of up to 15 solar rotations, as seen in Fig 17a, b. It is possible to incorporate this kind of effect into a 3D global flux-transport dynamo model by including unstable global tachocline modes into the dynamo, as inputs to the velocity field in the dynamo induction equation.

Recently McIntosh *et al* [44] have shown that many manifestations of solar activity show global periodicities with periods in the range 6-18 months. This range of periods may also be due to global dynamics and MHD of the solar tachocline. Dikpati [45] showed, using a nonlinear HD shallow water model of unstable latitudinal differential rotation in the tachocline, that there is a well defined nonlinear oscillation between energy in the differential rotation and in the low longitudinal wave number 'Rossby waves' that gain energy from it. When these waves extract enough energy from the differential rotation, it ceases to be unstable, and subsequently extracts energy from the disturbances, growing until it becomes unstable again. In this way, the differential rotation and perturbation energies oscillate out of phase with each other, with the total energy of the system remaining virtually constant. Similar oscillations should still occur when toroidal fields are added; we would expect that the maximum amount of new flux would emerge at the surface when the perturbations and therefore the bulges have maximum amplitude.

4 What should be done next?

The overriding priority for the 21st century solar cycle dynamo models is to build a 3D calibrated dynamo. 3D is essential for simulating the substantial departures from axisymmetry seen in typical synoptic magnetograms produced starting in the 1970's, and now available many times per solar rotation from both SDO/HMI and from ground-based observing systems such as SOLIS and MWO. Calibration, including for longitude-dependent features, is essential for determining that the model simulates typical solar field patterns through a solar cycle that the model can be used to simulate past cycles as well as applied to forecasting future cycles.

In principle, there are at least two approaches that could be taken. One is to use available full 3D convectively driven MHD dynamo models; the other is to generalize an already calibrated 2D Babcock-Leighton flux transport model to include longitude-dependent fields and flows. The latter approach is advantageous over the former, because no current full 3D dynamo model has been successfully calibrated for the Sun, and none seems likely to be for the foreseeable future. In addition, they are extremely expensive in computing time to operate, even without the inclusion of modern data assimilation algorithms to incorporate observations of solar magnetic fields and global velocities. In our view, therefore, the only practical approach to the global 3D solar dynamo problem is to generalize a successful 2D model.

What form should this generalization take? It should be strongly influenced by solar observations. For example, the typical solar synoptic magnetogram shows photospheric magnetic fields have truly global patterns, with high amplitude in low longitudinal wave-numbers m . Only a few low m 's can capture much of the important departures from the axisymmetric or $m = 0$ part of the field. In addition, as we go above the photosphere into the corona and beyond, the lowest m 's dominate even more in the magnetic structure. Solar 'sectors' generally are well represented by just $m = 1, 2$ modes. For these reasons, the model that is built should allow for picking just the lowest wave-numbers, and allow experimentation to determine how

few are needed. Therefore, fast-Fourier transform representation of interactions between modes should be avoided in the formulation.

It is straightforward to expand the kinematic induction equation to include longitude dependence of this type. In the linear case, one simply gets separate equations for the amplitude of each longitudinal wave-number, including the original 2D $m = 0$ mode. Linear solar dynamo models of this type were first studied in the early 1970's by Stix [46], but have not been pursued that much since. In these models, the flow fields included were still axisymmetric, as was the parameterized helical turbulence included. For future models, we have strong motivation, and theoretical support, for going far beyond this limited 3D model, to include velocity fields that are themselves dependent on longitude. The theoretical support is provided by the substantial theory of global HD and MHD instabilities in the solar tachocline that has been developed over the past 20 years [47-53,45]. This theory also favors the existence of global helical velocity and magnetic field disturbances in the tachocline, for which the same low longitudinal wave-numbers $m = 1, 2$ are the most unstable, and therefore the most likely to dominate in the tachocline and layers above. The only exception to that is if the toroidal field present there is confined to a narrow ($< 10^\circ$) band of latitudes. In this case, somewhat higher m 's are also excited.

Figure 18a shows a schematic framework for such a 3D model. In concentric spherical shells, starting from the innermost, the model would include 3D helical flow from global tachocline instabilities, from which bulges into the convection zone above would be generated. In the bulk of the convection zone there would be 2D differential rotation and meridional circulation, and residual 3D tachocline instability flow and 3D global fields. Near the outer boundary would reside the 3D Babcock-Leighton surface poloidal field source, generated from a spotmaker recipe that generates newly emerged flux from active longitude locations determined by global tachocline instability. Later versions could also contain a coronal domain above, with 3D field structures determined by the dynamo below. Figure 18b shows an example of the flow instability and bulges (red) and depressions (blue) generated in the tachocline. Red also corresponds to upward radial motion, blue to downward motion. Therefore the disturbance shown has kinetic helicity analogous to that used by Dikpati and Gilman [25] to derive a symmetry selection mechanism.

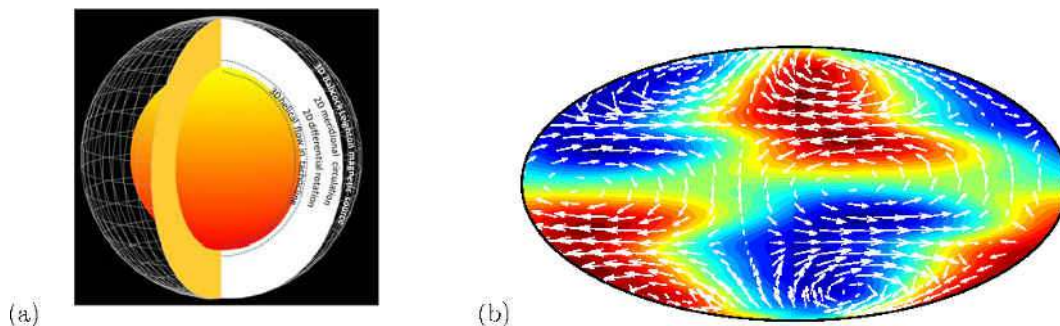


Figure 18. Frame (a): Schematic of proposed 3D flux-transport dynamo model for solar cycles, showing location in radius of various 2D and 3D processes to be included. Frame (b): An example from nonlinear simulations [45] showing a typical pattern of global HD instability flow and tachocline thickness (red:bulges; blue: depressions in the upper tachocline boundary). Red also implies upward flow, blue downward flow; arrows depict horizontal flow. The radial flows are clearly correlated with the radial component of vorticity of the horizontal flows, implying the existence of kinetic helicity in the flow. It is this kinetic helicity that was responsible for the preservation of dipole symmetry in the dynamo solutions in [25].

We see that unstable modes in the tachocline can have at least three important effects on the dynamo. First, they can 'imprint' longitude dependent patterns on the bottom of the convection zone, particularly

in magnetic field, which can be transmitted to the photosphere where they can be seen. Second, they can provide limited longitudinal bands where upward bulges in the tachocline can appear, which, when they coincide with the presence of strong toroidal fields, can provide favored sites for toroidal flux tubes that become buoyant and rise to the photosphere. Third, they also are a powerful source of kinetic helicity for the dynamo itself, supplementing the Babcock-Leighton type surface poloidal source in creating, amplifying and evolving the dynamo's poloidal field. All three of these effects could be important in the solar dynamo, so it is essential that their effects be included in the model.

It is feasible to incorporate all three of these effects into a global 3D flux-transport dynamo model. The imprinting of magnetic and velocity patterns at the bottom of the dynamo domain follows directly from including nonaxisymmetric flow and fields from solutions to the linear instability equations, with assigned amplitudes. Amplitudes for such modes can be calculated directly from nonlinear models for global tachocline instabilities. One such model already exists for HD instabilities [45], which can be generalized to include global MHD effects. This part of the dynamo model can in principle be calibrated to the Sun by comparing observations of persistence and longitudinal phase and phase speed of low longitudinal wavenumber magnetic patterns with those predicted by the linear and nonlinear tachocline instability.

In addition to the effects of global HD and MHD tachocline instabilities, at least one other well observed process that occurs on smaller longitudinal scales must be included in the model. This process is the emergence of solar active regions that are 'tilted' with respect to a latitude circle. This tilt is crucial, because it leads to preferential migration of surface poloidal flux of the polarity of follower sunspots toward the poles. Up to very recently, 2D dynamo models captured this effect in a parameterized Babcock-Leighton surface poloidal source. This approach works up to a certain point, but does not allow for fluctuations in the tilt angle with time, or from one active region to another, through a sunspot cycle. Since tilt angle, along with active region amplitude, largely determine how much net emerged flux of one sign gets to the poles, this effect can strongly influence the amplitude of polar fields with time; since in all dynamo models polar fields, which reverse near the maximum of each sunspot cycle, play an important role, capturing this emergence and tilting process in a way that calibrates well to the Sun is critical for 3D solar dynamos.

To do this requires including in the dynamo model a sequence of steps that capture (a) the longitude and latitude location of magnetic flux about to come into the bottom of the convection zone; (b) what tilt the flux starts with and how it is modified by Coriolis forces as it rises through the convection zone; (c) when and where it emerges at the photosphere; (d) what amplitude and tilt it has when it gets there; and (e) how it is evolved and transported poleward once it is on the surface. The outcome of this sequence of processes itself will need to be calibrated at least statistically against solar observations of active region location amplitude and tilt angle. In actual cycle simulations, data assimilation methods can be used to update the observed active region properties, compared to what the dynamo model itself produces, with appropriate updating guided by the difference between the observations and the model output.

Out of the sequence of processes just described will come a generalized Babcock-Leighton surface poloidal flux source that is longitude dependent. It can be Fourier analyzed to compute the $m = 0$ axisymmetric component, which is simply the traditional surface source for 2D flux-transport dynamo models, but with the addition of sources for each low nonzero wavenumber. This step allows the full effects of the flux emergence process to be represented in the same model framework as all other effects.

This 3D class of model will make much heavier use of solar observations, particularly at the surface; incorporating all this data into the model in a sensible way will require sophisticated data assimilation algorithms, most likely of the sequential type, to update the model integrations. The choice of update interval will be constrained by the time resolution of the observed data as well as the 'response' time of the dynamo model to changes in an input, such as meridional circulation. Experience so far with simpler 2D models with data assimilation point to an update interval of 15-30 days, which is consistent with the time cadence of the solar observations to be used.

The initial 3D models should be kinematic; they should solve just the induction equation. In later versions it should be possible, and desirable, to couple the induction equation to equations that calculate the global tachocline instability modes. The toroidal field produced by the dynamo as a function of time can be incorporated into linear and nonlinear models of tachocline instability to calculate how the unstable modes evolve. Then these can be included in computational updates of the induction equation, that include all changes in the dynamo inputs.

It is important to note that all of the above steps to build a global 3D solar dynamo are done without including any explicit calculations of solar convection on any spatial scale. This is done partly to see just how far one can go in producing a realistic global solar dynamo in which convection itself enters in only through estimates of turbulent magnetic diffusion, and implicitly through its maintenance of solar differential rotation. Since differential rotation is very nearly a constant throughout a solar cycle, not much can be gained by calculating it explicitly from convection; furthermore, such models do not yet reproduce the observed differential rotation that well, so using such a model can actually degrade the dynamo model's simulation of an actual solar cycle. It will be necessary to build better theoretical models of meridional circulation, probably starting from better measured quantities such as differential rotation; current global convective models do not do particularly well in simulating observed meridional circulation either.

Once a 3D global solar dynamo as described above is produced, which calibrates well to the Sun, then it will be possible to extend such a model into the corona to calculate dynamo-produced coronal structures with longitude dependence that can be compared with observed coronal structures. Then it should be possible to relate these structures to such features as coronal holes, polar structures, sectors, and high speed solar wind streams.

Finally, if the model calibrates well to the Sun, it can be tested for its ability to predict both longitude-averaged and longitude dependent features of future solar cycles, using real solar data.

Acknowledgments

SDO data courtesy of SDO (NASA) and the HMI and AIA consortium and WSO data courtesy. Many helpful comments/suggestions of an internal reviewer and an external reviewer of this paper are gratefully acknowledged. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

References

1. Parker E N, *Astrophys J*, 121(1955)293-314.
2. Charbonneau P, *Liv Rev Sol Phys*, 7(2010)3.
3. Charbonneau P, *Ann Rev Astron Astrophys*, 52(2014)251-290.
4. Hathaway D H, *Liv Rev Sol Phys*, 12(2015)4.
5. Petrie G, *Liv Rev Sol Phys*, 12(2015)5.
6. Charney J G, *Meteorology*, 4 (1947)136-162.
7. Eady E T, *Tellus*, 1(1949)33-52.
8. Trenberth K E, *Climate System Modeling*, (Cambridge Univ Press), 1992, 788pp.
9. Gilman P A, *Astrophys J Suppl*, 53(1983)243.
10. Lawson N, Strugarek A, Charbonneau P, *Astrophys J*, 813(2015)95.
11. Fan Y, Fang F, *Astrophys J*, 789(2014)35.
12. Miesch M S, Dikpati M, *Astrophys J Lett*, 785(2014) L8.
13. Dikpati M, Anderson J L, Mitra, D, *Geophys Res Lett*, 41(2014)L5361-L5369.
14. Hung C P, Jouve L, Brun A L, Fournier A, Talagrand O, *Astrophys J*, 814(2015)151.
15. Kalnay E, *Atmospheric modeling, Data assimilation and Predictability*, (Cambridge Univ Press), 2003, p 328.

16. Dikpati M, Gilman P A, de Toma G, Ghosh S S, *Solar Phys*, 245(2007)1-17.
17. Wang Y -M, Colaninno R C, Baranyi T, Li J, *Astrophys J*, 798(2015)50.
18. Babcock H W, *Astrophys J*, 133(1961)572-587.
19. Leighton R, *Astrophys J*, 140(1964)1547-1562.
20. Leighton R, *Astrophys J*, 156(1969)1-26.
21. Radler K -H, *Fluid Mechanics and its Applications*, 80(2007)55.
22. Dikpati M, Charbonneau P, *Astrophys J*, 518(1999)508-520.
23. Wang Y -M, Sheeley N R (Jr), Nash A G, *Astrophys J*, 383(1991)431-442.
24. Dikpati M, Gilman P A, *Astrophys J*, 649(2006)498-514.
25. Dikpati M, Gilman P A, *Astrophys J*, 559(2001)428-442.
26. Bonanno A, Elstner D, Rüdiger G, Belvedere G, *Astron Astrophys*, 390(2002)673-680.
27. Hotta H, Yokoyama T, *Astrophys J*, 709(2010)1009-1017.
28. Chatterjee P, Nandy D, Choudhuri A R, *Astron Astrophys*, 427(2004)1019-1030.
29. Dikpati M, Rempel M, Gilman P A, MacGregor K P, *Astron Astrophys*, 437(2005)699-702.
30. Dikpati M, de Toma G, Gilman P A, Arge C N, White O R, *Astrophys J*, 601(2004)1136-1151.
31. Jouve L, Brun A S, Arlt R, Arlt, Brandenburg A, Dikpati M, Bonanno A, Käpylä P J, Moss D, Rempel M, Gilman P, Korpi M J, Kosovichev A G, *Astron Astrophys*, 483(2008)949-960.
32. Dikpati M, de Toma G, Gilman P A, *Geophys Res Lett*, 33(2006)L05102.
33. Yeates A, Nandy D, and Mackay D, *Astrophys J*, 673(2008)544-556.
34. Choudhuri A R, Chatterjee P, Jiang J, *Phys Rev Lett*, 98(2007)131103.
35. Stix M, *Astron Astrophys*, 47(1976)243-254.
36. Choudhuri A R, Schüssler M, Dikpati M, *Astron Astrophys*, 303(1995)L29-L32.
37. Dikpati M, Gilman P A, de Toma G, Ulrich R K, *Geophys Res Lett*, 37(2010)L14107.
38. Zhao J, Bogart R S, Kosovichev A G, Duvall T L (Jr), Hartlep T, *Astrophys J Lett*, 774(2013)L29.
39. Kholikov S, Serebryansky A, Jackiewicz J, *Astrophys J*, 784(2014)145.
40. Belucz B, Dikpati M, Forgacs-Dajka E, *Astrophys J*, 806(2015)169.
41. Dikpati M, Anderson J L, *Astrophys J*, 756(2012)20.
42. Dikpati M, Anderson J L, Mitra D, *Astrophys J* (2016) submitted.
43. Dikpati M, Gilman P A, *Astrophys J Lett*, 635(2005)L193-L196.
44. McIntosh S, Leamon R J, Krista L D, Tittle A M, Hudson H S, Riley P, Harder J W, Kopp G, Snow M, Woods T N, Kasper J C, Stevens M L, Ulrich R K, *Nature Comm*, 6(2015)6491.
45. Dikpati M, *Astrophys J*, 745(2012)128.
46. Stix M, *Astron Astrophys*, 13(1971)203-208.
47. Gilman P A, Fox P A, *Astrophys J*, 484(1997)439-454.
48. Dikpati M, Gilman P A, *Astrophys J*, 512(1999)417-441.
49. Dikpati M, Gilman P A, *Astrophys J*, 551(2001)536-564.
50. Garaud P, *Mon Not Royal Astron Soc*, 324(2001)68-76.
51. Cally P S, Dikpati M, Gilman P A, *Astrophys J*, 582(2003)1190-1205.
52. Cally P S, *Mon Not Royal Astron Soc*, 339(2003)957-972.
53. Dikpati M, Gilman P A, Cally P S, Miesch M S, *Astrophys J*, 692(2009)1421-1431.

[Received: 19.2.2016; accepted: 20.3.2016]