

Reconstructing solar magnetic fields from historical observations

II. Testing the surface flux transport model

I. O. I. Virtanen¹, I. I. Virtanen¹, A. A. Pevtsov^{2, 1}, A. Yeates³, and K. Mursula¹

¹ ReSolVE Centre of Excellence, Space Climate research unit, University of Oulu, PO Box 3000, 90014 Oulu, Finland
e-mail: iiro.virtanen@oulu.fi

² National Solar Observatory, Sunspot, NM 88349, USA

³ Department of Mathematical Sciences, Durham University, Durham, DH1 3LE, UK

Received 10 January 2017 / Accepted 13 April 2017

ABSTRACT

Aims. We aim to use the surface flux transport model to simulate the long-term evolution of the photospheric magnetic field from historical observations. In this work we study the accuracy of the model and its sensitivity to uncertainties in its main parameters and the input data.

Methods. We tested the model by running simulations with different values of meridional circulation and supergranular diffusion parameters, and studied how the flux distribution inside active regions and the initial magnetic field affected the simulation. We compared the results to assess how sensitive the simulation is to uncertainties in meridional circulation speed, supergranular diffusion, and input data. We also compared the simulated magnetic field with observations.

Results. We find that there is generally good agreement between simulations and observations. Although the model is not capable of replicating fine details of the magnetic field, the long-term evolution of the polar field is very similar in simulations and observations. Simulations typically yield a smoother evolution of polar fields than observations, which often include artificial variations due to observational limitations. We also find that the simulated field is fairly insensitive to uncertainties in model parameters or the input data. Due to the decay term included in the model the effects of the uncertainties are somewhat minor or temporary, lasting typically one solar cycle.

Key words. Sun: magnetic fields – Sun: photosphere – Sun: activity

1. Introduction

Surface flux transport (SFT) models have been studied and developed for several decades, starting from Leighton (1964), and continuing to modern models capable of replicating fine details of magnetograms (Jiang et al. 2014; Yeates et al. 2015; Lemerle et al. 2015). They have been used for various purposes including, for example, studies of long-term solar activity over hundreds of years (Jiang et al. 2011), simulations of magnetic field between observations (Worden & Harvey 2000), and studies on how individual active regions affect poleward flux surges (Yeates et al. 2015). The models usually include meridional circulation, differential rotation and diffusion, and a term describing the emerging flux. Sometimes additional terms are included to describe the decay of the magnetic field. The main problem in SFT modeling is the parameterization of these processes. Meridional circulation and diffusion in particular are difficult to observe, and hence difficult to model accurately.

We aim to use SFT modeling to simulate the long-term evolution of the photospheric magnetic field over the past one hundred years from pseudo-magnetograms created from historical calcium II K line and sunspot observations (Pevtsov et al. 2016). In particular we are studying the polar fields, which are connected to coronal holes. Coronal holes are a source of open magnetic flux that extends into the heliosphere as the heliospheric magnetic field. They therefore play an important role in space weather in general, but are not included in the pseudo-magnetograms. The lack of information about polar areas and

the exact structure of active regions pose further challenges to SFT modeling.

Magnetic flux emergence can be included in an SFT simulation in many different ways, varying from creating artificial bipoles based on the observed sunspot number (Jiang et al. 2011) or the observed sunspot area (Jiang et al. 2010; Baumann et al. 2004) to assimilating observed active regions (Yeates et al. 2015) or full disk magnetograms (Worden & Harvey 2000) directly. The pseudo-magnetograms include active regions that can be directly assimilated into the simulation. However, the exact flux distributions inside these active regions are not very accurate, which leads to an additional source of uncertainty for pseudo-magnetograms compared to assimilating observed magnetograms.

The most widely used SFT model includes only meridional circulation, differential rotation, and supergranular diffusion (Mackay & Yeates 2012). In long simulations spanning multiple solar cycles, this model is in some cases unable to correctly reverse the polarity of the polar fields (Schrijver et al. 2002; Baumann et al. 2006; Lean et al. 2002). It has also been shown to be insufficient to simulate the decay of the axial dipole during the minimum between cycles 23 and 24 (Yeates 2014). Including an additional decay term that slowly weakens the polar fields over time can help eliminate these problems. Schrijver et al. (2002) and Yeates (2014) use simple exponential decay terms without a clear physical interpretation. Baumann et al. (2006) derive a similar decay term based on the

diffusion of the radial field in the convection layer. A significant increase in the speed of meridional circulation can also reduce the build up of flux at the poles and help weaken the unrealistically strong polar fields (Wang et al. 2002b; Yeates 2014), but the circulation speed required to do this is much higher than the observed circulation speed (Hathaway & Rightmire 2010).

The speed of meridional circulation can be measured, for example, by tracking magnetic features temporally in magnetograms or by determining it from helioseismic observations. Although many SFT models assume a constant meridional circulation profile, varying speed profiles have also been used and shown to improve the stability of the polar fields (Wang et al. 2002a). Upton & Hathaway (2014) use feature tracking to find the correct meridional flow and differential rotation speeds for each rotation. This approach leads to more accurate simulations, but cannot be used with historical observations, which do not contain information about the speed of surface flows.

Supergranular diffusion is a challenging process to parametrize, because it cannot be measured directly. Early studies suggested diffusivities as large as $770\text{--}1540\text{ km}^2\text{ s}^{-1}$ (Leighton 1964), but later estimates of the optimal value of supergranular diffusivity have been significantly lower. For example, values of $600\text{ km}^2\text{ s}^{-1}$ (Wang et al. 1989), $500\text{ km}^2\text{ s}^{-1}$ (Wang et al. 2002b), and $350\text{ km}^2\text{ s}^{-1}$ (Lemerle et al. 2015) have been suggested. The large uncertainties in supergranular diffusivity imply that the sensitivity of the model to this parameter should be investigated.

In this paper we use the SFT model by Yeates et al. (2015) extended with the decay term by Baumann et al. (2006) to simulate the radial photospheric magnetic field of the past three solar cycles using synoptic maps from the National Solar Observatory (NSO). The data assimilation technique of the model can be easily adapted to be compatible with pseudo-magnetograms, and the model uses a meridional circulation profile that is constant in time. These properties make it particularly suitable for use with pseudo-magnetograms. In this paper we show that even if there are significant uncertainties related to meridional circulation, supergranular diffusion, and the structure of active regions, the model is fairly robust, and can reliably simulate the large scale evolution of polar fields. We also demonstrate the necessity of the additional decay term. The paper is organized as follows. In Sect. 2 we describe the data used in this paper and the instrumentation used to obtain the data. In Sect. 3 we discuss the SFT model and the additional decay term in detail. In Sect. 4 we run various tests with the model to assess how sensitive it is to uncertainties in meridional circulation, supergranular diffusion, the exact structure of active regions, and the initial polar field. In Sect. 5 the simulated magnetic field is compared to observations, and in Sect. 6 we discuss the results in light of the intended use of the model with pseudo-data constructed from historical observations.

2. Data

The SFT model uses synoptic Carrington rotation maps of the magnetic field as input. The photospheric longitudinal magnetic field has been measured in the NSO at Kitt Peak (KP) since early 1970s. In 1970–1974 (Carrington rotations 1558–1611) observations were carried out with a 40-channel photoelectric magnetograph using the FeI 523.3 nm spectral line (Livingston & Harvey 1971). From February 1974 (CR 1625) until March 1992 (CR 1853) the instrument was a 512-channel diode array magnetograph using the Fe I 868.8 nm spectral line (Livingston et al. 1976). The CCD spectromagnetograph

(SPMG) started operating in 1992, first using the Fe I 550.7 nm spectral line from April to October 1992 (CR 1855–1862) and then the same Fe I 868.8 nm line as the previous instrumentation (Jones et al. 1992). These observations terminated in August 2003 (CR 2006), when the Synoptic Optical Long term Investigations of the Sun (SOLIS) vector spectromagnetograph (VSM) replaced the old instrumentation (Keller et al. 2003). In 2006 the VSM modulators were replaced, and in 2010 the original Rockwell cameras, which had a pixel size of $1.125''$, were replaced with Sarnoff cameras that have a smaller pixel size of $1''$ (Pietarila et al. 2013). The size of the full-disk images of the 512-channel magnetograph is 2048 by 2048 pixels and of the SPMG 1788 by 1788 pixels. The size of the SOLIS full-disk images is 2048 by 2048 pixels. The full resolution images were used to create pseudo-radial synoptic maps with a reduced resolution of one degree in longitude and $1/90$ in sine of latitude, resulting in a synoptic map with a size of 360 by 180 pixels in longitude-sine-latitude, under the assumption that the magnetic field in the photosphere is radial. These are the maps used in the simulations of this work.

Thus, the measurements at the highest latitudes are less reliable due to a combination of projection effects and noise. Also, due to the inclination of the orbit of the Earth relative to the solar equator, solar poles are not observable during some periods of time, and it is customary to fill the missing pixels with interpolated data. The polar areas in these maps were filled using cubic spline fitting over the periods when the poles are not well observed, but the values are known to be partly erroneous (Harvey 2000; Harvey & Munoz-Jaramillo 2015). There are several periods in KP data with known errors, especially during the 512-channel magnetograph era. The polar fields in late 1980s and early 1990s are unreliable. The average strength of the southern polar field increases in a step-like manner by almost four Gauss in 1985 and remains very strong for several years. During the maximum of cycle 22 the average strengths of both polar fields often change by several Gauss between two rotations. Similar behaviour can also be seen during the maximum of cycle 21. However, the problems mainly relate to polar fields (Virtanen & Mursula 2016), not active regions, which means that our simulations are not affected by this problem. Therefore one could expect that the amplitudes of polar fields obtained from an SFT simulation might not agree very closely with those of the original synoptic maps. We note that there is an ongoing effort to recalibrate Kitt Peak measurements (Harvey & Munoz-Jaramillo 2015), but the recalibrated data is not yet available. Because of many errors in data we excluded the period of the 40-channel photoelectric magnetograph, 1970–1974, and the first years of the 512-channel diode array magnetograph, 1974–1978, from this study.

The SFT simulations are sensitive to the polar fields of the first rotation, but after that only observations of active regions are used as input and can affect the results. We started the simulation from January 1978 (CR 1664), which has no data gaps or obvious errors, meaning that there were no missing longitude strips or large areas of unrealistically strong magnetic fields.

Figure 1 shows the KP and SOLIS data sets over the time interval studied here. Each vertical line was formed by averaging a synoptic map over longitude. We call this presentation a super-synoptic map. The data set started from January 1978 (CR 1664) and ended in May 2016 (CR 2177). The switch from the earlier Kitt Peak instrumentation to SOLIS happened in August 2003 between rotations 2006 and 2007. On July 21, 2014, SOLIS was relocated from Kitt Peak to Tucson, Arizona, leading to the longest data gap seen also in Fig. 1. As described above,

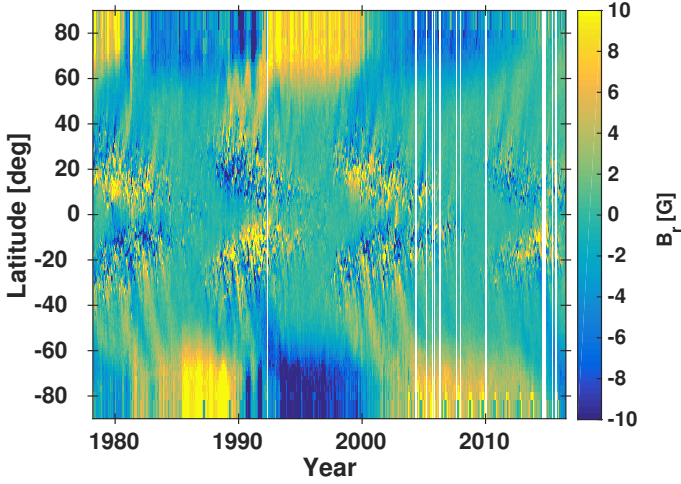


Fig. 1. Super-synoptic map of the KP and SOLIS photospheric magnetic field observations from January 1978 (CR 1664) to May 2016 (CR 2177). Yellow and blue tones correspond to positive and negative polarity. White vertical lines are periods with no observations.

the synoptic maps contain polar field values obtained by filling. However, the filling method is not perfect, as indicated by the rapidly alternating colors of the polar-most bins of the super-synoptic map. This does not affect our results because the polar fields were obtained from simulations with the SFT model. The white vertical lines are periods with no observations. The data gaps were not filled, meaning that no active regions were assimilated into the simulation during rotations with no observations.

3. The SFT model

We used the SFT model of Yeates et al. (2015) to model the evolution of the photospheric magnetic field and compare it with the observed evolution. The radial magnetic field $B_r(\theta, \phi, t)$ can be written in a spherical coordinate system as a function of the two-component vector potential $[A_\theta, A_\phi]$:

$$B_r(\theta, \phi, t) = \frac{1}{R \sin(\theta)} \left(\frac{\partial}{\partial \theta} (\sin(\theta) A_\phi) - \frac{\partial A_\theta}{\partial \phi} \right) \quad (1)$$

where R is the radius of the Sun, θ is the colatitude, ϕ is the azimuth angle, and t is the time. The vector potential evolves in time as follows (Yeates et al. 2015):

$$\frac{\partial A_\theta}{\partial t} = \omega(\theta) R \sin(\theta) B_r(\theta, \phi, t) - \frac{D}{R \sin(\theta)} \frac{\partial B_r(\theta, \phi, t)}{\partial \phi} + S_\theta(\theta, \phi, t) \quad (2)$$

$$\frac{\partial A_\phi}{\partial t} = -u_\theta(\theta) B_r(\theta, \phi, t) + \frac{D}{R} \frac{\partial B_r(\theta, \phi, t)}{\partial \theta} + S_\phi(\theta, \phi, t) \quad (3)$$

where $\omega(\theta)$ is the differential angular velocity of rotation, D is the supergranular diffusivity coefficient, $u_\theta(\theta)$ is the velocity of the meridional circulation, and S_θ and S_ϕ are the source terms that represent the emergence of new flux.

We used the angular velocity of differential rotation $\omega(\theta)$ (in the Carrington rotation frame) of Yeates et al. (2015):

$$\omega(\theta) = (0.521 - 2.396 \cos^2(\theta) - 1.787 \cos^4(\theta)) \text{ deg/day}. \quad (4)$$

We did not investigate here the effects of possible variations or uncertainties in differential rotation speed because differential rotation parameters are better constrained than the parameters

studied here. Moreover, if the parameters are kept within a sensible parameter range they do not significantly affect the average strength of polar fields in SFT simulations (Baumann et al. 2004).

For meridional circulation, the profile of Schüssler & Baumann (2006) was used. It is a semi-empirical profile that has been adapted to the profile obtained from helioseismic measurements:

$$u_\theta(\theta) = u_0 \sin\left(2\left(\frac{\pi}{2} - \theta\right)\right) \exp\left(\pi - 2\left|\frac{\pi}{2} - \theta\right|\right) \quad (5)$$

where u_0 is the peak velocity. The optimization of these parameters has been studied by, for example, Yeates (2014) and Lemerle et al. (2015), and is further discussed in Sect. 4.1 (see also Whitbread et al., in prep.).

The source terms S_θ and S_ϕ represent the emerging flux of active regions. It should be noted that they are included in the equations only for the sake of clarity and have no functional form. In practice the active regions are copied from the observations directly to the simulated radial field $B_r(\theta, \phi, t)$, not via the vector potential.

The radial magnetic field at the start of the simulation was defined by the first observed synoptic map. Thereafter only active regions of the remaining maps are used as input. The simulation computed the vector potential from the radial field by setting $A_\phi = 0$ every time the vector potential was solved. Because of freedom of gauge we could make this choice as long as A_θ is chosen so that $\mathbf{B} = \nabla \times \mathbf{A}$. Equation (1) then gives:

$$A_\theta = -R \sin(\theta) \int_0^\phi B_r(\phi', \theta) d\phi'. \quad (6)$$

The vector potential was evolved in time according to Eqs. (2) and (3). Due to Eq. (3), A_ϕ becomes non-zero, even though it was set to zero at the beginning of each time step. The new radial field is then computed from Eq. (1). Active regions, the source of emerging flux in the simulation, were added to the radial field at the time when they crossed the central meridian.

The differential equations were solved using a finite-difference method. To ensure convergence, the time step of the numerical integration must be small enough to fulfill the Courant-Friedrichs-Lewy condition (Courant et al. 1967):

$$\Delta t \leq \frac{\Delta x}{u} \quad (7)$$

where Δt is the time step, Δx is the spatial step, and u is the velocity of a process. Therefore the maximum value of Δt depends on the parameters of the model and the spatial resolution of the data. We computed the maximum value of Δt that satisfied Eq. (7) for differential rotation and meridional circulation. For supergranular diffusion D the criteria is:

$$\Delta t \leq \frac{(\Delta x)^2}{D}. \quad (8)$$

We set the time step of the simulation to be 20% of the smallest acquired value to ensure convergence. In the simulations presented in this paper the time step is typically between 10 and 20 min.

3.1. Additional decay term

The above model, which we initially used, is successful in reproducing the properties of the polar fields of cycles 21–23. However, we later found that the model fails to reproduce the polarity

Table 1. Decay times τ_l .

l	1	2	3	4	5	6	7	8	9
τ_l [day]	1888	1623	1340	1088	881	718	591	491	413

reversal of cycle 24. This will be shown later in this work. We attribute this failure to the weak sunspot activity of cycle 24, which was insufficient to cancel out the magnetic flux of the previous, stronger cycle. Thus, we extend the SFT model of Yeates et al. (2015) with the decay term suggested by Baumann et al. (2006). The purpose of this term is to remove the long-term memory of the model. Physically, this corresponds to the radial diffusion of the magnetic flux into the convection zone. Without it, a weak cycle might not be able to reverse the polarity of strong polar fields generated during a previous stronger cycle. In SFT models the evolution of the magnetic field is governed by supergranular diffusion, which simply spreads the magnetic field over a larger area, and the meridional flow, which transports flux to polar regions. Without emergence of new flux, these two processes will lead to a preservation of magnetic flux in polar areas. Even if new flux emerges, the flux transported to the poles may be insufficient to reverse the polarity of the magnetic field. Additional diffusion is necessary to overcome this limitation.

Baumann et al. (2006) assume that the magnetic field \mathbf{B} is radial at the surface and that the radial component disappears at the bottom of the convection layer at $0.7R$. The poloidal field can be written in terms of a scalar function $S(r, \theta, \phi, t)$:

$$\mathbf{B} = -\nabla \times (\mathbf{r} \times \nabla S), \quad (9)$$

which obeys the diffusion equation:

$$\eta \nabla^2 S - \frac{\partial S}{\partial t} = 0. \quad (10)$$

The solution of the diffusion equation is written as a sum of orthogonal decay modes:

$$S(r, \theta, \phi, t) = \sum_{n=0}^{\infty} \sum_{l=1}^{\infty} \sum_{m=-l}^l R_{ln}(r) Y_{lm}(\theta, \phi) T_{ln}(t) \quad (11)$$

where $Y_{lm}(\theta, \phi)$ are spherical harmonics. The monopole term $l = 0$ has been omitted. Equation (10) implies that $T_{ln}(t)$ has the following time dependence:

$$T_{ln}(t) = \exp(-\eta k_{ln}^2 t) \quad (12)$$

where $\eta = 100 \text{ km}^2 \text{ s}^{-1}$ is the volume diffusion coefficient. This leads to an exponential decay of each mode characterized by l and n , and the decay time $\tau_{ln} = 1/(\eta k_{ln}^2)$ depends on the value of k_{ln} , which can be solved numerically. We considered only the modes where $n = 0$ and $l \leq 9$. The decay times τ_l for these modes are shown in Table 1. The decay times decrease for larger values of l . The highest mode used in the computation of the decay term, $l = 9$, had a decay time of about a year. Higher modes could be omitted because they do not affect the long-term evolution of the field, which we are studying. Radial modes higher than $n = 0$ have even shorter decay times and could be omitted for the same reason (Baumann et al. 2006).

The radial surface field can be presented with spherical harmonics:

$$B_r(\theta, \phi, t) = \sum_{l=1}^{\infty} \sum_{m=-l}^{m=l} c_{lm}(t) Y_{lm}(\theta, \phi), \quad (13)$$

where c_{lm} are coefficients that are evaluated for the synoptic map observed at time t . Combining this with the above decay times leads to the following decay rate $D(\theta, \phi, t)$ of the radial surface field:

$$D(\theta, \phi, t) = \sum_{l=1}^{l=9} \frac{1}{\tau_l(k_l)} \sum_{m=-l}^{m=l} c_{lm}(t) Y_{lm}(\theta, \phi). \quad (14)$$

The decay term was computed and subtracted from the simulated field every 100th time step, meaning that this part of the simulation ran at a lower time resolution than the rest. Because one time step is typically 10 to 20 min, the decay was computed roughly once a day. This was done to save computation time and doing so did not reduce the accuracy of the simulation because the decay process is very slow compared to the other processes included in the model, as can be seen from the decay times in Table 1.

3.2. Emerging flux

We identified active regions from each synoptic map of the KP and SOLIS data sets with the following method. The overall flux imbalance, meaning the total sum of signed flux of the map, was first removed by subtracting the overall mean value of the synoptic map from every pixel. The maps were then combined to form a continuous series of observations from the beginning of the first map in the data set to the end of the last, effectively creating one large map. This allowed us to treat active regions that are located on longitude zero or 360 degrees, meaning those that reside on two different maps at the same time, as one active region rather than two regions. The combined maps were then transformed to absolute value maps and smoothed with a Gaussian filter. The width of the filter was 4 pixels in both latitude and longitude and the standard deviation was 2 pixels. The purpose of the filtering was to smooth out small gaps between areas of negative and positive polarities within an active region. Otherwise the negative and positive parts could be identified as two separate active regions, which would then be discarded during the flux balancing phase, which will be discussed later.

We used a threshold of 50 G to identify active regions from the smoothed map. Any connected group of pixels with a magnetic field greater than the threshold counted as an active region. The identified regions were then selected from the unsmoothed map pixel by pixel. The filter width and threshold were selected with trial and error. As long as the threshold was clearly above the strength of the background field and the filter was not excessively wide they had only a minor effect on how the edges of the active regions were defined and how much flux was contained within each active region.

Flux imbalance was removed from each active region by subtracting the active region mean value from every pixel inside the region. If the total imbalance of the region was more than half of the total unsigned flux inside it, meaning if the weaker field was less than one third of the stronger, the region was discarded. This prevented areas of strong magnetic fields in the polar regions and inside poleward flux surges from being identified as active regions, and left only the newly emerged active regions, which are typically fairly balanced in negative and positive flux. Prior to applying the flux balance condition the majority of the selected regions were not emerging active regions in the activity belts, but small areas of strong unipolar magnetic field at higher latitudes. This was especially true for the KP observations, which often contained areas of very strong magnetic fields at high latitudes.

Every remaining active region was added to the simulation at the time when the center of the region crossed the central meridian. The added region replaced the values of the corresponding pixels in the simulation. Any flux imbalance that existed inside the area in the simulation was kept by adding the average imbalance to every pixel of the substitute region. This was done to preserve the overall flux balance of the simulation.

4. Testing the SFT model

In light of the intended future use of the SFT model on historical data, we investigated the sensitivity of the model when varying the values of some of the main parameters. In particular, we investigated the outcome of the model for different diffusion coefficients, different meridional circulation speeds, different initial polar fields, and a simplified representation of emerging flux.

4.1. Testing parameters for meridional circulation and diffusion

Figure 2 shows nine simulations that were run using different peak velocities u_0 of meridional circulation and different supergranular diffusion coefficients D . Peak velocity increases from left column, $u_0 = 9 \text{ m s}^{-1}$, to middle column, $u_0 = 11 \text{ m s}^{-1}$, and right column, $u_0 = 13 \text{ m s}^{-1}$. Diffusion coefficient increases from top row, $D = 300 \text{ km}^2 \text{ s}^{-1}$, to middle row, $D = 400 \text{ km}^2 \text{ s}^{-1}$, and bottom row, $D = 500 \text{ km}^2 \text{ s}^{-1}$. The parameter range was selected based on earlier studies on SFT parameters (Yeates 2014; Lemerle et al. 2015) and meridional circulation (Hathaway & Rightmire 2010).

Increasing supergranular diffusion coefficient D makes the poleward surges wider and causes them to merge into a polar field at a lower latitude. Increasing meridional circulation speed has the opposite effect, making the flux surges reach the poles faster, hence giving them less time to diffuse.

In the top right panel, where the meridional circulation speed is high and supergranular diffusion coefficient low, the polar fields exist only at high latitudes and decay fairly quickly. This is especially true for the southern polar field of cycle 22, which is formed by a single strong poleward surge. On the other hand, in the bottom left panel, where supergranular diffusion is strong and meridional circulation slow, the polar fields form at fairly low latitudes. Both of these situations are inconsistent with the observations. Based on this visual analysis and comparison of Figs. 1 and 2 we selected $D = 400 \text{ km}^2 \text{ s}^{-1}$ and $u_0 = 11 \text{ m s}^{-1}$ (middle panel) as the reference parameters. This is also consistent with Yeates (2014) and Hathaway & Rightmire (2010).

However, we also find that the simulation is not very sensitive to the exact values of u_0 and D . Even though the relative change in parameter values between the panels of Fig. 2 is significant, there are no dramatic changes in the simulated field, and several parameter pairs produce almost identical simulations. The u_0 and D parameters mostly affect how the poleward surges develop and polar fields spread. The polar field strength and the times of polarity reversals are not significantly affected.

4.2. Testing the structure of active regions

To study how the shape of the active regions affects the simulation we replaced all active regions identified in input magnetograms (see Sect. 3.2) with artificially constructed active regions. For this exercise, we took into account the location, tilt angle, and total flux of each observed active region. The tilt angle

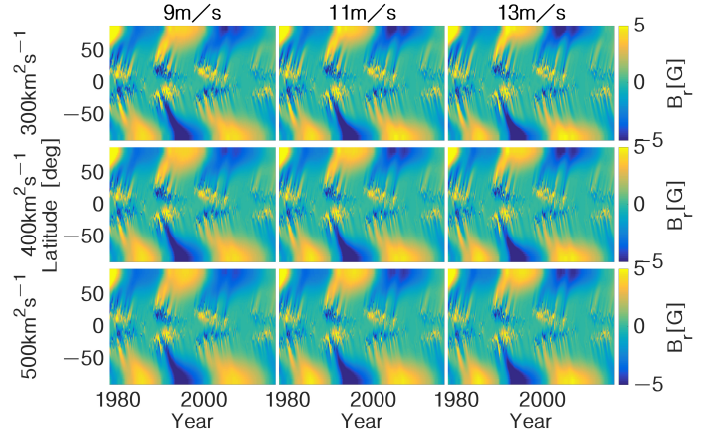


Fig. 2. Super-synoptic maps from nine SFT simulations from January 1978 (CR 1664) to May 2016 (CR 2177) using different values of meridional circulation peak velocity u_0 and supergranular diffusion coefficient D . Yellow and blue tones correspond to positive and negative polarity.

was defined as the angle between the line that connects the centroids of opposing polarities and the east-west direction. Figure 3 shows one active region from February 1978 (CR 1665) and the generated artificial substitute as an example.

It should be noted that the purpose of this exercise is not to create artificial regions that are identical to observed regions, but to demonstrate that the simulation is not sensitive to the exact structure of the active regions. The parameters that define the properties of the artificial active regions were selected by trial and error and are not based on observations. The artificial substitute regions were created by the following method:

We created a bipolar region consisting of two circles of equal radius r , one containing positive flux and one containing negative flux. In the following, Φ_{AR} is the total flux of the observed active region. To define the radius r we took the absolute flux in every bipole pixel i to be $\Phi_i = \Phi_0 = 200 \text{ G}$ and the area of the active region to be $2\pi r^2$. These assumptions were made only to define the radius and were not used in the construction of the final substitute region. The radius r was then

$$r = \sqrt{\frac{\Phi_{AR}}{400\pi}}. \quad (15)$$

The distance between the centers of the circles was set to be the radius r , which meant that the circles partly overlap. The tilt angle, meaning the angle between the line connecting the centers of the circles and the east-west direction, was set to be the same as the tilt angle of the observed region.

We then set the flux distribution of both the positive and negative part to be a two-dimensional normal distribution centered at the center of the circle. The variance of the distribution in either direction was $1.5r$. The distributions were normalized so that the absolute flux inside the region was the same as the flux inside the observed active region. In the region where the negative and positive circles overlapped the distributions were added together.

Finally we added random variation to the artificial active region by adding to each pixel a value that is 25% of the maximum field intensity of the artificial region times a random number drawn from the standard normal distribution. After this the region was again normalized so that the total absolute flux matched the observations and was placed on the centroid of the observed region.

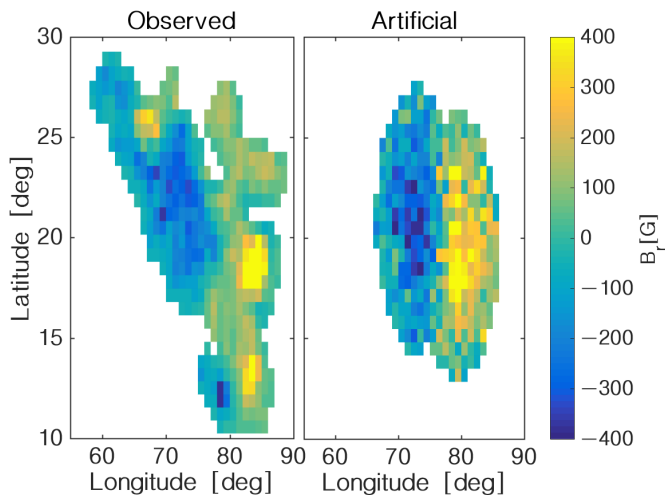


Fig. 3. An active region in February 1978 (CR 1665) and its artificial substitute. The artificial active region consists of two circles, which appear as ellipses due to unequal scaling in x and y axes. Yellow and blue tones correspond to positive and negative polarity.

As seen in Fig. 3, this leads to artificial regions that have a size that is roughly comparable to the size of the real active regions. The substitute may be smaller or larger than the observed region depending on the flux distribution of the observed region. In the case of the region shown in Fig. 3 the substitute is slightly smaller than the observed region and has a simpler flux distribution.

The top panel of Fig. 4 shows the photospheric field from a simulation run using the artificial substitutes of the observed active regions as input. For comparison, the bottom panel shows the field from a simulation that uses the observed active regions. The observed field of January 1978 (CR 1664) was used as the starting point. The peak velocity of meridional circulation is 11 m s^{-1} , and the diffusion coefficient is $D = 400 \text{ km}^2 \text{ s}^{-1}$. The most notable difference between the two simulation runs is the second strong southward flux surge that starts in January 1992 (CR 1851) and is marked with a black arrow in the top panel of Fig. 4 but is not present in the observations. In this case, the reshaping of the active regions to their artificial substitutes is momentarily producing an intense surge of negative flux that is transported southward. This is caused by large active regions during this period of high activity, which have complex structures that are very different from the simple artificial regions. We note, however, that the simulation with the observed active regions shows a number of smaller surges of negative flux at the same time, roughly balancing the effects of the intense surge of the artificial run to the overall polar field strength and the timing of the following polarity reversal. The long-term polar field contribution depends predominantly on the axial dipole moment of the active region, because it is the slowest decaying mode.

As an additional note, the tests of replacing actual active regions with simple bipoles also suggest that the procedure of preserving the observed flux distribution described in Sect. 3.2 is not critical for successful modeling of polar fields. Model runs that replace actual field distributions with balanced artificial bipolar distributions are capable of reproducing the polar fields very well.

4.3. Testing the effect of the initial polar field

The magnetic field at the start of all the above simulations was defined by the observed synoptic map of the first rotation. To

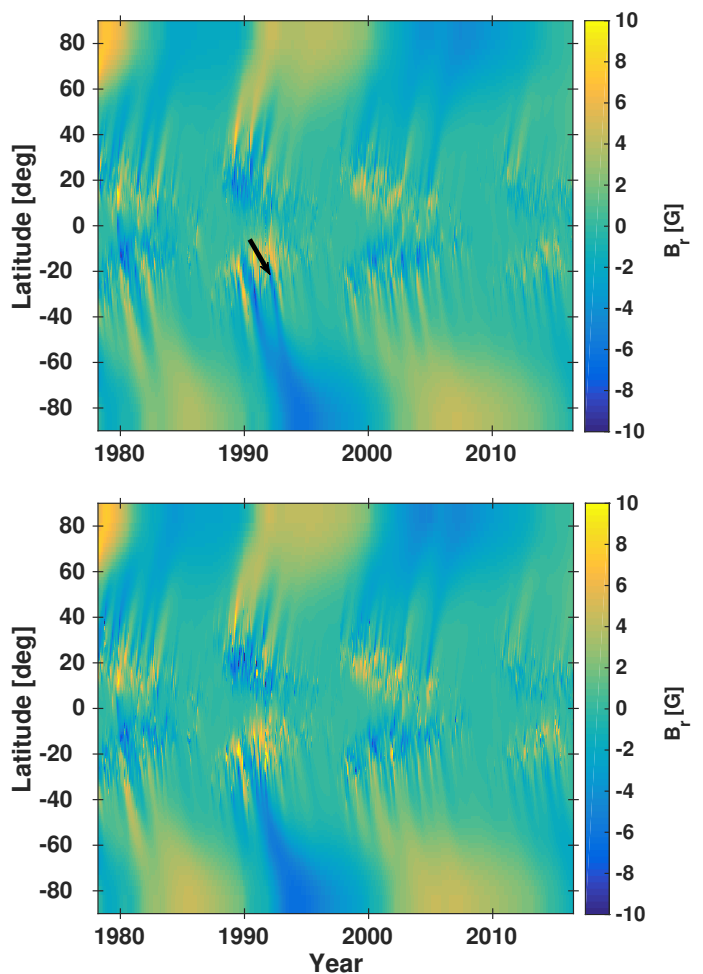


Fig. 4. Super-synoptic maps of SFT simulations run using artificial active regions (*top panel*) and observed active regions (*bottom panel*). Yellow and blue tones correspond to positive and negative polarity. The black arrow marks a poleward surge that appears in the top panel but not in the bottom panel.

determine how sensitive the simulation is to the change of the first map we ran simulations using artificial synoptic maps as starting points. Figure 5 shows three super-synoptic maps from three simulations that used artificial initial maps with a very simple structure of the overall magnetic field. The artificial maps have a constant polar field above 60° and below -60° latitude and are zero elsewhere. The strength of the polar field is $\pm 2.5 \text{ G}$ in the top panel, $\pm 5.0 \text{ G}$ in the middle panel and $\pm 7.5 \text{ G}$ in the bottom panel. The field is positive in the north and negative in the south. All three simulations start from January 1978 (CR 1664). The average strength of the simulated polar fields is shown in Fig. 6.

The differences in polar field strength, which are very clear at the beginning of the simulations, gradually disappear with time. The simulations start from the ascending phase of cycle 21, and the polar field strengths have converged to almost identical values by the polarity reversal of cycle 22. This convergence of simulations with different starting field strengths is mostly due to the additional decay term in the model (see Sect. 3.1). To demonstrate the importance of the decay term, Fig. 6 also includes polar fields from a simulation that was run without the additional decay term and with an initial polar field of $\pm 7.5 \text{ G}$. In this case, the field does not converge to the fields of the other simulations. The strong initial polar field causes the simulation to alternate

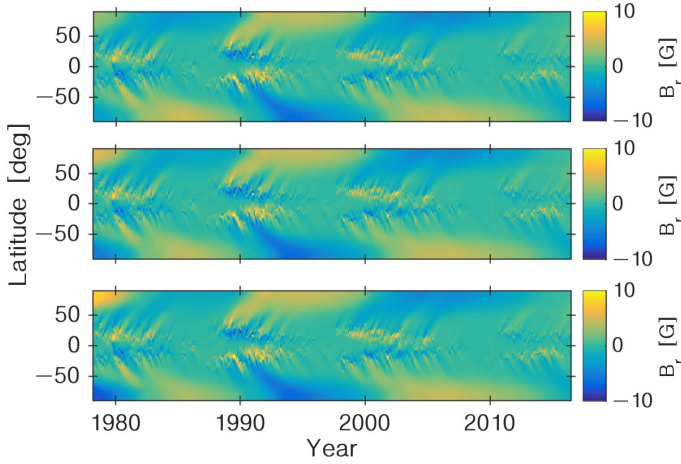


Fig. 5. Super-synoptic maps of three different SFT simulations that use artificial synoptic maps for the first rotation. The simulations start from January 1978 (CR 1664). *Top panel:* field strength is ± 2.5 G above and below $\pm 60^\circ$ latitude and zero elsewhere. *Middle panel:* ± 5.0 G above and below $\pm 60^\circ$, zero elsewhere. *Bottom panel:* ± 7.5 G above and below $\pm 60^\circ$, zero elsewhere. Yellow and blue tones correspond to positive and negative polarity.

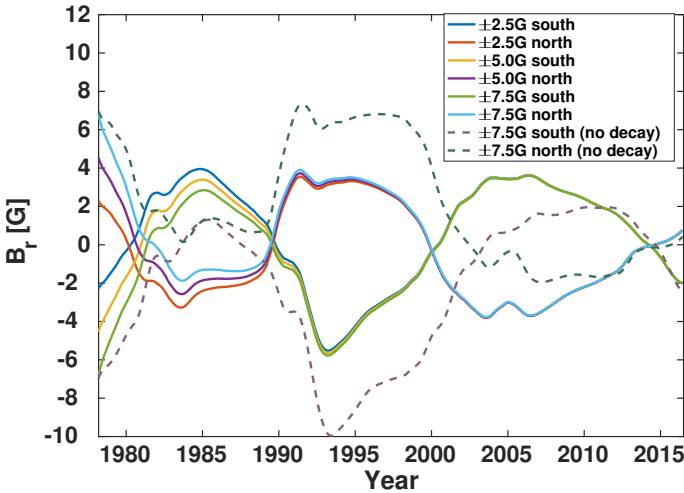


Fig. 6. Average field strength in the polar regions (above and below $\pm 60^\circ$). Solid lines are from the simulations shown in Fig. 5. Dashed lines are from a simulation without the additional decay term.

between strong and weak polar fields, and this pattern does not fade away within the time frame of the simulation.

Figure 7 shows a simulation that was started in March 1982 (CR 1720) when the polar fields were weak and about to reverse their polarity. The simulation starts with an empty map, meaning that the magnetic field is set to zero at CR 1720. In this case, the polar fields remain very weak in the declining phase of cycle 21. This is because the strong poleward surges that started in the ascending phase of the cycle around 1980, as can be seen in Fig. 1, are not included in the simulation. The active regions that created the surges appeared before the start of the simulation and the surges themselves are not included in the first synoptic map of this run. The polar fields start forming only after the flux from the first active regions that appear after CR 1720 reach high latitudes. By the declining phase of cycle 22, in the early 1990s, the polar fields are as strong as in Fig. 6, meaning that the uncertainty in the initial field affects the simulation for only one solar cycle.

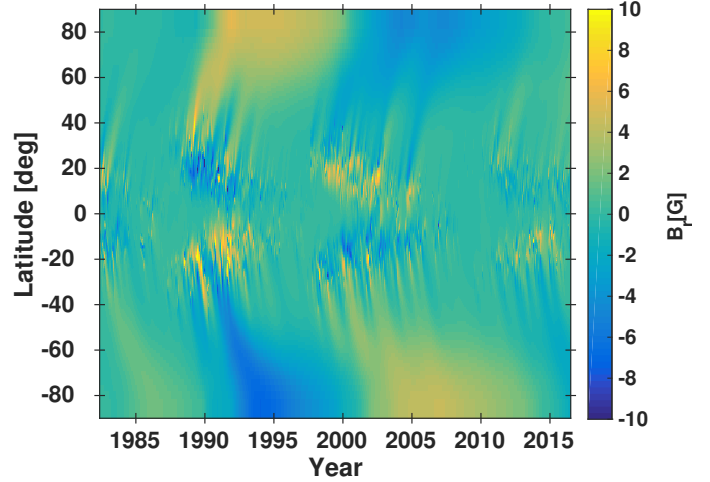


Fig. 7. Super-synoptic map from an SFT simulation that was started from March 1982 (CR 1720) with no initial magnetic field.

5. Comparing modeled and observed field evolution

The top panel of Fig. 8 shows a larger version of the middle panel in Fig. 2, meaning the simulation where $D = 400 \text{ km}^2 \text{ s}^{-1}$ and $u_0 = 11 \text{ m s}^{-1}$, and the bottom panel shows the observed field for comparison. The activity belts of the modeled and observed magnetic fields are almost identical, which is expected because the observed active regions were used in the simulation. Most of the poleward surges produced by the active regions behave similarly at mid-latitudes, having similar strengths and slopes and causing the polar fields to change polarity at roughly the same times in simulations and observations.

There are some differences in how the surges connect to the polar fields at high latitudes, especially around and after the maximum of cycle 22. However, rather than reflecting problems in the simulation, these differences may be due to problems in the polar fields of the KP maps, which are known to have errors (Virtanen & Mursula 2016). This interpretation is even likely, because KP data depict uniquely large annual variations at this time, as seen in the bottom panel of Fig. 8. Also, the several sudden polarity reversals, especially in the northern polar field around 1990, are known to be errors in the data (see, e.g., Virtanen & Mursula 2016), which is why the smooth transition of the simulation is probably closer to reality.

The simulation starts in the ascending phase of cycle 21, about forty rotations before the polarity reversal of cycle 21. The large active regions in the beginning of the simulation cause large poleward surges of flux in both hemispheres that eventually turn the polarity of both poles. During the next solar minimum, the maximum polar field is stronger in the south than in the north. The northern field is weakened by a large surge of positive flux around December 1982 (CR 1730). This development is more clearly presented by simulations than observations, in which weakening of the northern field is less systematic, most likely due to the vantage point effect and other problems related to polar fields. We note also that the observed southern field is intensified in a step-like manner in March 1985 (CR 1760), indicating another error in KP data (Virtanen & Mursula 2016) and leading to artificially large fields at the southern pole for many subsequent years.

In simulations, the polarity reversal of cycle 22 takes place ten to twenty rotations earlier at the north pole than at the south pole. Both reversals are caused by large clusters of active regions and the subsequent poleward surges, the northward surge

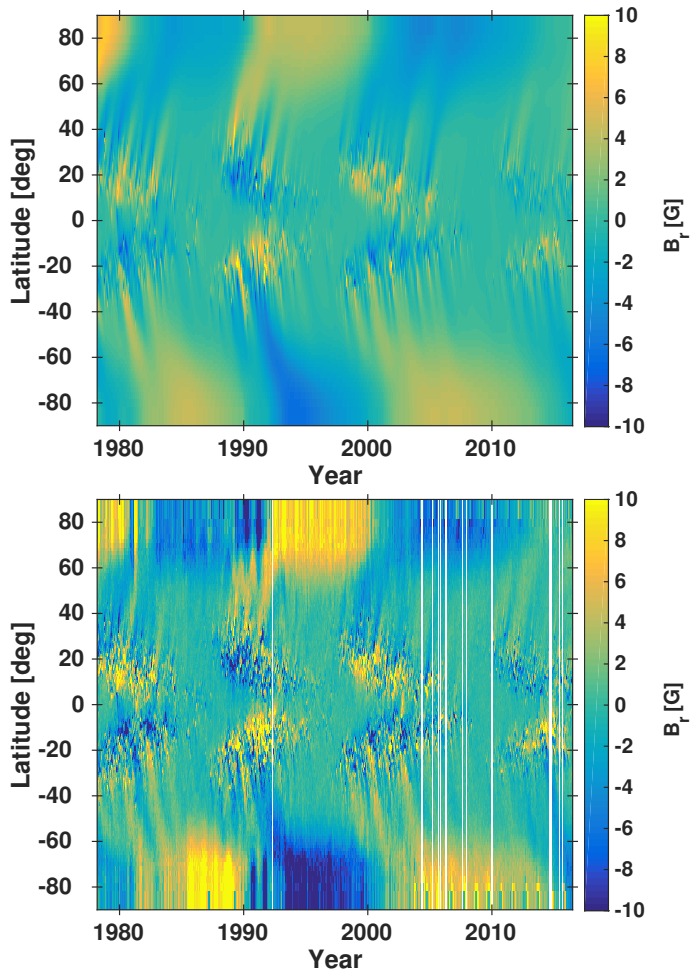


Fig. 8. Super-synoptic map of the simulated photospheric magnetic field from February 1978 (CR 1665) to May 2016 (CR 2177) (*top panel*) and the observed field (*bottom panel*). Yellow and blue tones correspond to positive and negative polarity. White vertical lines are periods with no observations.

of positive flux starting in April 1989 (CR 1815), and the southward surge of negative flux in June 1990 (CR 1830). The polar field again grows to be stronger in the south than in the north during the late declining to minimum phase, but field intensity maxima in both hemispheres are stronger than during the previous minimum. This hemispheric asymmetry of polar fields is also seen in the measured KP data (Virtanen & Mursula 2016) and is one demonstration of the "bashful ballerina" phenomenon (Mursula & Hiltula 2003).

The polarity reversal of cycle 23 also takes place earlier in the north than in the south. There is a continuous stream of weak poleward surges throughout the cycle. This is most clearly seen in the southern hemisphere. On the other hand, during cycle 22 the poleward transportation of flux was more clearly connected to a few strong surges that were also responsible for the polarity reversal and the formation of the polar fields.

During the deep minimum between cycles 23 and 24 the polar fields decay steadily in the absence of sufficiently large active regions. In 2014 the first poleward surges of cycle 24 reach the poles and reverse the polarity. However, because the cycle is weak and there are only a few large active regions, the polar fields remain weak until the end of the simulation.

Figure 9 shows the average field strength above 60° latitude and below -60° latitude, meaning the mean field strength in the

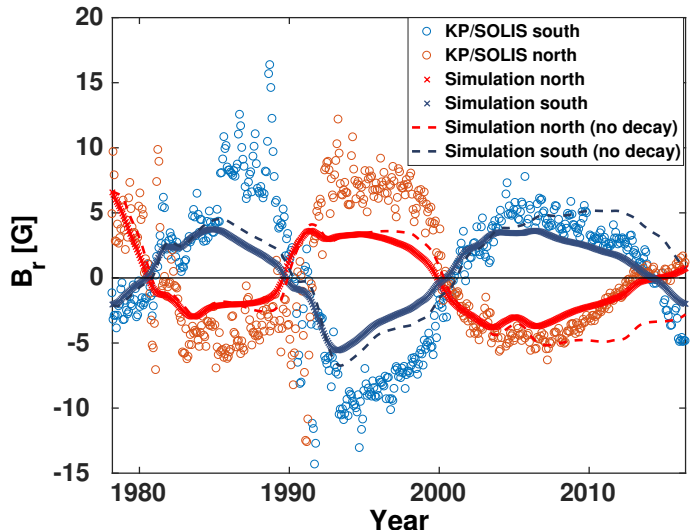


Fig. 9. Average field strength above 60° and below -60° latitude from the reference simulation shown in the top panel of Fig. 8 (crosses), a simulation without the decay term (dashed line) and the KP and SOLIS data (open circles). The solid black line is the zero-line.

polar region, for simulations with and without the decay term and for observations. The differences between the simulations and the observations are clearly visible. The largest differences occur in the southern hemisphere in 1985–1989. These are at least partly caused by the sudden and significant strengthening of the polar field in the KP data in May 1985 (CR 1762), which is also clearly visible in the bottom panel of Fig. 8. This is most likely an error in the KP data.

The maximum value of the simulated field of cycle 21 is higher in the south, reaching 3.7 G, whereas the corresponding value in the north is -3.0 G. The same asymmetry is true, even more clearly, for cycle 22, where the southern field reaches -5.6 G and the northern 3.6 G. However, during cycle 23 there is no significant hemispheric asymmetry in the simulated fields. The maximum value is 3.6 G in the south and 3.8 G in the north.

The polarity reversals of cycles 21–23 happen from a few months to a few years earlier in simulations than in observations, but the exact timings of the observed reversals of cycles 21 and 22 are difficult to determine because of the very large and erroneous variations of field strength from rotation to rotation that cause the polar fields to change polarity multiple times over several years. The first reversal of the northern field during cycle 24 seems to happen earlier in the observations, in 2013, whereas the simulated field does not change polarity until 2015. However, there is large variance in the observed values during this time, and some of them are still negative in 2015. The final reversal in the north is observed in 2015. The reversal of the southern field happens within only a few rotations in simulation and observations in 2014.

The necessity of the decay term is clearly demonstrated in Fig. 9. Without it the simulation starts to significantly deviate from the observations during the deep minimum between cycles 23 and 24. Without a sufficient emergence of active regions in the ascending phase of cycle 24, and without the additional decay term, the polar fields remain almost constant for several years of cycle 24 and fail to reverse polarity by the end of simulations.

The polar fields are on average weaker in the simulation than in the observations. This is at least partly due to the errors that affect the observed polar fields especially in the 1980s and early 1990s and are discussed above. During cycle 23, which is

covered by the more reliable SOLIS observations, the simulation follows the observed polar fields more closely than during cycles 21 and 22.

One should also note that the observations used in this work are from three different instruments and their mutual intensity scaling is not known in detail. The KP 512-channel magnetograph and SPMG operated simultaneously for a short period in 1992–1993. The linear scaling factor of 1.46 has been used to scale the KP-512 full-disk observations to SPMG level. However, based on 22 days of observations between November 28, 1992 and April 10, 1993, scaling factors in the range of 1.38–1.63 were derived (Wenzler et al. 2006). An alternative method, where KP-512 observations are corrected without comparisons with other instruments (Arge et al. 2002) gives a lower scaling factor of 1.242 (Wenzler et al. 2006).

The first version of the synoptic SOLIS data set was calibrated to match the magnetic flux of the SPMG synoptic maps, but the calibration process was changed after the modulator update in 2006. All SOLIS observations prior to 2006 were recalibrated using the new procedure in 2016, which increased the magnetic field intensity of the synoptic maps in 2003–2006. Therefore the SOLIS data set used in this work is not scaled to SPMG level. A comparison of SPMG and SOLIS observations based on harmonic expansions (Virtanen & Mursula 2017) indicates that the correct SPMG to SOLIS scaling factor would be 1.3–1.4.

There is an ongoing recalibration effort that will cause changes to both the KP 512-channel magnetograph and SPMG observations (private communication with Jack Harvey). Because the scaling and its effects in the simulation are non-linear, we are at the moment unable to estimate in detail how much the recalibration will change the active region flux densities and the simulated polar fields, but it is clear that polar field strength during the KP-512 period in 1974–1992 and the SPMG period in 1992–2003 will increase. We expect that more detailed information about the mutual scaling of the NSO instruments will be available in near future.

6. Discussion and conclusions

We study here the applicability of SFT models for future modeling of historical data, which may have uncertainties in several main parameters, including the initial magnetic field, supergranular diffusion, speed of meridional flow, and the precise distribution of magnetic field in active regions. We tested peak meridional flow speeds u_0 ranging from 9–13 m s⁻¹ and supergranular diffusion coefficients D ranging from 300–500 km² s⁻¹. Parameters within this range produce simulations that agree fairly well with observations. Selecting an optimal parameter set is somewhat difficult, because the two parameters can have similar effects. As a further development, Whitbread et al. (in prep.) has been developing a new automated optimisation technique. It is known that the meridional circulation may have varied in time (Hathaway & Rightmire 2010). However, we show that the simulations are not very sensitive to the exact values of the D and u_0 parameters, and that changes within the parameter range studied here, although impacting on the strength of poleward surges and polar fields, lead to a fairly similar solar cycle and long-term evolution.

We show that changing the distribution of magnetic flux inside active regions while preserving the tilt angles does not cause significant difference in model outcome (see Fig. 4). The structure of the polar fields remains intact, and significant differences

are observed only during certain short time periods involving large active regions with complex structures. The effect of these regions is temporary and does not affect the long-term evolution of the simulated field.

The structure and even the strength of the first synoptic map affects the simulation typically for only one solar cycle. Within this time period any uncertainties caused by the initial fields disappear and the field becomes completely determined by the emerging flux of the active regions.

We also verify that the additional decay term (Baumann et al. 2006) is necessary. Without it the simulations are somewhat sensitive to errors in the initial polar fields and the model does not reproduce the polarity reversal of cycle 24 correctly. The additional decay term is necessary for the polar fields to weaken.

Comparing the simulated magnetic field with the observations, we find that the activity belts and poleward surges are very similar, but the polar fields are sometimes different. During the time period from May 1985 (CR 1762) until the polarity reversal of cycle 22 in 1990 there are large differences in field strength, which are at least partly due to errors in the polar fields of the KP maps. Simulations may yield a more realistic view than observations during, for example, the polarity reversals of cycles 21 and 22 and in the mid-1980s. Simulations and observations agree with each other considerably better during the SOLIS period. We also note that simulations verify the larger maximum field strength in the southern hemisphere during the declining to minimum phase of cycles 21 and 22, as well as in the early declining phase of cycle 24, in agreement with the “bashful ballerina” (Mursula & Hiltula 2003; Mursula & Virtanen 2011; Virtanen & Mursula 2016).

We conclude that the SFT model is stable and reliable even without accurate information about the initial polar fields or the exact structure of the active regions. The initial field affects the simulation for roughly one solar cycle, whereas the structure of active regions can affect individual poleward surges but does not change the long-term evolution of the polar fields. The model is also not very sensitive to the exact values of parameters of meridional circulation or supergranular diffusion. Although the model is not able to reproduce the finer details of the field, it is robust enough to be used as a tool to study the long-term evolution of the polar fields.

Acknowledgements. We acknowledge the financial support by the Academy of Finland to the ReSoLVE Centre of Excellence (project No. 272157). The National Solar Observatory (NSO) is operated by the Association of Universities for Research in Astronomy, AURA Inc under cooperative agreement with the National Science Foundation (NSF).

References

- Arge, C. N., Hildner, E., Pizzo, V. J., & Harvey, J. W. 2002, *J. Geophys. Res. (Space Phys.)*, **107**, 1319
- Baumann, I., Schmitt, D., Schüssler, M., & Solanki, S. K. 2004, *A&A*, **426**, 1075
- Baumann, I., Schmitt, D., & Schüssler, M. 2006, *A&A*, **446**, 307
- Courant, R., Friedrichs, K., & Lewy, H. 1967, *IBM J. Res. & Dev.*, **11**, 215
- Harvey, J. 2000, Kitt Peak synoptic data documentation, <ftp://vso.nso.edu/kpvt/synoptic>
- Harvey, J., & Munoz-Jaramillo, A. 2015, in *AAS/AGU Triennial Earth-Sun Summit*, **1**, 11102
- Hathaway, D. H., & Rightmire, L. 2010, *Science*, **327**, 1350
- Jiang, J., Cameron, R., Schmitt, D., & Schüssler, M. 2010, *ApJ*, **709**, 301
- Jiang, J., Cameron, R. H., Schmitt, D., & Schüssler, M. 2011, *A&A*, **528**, A82
- Jiang, J., Hathaway, D. H., Cameron, R. H., et al. 2014, *Space Sci. Rev.*, **186**, 491
- Jones, H. P., Duvall, Jr., T. L., Harvey, J. W., et al. 1992, *Sol. Phys.*, **139**, 211

- Keller, C. U., Harvey, J. W., & Giampapa, M. S. 2003, in *Innovative Telescopes and Instrumentation for Solar Astrophysics*, eds. S. L. Keil, & S. V. Avakyan, *Proc. SPIE*, **4853**, 194
- Lean, J. L., Wang, Y.-M., & Sheeley, N. R. 2002, *Geophys. Res. Lett.*, **29**, 77
- Leighton, R. B. 1964, *ApJ*, **140**, 1547
- Lemerle, A., Charbonneau, P., & Carignan-Dugas, A. 2015, *ApJ*, **810**, 78
- Livingston, W., & Harvey, J. 1971, in *Solar Magnetic Fields*, ed. R. Howard, *IAU Symp.*, **43**, 51
- Livingston, W. C., Harvey, J., Slaughter, C., & Trumbo, D. 1976, *Appl. Opt.*, **15**, 40
- Mackay, D. H., & Yeates, A. R. 2012, *Liv. Rev. Sol. Phys.*, **9**, 6
- Mursula, K., & Hiltula, T. 2003, *Geophys. Res. Lett.*, **30**, 2135
- Mursula, K., & Virtanen, I. I. 2011, *A&A*, **525**, L12
- Pevtsov, A. A., Virtanen, I., Mursula, K., Tlatov, A., & Bertello, L. 2016, *A&A*, **585**, A40
- Pietarila, A., Bertello, L., Harvey, J. W., & Pevtsov, A. A. 2013, *Sol. Phys.*, **282**, 91
- Schrijver, C. J., De Rosa, M. L., & Title, A. M. 2002, *ApJ*, **577**, 1006
- Schüssler, M., & Baumann, I. 2006, *A&A*, **459**, 945
- Upton, L., & Hathaway, D. H. 2014, *ApJ*, **780**, 5
- Virtanen, I., & Mursula, K. 2016, *A&A*, **591**, A78
- Virtanen, I., & Mursula, K. 2017, *A&A*, in press
DOI: [10.1051/0004-6361/201730863](https://doi.org/10.1051/0004-6361/201730863)
- Wang, Y.-M., Nash, A. G., & Sheeley, Jr., N. R. 1989, *Science*, **245**, 712
- Wang, Y.-M., Lean, J., & Sheeley, Jr., N. R. 2002a, *ApJ*, **577**, L53
- Wang, Y.-M., Sheeley, Jr., N. R., & Lean, J. 2002b, *ApJ*, **580**, 1188
- Wenzler, T., Solanki, S. K., Krivova, N. A., & Fröhlich, C. 2006, *A&A*, **460**, 583
- Worden, J., & Harvey, J. 2000, *Sol. Phys.*, **195**, 247
- Yeates, A. R. 2014, *Sol. Phys.*, **289**, 631
- Yeates, A. R., Baker, D., & van Driel-Gesztelyi, L. 2015, *Sol. Phys.* **290**, 3189