Surface Flux Transport

Anthony R. Yeates[1*], Mark C. M. Cheung[2], Jie
Jiang[3], Kristof Petrovay[4] and Yi-Ming Wang[5]

[1*]Department of Mathematical Sciences, Durham University,
Durham, UK.
[2]CSIRO, Space & Astronomy, Marsfield, NSW, Australia.
[3]School of Space and Environment, Beihang University, Beijing,
People's Republic of China.
[4]Department of Astronomy, Eötvös Loránd University, Budapest,
Hungary.
[5]Space Science Division, Naval Research Laboratory,
Washington, DC, USA.

*Corresponding author(s). E-mail(s):
anthony.yeates@durham.ac.uk;
Contributing authors: mark.cheung@csiro.au;
jiejiang@buaa.edu.cn; k.petrovay@astro.elte.hu;
yi.wang@nrl.navy.mil;

Abstract

We review the surface flux transport model for the evolution of magnetic flux patterns on the Sun's surface. Our underlying motivation is to understand the model's prediction of the polar field (or axial dipole) strength at the end of the solar cycle. The main focus is on the "classical" model: namely, steady axisymmetric profiles for differential rotation and meridional flow, and uniform supergranular diffusion. Nevertheless, the review concentrates on recent advances, notably in understanding the roles of transport parameters and – in particular – the source term. We also discuss the physical justification for the surface flux transport model, along with efforts to incorporate radial diffusion, and conclude by summarizing the main directions where researchers have moved beyond the classical model.

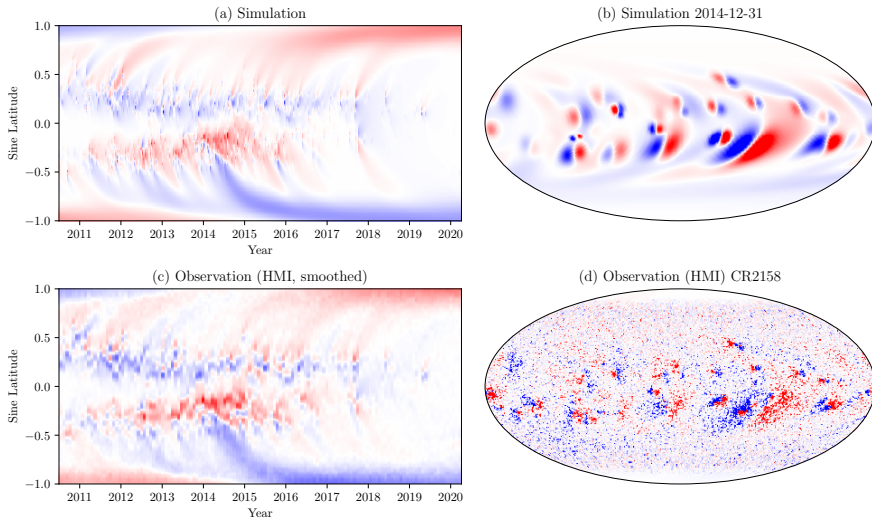**Keywords:** Sun, Solar magnetic field, Solar photosphere, Solar activity

Fig. 1 An SFT model for Solar Cycle 24 with emerging regions derived from SDO/HMI SHARPs data (following the method of Yeates, 2020). Panel (a) shows the longitude-averaged field $\langle B_r \rangle$ in the simulation, and (b) shows a snapshot of the two-dimensional field B_r on 31 December 2014. For comparison, (c) shows a magnetic butterfly diagram (or super-synoptic map) constructed from SDO/HMI pole-corrected synoptic maps (Sun, 2018), smoothed to a comparable resolution to the simulation. The individual, unsmoothed synoptic map for Carrington rotation CR2158 is shown in (d). Red/blue denote positive/negative values, capped at $\pm 10\,\mathrm{G}$ in (a,c) and $\pm 50\,\mathrm{G}$ in (b,d).

1 Introduction

The surface flux transport (hereafter SFT) model is based on an elegant and simple idea, originally formulated by Leighton (1964): radial magnetic flux on the solar surface behaves like a passive scalar field. In other words, flux is carried around by horizontal plasma flows but with no back reaction on these flows.

Despite its simplicity, the SFT model has proven remarkably successful at replicating the magnetic flux patterns on the real solar surface (photosphere). Figure 1 shows an example SFT simulation for Solar Cycle 24, where new active regions have been inserted based on magnetograph observations. With appropriate parameters, the time-latitude "magnetic butterfly diagram" in the SFT model (Figure 1a) is a good match for the observed time-latitude plot (Figure 1c) at all latitudes. In general, the success of the SFT model has led to important applications both as (i) an inner boundary condition for extrapolations of the magnetic field in the solar atmosphere, and (ii) an outer boundary constraint on models for the solar interior dynamo.

In this review, our focus is on understanding the model itself: both its key ingredients and fundamental behaviour when applied in the solar regime. Details about applications, particularly to the solar atmosphere, may be found in previous review articles (Sheeley Jr., 2005; Mackay and Yeates, 2012; Wang,

2017). In the solar dynamo context, the SFT model has been used to constrain theories and models of the magnetic field in the solar interior (e.g., Cameron et al, 2012; Cameron and Schüssler, 2015; Jiang et al, 2014b; Lemerle and Charbonneau, 2017; Whitbread et al, 2019; Hazra, 2021). But it is also a valuable practical tool for solar cycle prediction, enabling predictions to be made of the polar field at the end of the current solar cycle, and hence – through well-established correlations – the amplitude of the following solar activity cycle (e.g., Cameron et al, 2016; Iijima et al, 2017; Jiang et al, 2018; Upton and Hathaway, 2018; Bhowmik and Nandy, 2018; Jiang et al, 2022). Understanding the origin and limitations of such polar field predictions requires an understanding of the SFT model itself, which is what we seek to provide here.

The review is organised as follows. In Section 2, we present the basic equations of the "classical" SFT model. Section 3 discusses the imposed flows in the model, including the importance of including meridional flow and recent work on constraining the flow parameters. Section 4 discusses the source term representing new flux emergence, which is fundamental to the flux patterns that the model predicts. Section 5 examines the important question of whether the SFT model – usually seen as purely phenomenological – can be derived from physical principles. We conclude in Section 6 with an overview of model features beyond our "classical" version.

2 Fundamentals of the Classical Model

Denoting the radial magnetic field distribution by $B_r(\theta, \phi, t)$, the equation for a passive scalar field is

$$\frac{\partial B_r}{\partial t} + \nabla_h \cdot \left(\mathbf{u}_h B_r \right) = \eta \nabla_h^2 B_r + S, \tag{1}$$

where $\mathbf{u}_h$ is the imposed advection velocity, and η is the diffusivity. In the classical model, B_r represents the large-scale mean field, and η is a turbulent diffusivity modelling the net effect of unresolved supergranular-scale convective motions. For SFT it is necessary to include also a prescribed source term $S(\theta, \phi, t)$ that describes the emergence of new magnetic flux, typically in the form of active regions. In a more complete physical model, S would arise self-consistently through Faraday's induction equation (to be discussed in Sections 5 and 6), but in the classical SFT model it is a prescribed model input. Throughout we will use subscript h to denote the "horizontal" components of a vector, meaning those tangential to the solar surface.

In the classical SFT model, the diffusivity η is uniform and constant, and most authors assume a steady, axisymmetric imposed velocity of the form

$$\mathbf{u}_h(\theta) = R_\odot \sin\theta \, \Omega(\theta) \mathbf{e}_\phi + u_\theta(\theta) \mathbf{e}_\theta. \tag{2}$$

Thus $\Omega(\theta)$ represents the angular velocity of solar differential rotation, and $u_\theta(\theta)$ represents the meridional circulation. The choice of these flows is important and will be discussed further in Section 3. Relaxing the classical assumptions is considered in Section 6 (except for the addition of an exponential decay term which is discussed in Section 5).

2.1 Dimensionless Form

Ignoring S, we can consider non-dimensionalization of equation (1) by defining dimensionless variables $\mathbf{u}'_h = \mathbf{u}_h/U_0$, $\nabla'_h = R_\odot \nabla_h$ and $t' = tU_0/R_\odot$, where U_0 is a typical flow speed. Then (1) becomes

$$\frac{\partial B_r}{\partial t'} + \nabla'_h \cdot \left(\mathbf{u}'_h B_r\right) = \frac{1}{\text{Rm}} {\nabla'_h}^2 B_r, \tag{3}$$

suggesting that the behaviour (in the absence of new emergence) is controlled by the dimensionless magnetic Reynolds number

$$\text{Rm} = \frac{R_\odot U_0}{\eta}. \tag{4}$$

In effect, it is only the relative speed of advective to diffusive transport that matters.

2.2 Explicit Form

Writing out (1) explicitly in spherical coordinates, and assuming (2), gives the standard SFT equation

$$\frac{\partial B_r}{\partial t} + \frac{1}{R_\odot \sin\theta} \frac{\partial}{\partial \theta}\left(\sin\theta\, u_\theta B_r\right) + \Omega(\theta)\frac{\partial B_r}{\partial \phi} =$$
$$\frac{\eta}{R_\odot^2 \sin\theta} \frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial B_r}{\partial \theta}\right) + \frac{\eta}{R_\odot^2 \sin^2\theta} \frac{\partial^2 B_r}{\partial \phi^2} + S. \tag{5}$$

In some applications it suffices to consider the longitude-averaged field,

$$\langle B_r\rangle(\theta, t) = \frac{1}{2\pi} \int_0^{2\pi} B_r(\theta, \phi, t)\, \mathrm{d}\phi. \tag{6}$$

Integrating (5), we find that $\langle B_r\rangle$ obeys the one-dimensional equation

$$\frac{\partial \langle B_r\rangle}{\partial t} + \frac{1}{R_\odot \sin\theta} \frac{\partial}{\partial \theta}\left(\sin\theta\, u_\theta \langle B_r\rangle\right) = \frac{\eta}{R_\odot^2 \sin^2\theta} \frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \langle B_r\rangle}{\partial \theta}\right) + \langle S\rangle, \tag{7}$$

showing in particular that differential rotation has no effect on the evolution of $\langle B_r\rangle$ (Leighton, 1964). On the other hand, the differential rotation – being
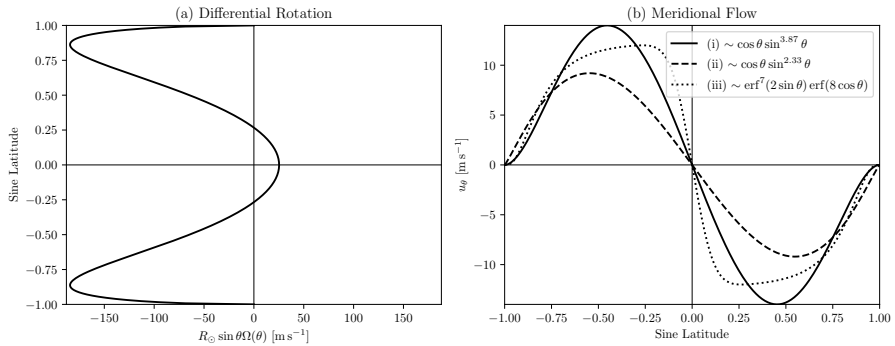
Fig. 2 Velocity profile of differential rotation in the Carrington frame (a), and some example optimized profiles of meridional flow velocity (b), including (i) the simulation shown in Figure 1; (ii) the Cycle 21 simulation of Whitbread et al (2017); and (iii) the Cycle 21 simulation of Lemerle et al (2015). The corresponding values of Δ_u are (i) $0.7 \times 10^{-7}\,\mathrm{s}^{-1}$, (ii) $0.4 \times 10^{-7}\,\mathrm{s}^{-1}$, and (iii) $1.6 \times 10^{-7}\,\mathrm{s}^{-1}$.

the fastest flow – plays an important role in determining the two-dimensional flux patterns seen on the solar surface. By increasing the length of the polarity inversion lines in and between active regions, it also speeds up the diffusive cancellation of non-axisymmetric components of B_r (Sheeley Jr. and DeVore, 1986).

2.3 Implementation

Although some analytical analysis is possible (see Sheeley Jr. and DeVore, 1986; DeVore, 1987, and also Section 4 below), for most applications it is usual to solve (5) or (7) with numerical methods. This dates right back to the original paper of Leighton (1964). The most natural numerical approach would be a spectral method based on spherical harmonics, as implemented for example by Mackay et al (2002) or Baumann et al (2004) (see Baumann, 2005, for more details). However, care is needed in treating the source term S, since newly-emerging active regions are typically highly localized in space and usually require filtering in spectral space to avoid the Gibb's phenomenon ("ringing"). A more straightforward approach is to use a simple explicit finite-volume method designed to conserve magnetic flux (i.e., preserve $\int_0^{2\pi} \int_0^{\pi} B_r(\theta, \phi, t) \sin\theta \, \mathrm{d}\theta \mathrm{d}\phi = 0$). The resulting time-step restriction is typically not a severe problem on modern machines, given the two-dimensional nature and modest resolutions typically used (for example, a 360×180 mesh). Much higher resolutions would not be consistent with the mean-field assumption of the classical model (alternatives are discussed in Section 6).
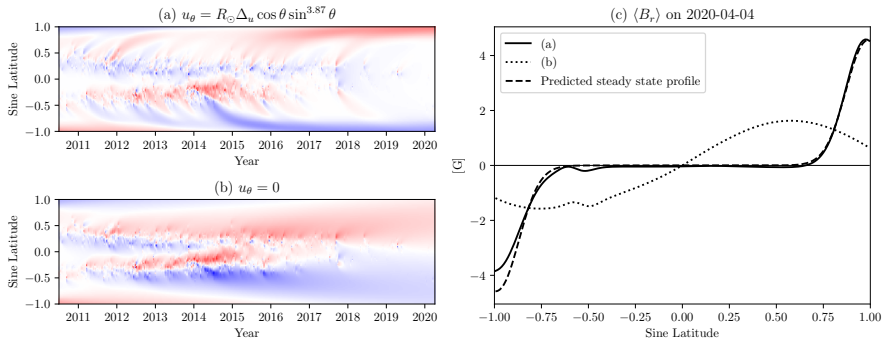
Fig. 3 Effect of meridional flow in the simulation from Figure 1, showing latitude-time plots of $\langle B_r \rangle$ when the flow is included (a) or omitted (b). Panel (c) shows the latitudinal profiles of $\langle B_r \rangle$ at the end of the simulation. The dashed curve shows the (near) steady-state profile (11) for the case with flow. (After Figure 3 of Sheeley Jr., 2005).

3 Flows

3.1 Differential Rotation

The solar surface differential rotation is well constrained observationally and usually treated as a fixed constraint. Typically, SFT models use a steady axisymmetric angular velocity profile such as

$$\Omega(\theta) = 0.18 - 2.396 \cos^2 \theta - 1.787 \cos^4 \theta \quad \left[{}^\circ \, \mathrm{day}^{-1} \right] \tag{8}$$

as determined by Snodgrass and Ulrich (1990). The constant term here is written in the Carrington frame that is usually adopted for SFT simulations. The resulting velocity profile is shown in Figure 2(a). As mentioned above, the differential rotation affects only the non-axisymmetric component of B_r, not the axisymmetric component $\langle B_r \rangle$, and will not be discussed further.

3.2 Meridional Flow

Although the only large-scale flow included by Leighton (1964) was the differential rotation, it became clear from subsequent investigation of the SFT model that adding a meridional flow gives more realistic magnetic flux distributions (DeVore et al, 1984). In particular, a poleward flow is needed in order to concentrate the magnetic field into polar caps at the end of the solar cycle – compare Figures 3(a) and (b). Otherwise, once B_r has become approximately axisymmetric it will tend to the slowest decaying ($\ell = 1$) eigenmode of the diffusion operator, which is the dipole $B_r \sim \cos\theta$. (A pure dipole is not seen in Figure 3c because it requires a few more years: the decay time for the next higher mode, $l = 2$, is $R_\odot^2 / [\eta l(l+1)] \approx 6\,\mathrm{yr}$.)

Observational evidence now clearly supports the existence of a surface meridional flow (Hanasoge, 2022) although it is much slower than the differential rotation and potentially more variable. As such, different modellers have

used different flow profiles. A typical example (from Whitbread et al, 2017) has a single peak in each hemisphere,

$$u_\theta(\theta) = R_\odot \Delta_u \cos\theta \sin^p\theta, \qquad (9)$$

illustrated by profiles (i) and (ii) in Figure 2(b). Larger values of the parameter p lead to a flow that is more concentrated at the equator. Here Δ_u is the flow divergence at the equator; the corresponding peak flow speed is

$$\max_\theta u_\theta(\theta) = \pm \frac{R_\odot \Delta_u \, p^{p/2}}{(1+p)^{(1+p)/2}} \qquad (10)$$

at $\cos\theta = \pm(1+p)^{-1/2}$.

It is non-trivial to determine the precise eigenmodes of equation (7) when meridional flow is included (DeVore, 1987), even with a simple flow profile such as (9). However, one can determine a useful approximation by seeking a perfectly axisymmetric steady state $B_r(\theta)$ that balances the poleward advection with diffusion. For example, for the flow profile (9), equation (7) can be solved in an individual hemisphere to give the steady state solution

$$B_r(\theta) = B_r(0) \exp\left[-\frac{\text{Rm}_0 \sin^{1+p}\theta}{(1+p)} \right]. \qquad (11)$$

Here $\text{Rm}_0 = R_\odot^2 \Delta_u / \eta$, which is the magnetic Reynolds number Rm from (4) with the specific choice $U_0 = R_\odot \Delta_u$, highlighting explicitly the dependence of the solution on the magnetic Reynolds number. The amplitude $B_r(0)$ will depend on the initial condition and source term S and cannot be determined directly. The solution (11) can only be an approximation to the slowest-decaying eigenfunction because it is necessarily non-zero at the equator, and will therefore generate a discontinuity at the equator when applied in both hemispheres with opposite sign. However, this discontinuity is small for typical values of Rm_0 and will lead to diffusive cancellation only on a timescale much longer than the solar cycle (cf. Cameron et al, 2010). Indeed, Figure 3(c) shows that (11) gives an excellent approximation to the latitudinal B_r profile at the end of the example simulation in Figure 3(a), particularly in the Northern hemisphere. (In the Southern hemisphere there is a remnant active region at low latitude that modifies the profile.) This simulation used $\eta = 425\,\text{km}^2\text{s}^{-1}$, $p = 3.87$, $\Delta_u = 6.9 \times 10^{-8}\,\text{s}^{-1}$, and consequently $\text{Rm}_0 \approx 79$.

3.3 Parameter Optimization

The key flow parameter to choose is the meridional flow profile $u_\theta(\theta)$, along with the diffusivity coefficient η. The basic effects of varying these parameters were investigated in the 1980s (DeVore et al, 1984; Wang et al, 1989). A more systematic parameter study was published by Baumann et al (2004), who explored the results of varying both η and the meridional flow amplitude (in

addition to properties of the source term), albeit varying only one parameter at a time and not the shape of the meridional flow profile.

More recent studies have explored the parameter space more widely, and have also attempted to optimize the parameters directly against synoptic magnetogram observations. The two most general studies are Lemerle et al (2015) and Whitbread et al (2017), who both allow the strength and shape of $u_\theta(\theta)$ to vary, in addition to η. For u_θ, Whitbread et al (2017) allowed for profiles of the form (9), whereas Lemerle et al (2015) allow for the more general (but still single-peaked) form

$$u_\theta(\theta) = -u_0 \operatorname{erf}^q(\nu \sin\theta) \operatorname{erf}^n(w \cos\theta). \qquad (12)$$

The optimal profiles from both studies for data from Cycle 21 are shown in Figure 2. At present, it is not possible to select confidently between these solutions using observations, though helioseismic measurements of the plasma flow suggest equatorial slopes Δ_u in the range $[0.6-1.2] \times 10^{-7}\,\mathrm{s}^{-1}$ – somewhere between profiles (i) and (iii) in Figure 2. Measurements based on magnetic feature tracking give lower equatorial slopes more like that of profile (ii), but it has been suggested that these are contaminated by supergranular diffusion (Dikpati et al, 2010; Wang, 2017). A recent list of observations is given in Jiang et al (2022).

Both of the above studies used the same genetic optimization algorithm, PIKAIA (Charbonneau and Knapp, 1995). They differed in their chosen goodness-of-fit functions, although both were ultimately derived from comparing to observed $B_r(\theta, \phi)$ maps. Whitbread et al (2017) gave more weight to lower latitudes (where magnetogram observations are more reliable), whereas Lemerle et al (2015) gave additional weight to the mid-latitude "transport regions" (because they represent the result of the model evolution rather than only the active region emergence) and to the axial dipole strength. At the other extreme, Petrovay and Talafha (2019) carried out another parameter study in the 1D model but focusing on only the high latitude (polar) field. This study used a synthetic (averaged) source term and fitted to average cycle properties from Wilcox Observatory polar field measurements, such as reversal time or width of the polar cap.

A robust finding in these optimization studies is a degeneracy between η and the amplitude of u_θ. This is illustrated by Figure 4, which shows that there is a long ridge of near-optimal solutions in parameter space. Increasing both parameters together tends to lead to a equally (or nearly equally) well-matched solution, perhaps explaining why different groups have been able to use quite different values of η – for example, Cameron et al (2010) use $\eta = 250\,\mathrm{km}^2\mathrm{s}^{-1}$ as their standard value whereas the simulation in Figure 1 used $\eta = 425\,\mathrm{km}^2\,\mathrm{s}^{-1}$. This degeneracy makes sense given the appearance of the magnetic Reynolds number Rm in equation (3), which is essentially the ratio of η to $|u_\theta|$. It means that SFT simulations can not be used to constrain both the meridional flow and diffusion from magnetogram observations alone.
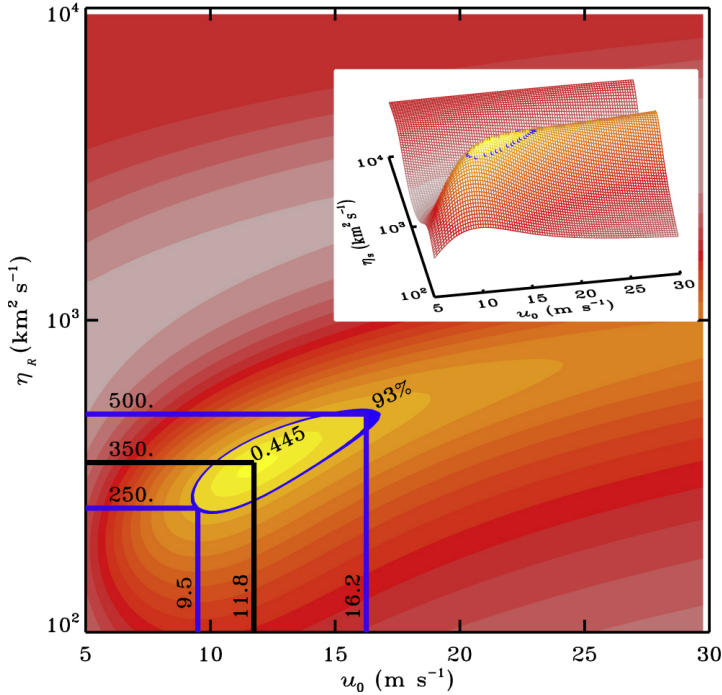
Fig. 4 Fitness function χ^{-2} as a function of meridional flow amplitude $u_0 = \max_\theta |u_\theta|$ (horizontal axis) and diffusivity $\eta_R \equiv \eta$ (vertical axis), from the optimization study of Lemerle et al (2015). Black lines show the optimum value and blue lines the limit of the acceptable region ($\chi^{-2} \geq 93\%\chi_{\max}^{-2}$). (© AAS. Reproduced with permission. Original article: http://dx.doi.org/10.1088/0004-637X/810/1/78)

When optimizing the model individually for different solar cycles, Whitbread et al (2017) found some cycle-to-cycle variation in the optimal speeds and diffusivities. This is understandable given the phenomenological nature of the model (to be discussed further in Section 5). Indeed, when simulating multiple cycles, Wang et al (2002) had previously varied the meridional flow speed from cycle to cycle so as to avoid unrealistic drift of the polar field over time. On the other hand, other authors have avoided this problem by varying instead the tilts of emerging active regions (Cameron et al, 2010, see also Section 4.3), or adding an addition decay term (to be discussed in Section 5). In reality it is likely that the effective mean-field meridional flow varies even over the course of a single Solar Cycle (see Section 6). Interestingly, Hung et al (2017) have shown – in the context of a flux-transport (interior) dynamo model – that a time-dependent meridional flow may be recovered from surface magnetic data through variational data-assimilation, and in future this approach could also be tried for SFT.
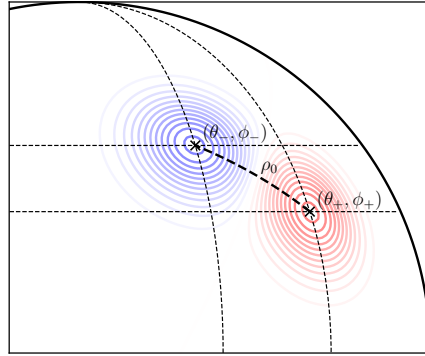
Fig. 5 Positive and negative contours of B_r for a BMR of the Van Ballegooijen et al (1998) form (14). The size is exaggerated ($\rho_0 = 25°$) compared to a real active region. This example follows Joy's Law in that the leading (rightmost) polarity is closest to the equator.

4 The Source Term

The magnetic flux patterns in the SFT model are determined in large part by the source term $S(\theta, \phi, t)$, which – in the classical mean-field model – represents the emergence of new macroscopic active regions on the solar surface. Since the classical SFT equation (1) is linear in B_r, the solution is a superposition of solutions for each individual active region, so it is insightful to consider the evolution of one of these regions in isolation. Since most SFT simulations follow the evolution for periods of years, it is usual to emerge each active region instantaneously in time, so that

$$S(\theta, \phi, t) = \sum_i B_r^{(i)}(\theta, \phi)\delta(t - t^{(i)}), \tag{13}$$

where $B_r^{(i)}(\theta, \phi)$ is the magnetic field of an individual active region emerging at $t = t^{(i)}$.

Traditionally, SFT models treat each active region as a bipolar magnetic region (BMR). Figure 5 shows the shape used by Van Ballegooijen et al (1998), with circular flux patches centred on the poles (θ_-, ϕ_-) and (θ_+, ϕ_+) and having the form

$$B_r(\theta, \phi) = B_0 \left\{ \exp\left[-\frac{2(1 - \cos \beta_+)}{(b\rho_0)^2}\right] - \exp\left[-\frac{2(1 - \cos \beta_-)}{(b\rho_0)^2}\right] \right\}, \tag{14}$$

where

$$\cos \beta_\pm = \cos \theta_\pm \cos \theta + \sin \theta_\pm \sin \theta \cos(\phi - \phi_\pm), \tag{15}$$

$$\cos \rho_0 = \cos \theta_+ \cos \theta_- + \sin \theta_+ \sin \theta_- \cos(\phi_+ - \phi_-). \tag{16}$$
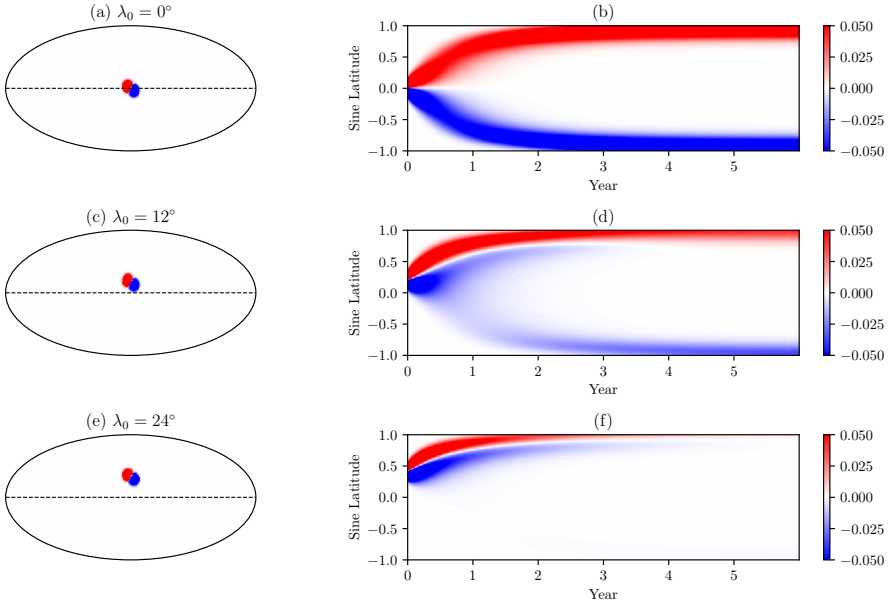
Fig. 6 Long term evolution of three identical BMRs emerged at different latitudes ($\lambda_0 = \pi/2 - \theta_0$) in the SFT model. Left column shows the initial BMRs and right column the time-latitude plot of $\langle B_r \rangle$ in each case.

Thus $\beta_\pm(\theta, \phi)$ denote the heliocentric angles from each pole, and ρ_0 the heliocentric angle between them. Van Ballegooijen et al (1998) took $b = 0.4$, and chose B_0 so that $\int_S |B_r| \sin \theta \, d\theta \, d\phi = 2\Phi_0$. For some purposes, one can approximate (14) with a pair of Dirac-delta sources,

$$B_r(\theta, \phi) = \frac{\Phi_0}{R_\odot^2 \sin \theta} \Big[\delta(\theta - \theta_+)\delta(\phi - \phi_+) - \delta(\theta - \theta_-)\delta(\phi - \phi_-) \Big]. \qquad (17)$$

Although the precise chosen shape for BMRs varies between implementations (see Yeates, 2020, for another variation), the key properties are the magnetic flux, Φ_0, and pole locations, (θ_-, ϕ_-) and (θ_+, ϕ_+). The latter may equivalently be specified by giving the coordinates of the BMR centre (θ_0, ϕ_0) along with the separation ρ_0 as in (16) and tilt angle γ_0, typically defined by

$$\tan \gamma_0 = \frac{\theta_+ - \theta_-}{\sin \theta_0 (\phi_+ - \phi_-)}. \qquad (18)$$

Together these BMR properties determine both the short-term and long-term evolution of the region.

After a new region emerges, much of its magnetic flux cancels by supergranular diffusion. Physically, this models the observed process of flux cancellation at the polarity inversion line (PIL) between the positive and negative polarities. This cancellation rate is enhanced as the region is sheared by differential

rotation and the PIL lengthened. On short-timescales (days) it is possible to approximate the solar surface as a Cartesian plane. Assuming a linear shear flow profile for the differential rotation, Lagrangian variables can be used to solve the Cartesian form of (1) for the exact evolution $B_r(\theta, \phi, t)$ of a tilted BMR (we will see an example in Section 4.1). On longer timescales, it is necessary to follow the evolution numerically.

It takes approximately 2 years for the non-axisymmetric component of B_r to cancel completely (Wang and Sheeley, Jr., 1991). Whether or not any axisymmetric B_r remains on a longer timescale depends on how much flux of one polarity escapes across the equator so that the two polarities are pushed to opposite poles by the meridional flow. An untilted region will send both polarities equally to each pole and so leave no asymptotic contribution at the end of the solar cycle. In a similar way, a (tilted) region that is nearer to the equator will produce a greater asymptotic contribution, because more flux escapes across the equator before being cancelled. This important effect is illustrated in Figure 6, where the same BMR is inserted at three different latitudes.

4.1 Dipole Amplification Factor of a BMR

A common way to measure the end-of-cycle contribution of an individual BMR is through its axial dipole strength

$$b_{1,0}(t) = \frac{3}{4\pi} \int_0^{2\pi} \int_0^{\pi} B_r \cos\theta \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi = \frac{3}{2} \int_0^{\pi} \langle B_r \rangle \cos\theta \sin\theta \, \mathrm{d}\theta. \quad (19)$$

By linearity of the classical SFT model, the total axial dipole strength will be the sum of the individual contributions from all of the active regions.

At the time of emergence, a BMR with the simple form (17) has

$$b_{1,0}(t_{\mathrm{em}}) = \frac{3\Phi_0}{4\pi R_\odot^2} \int_0^{\pi} \left[\delta(\theta - \theta_+) - \delta(\theta - \theta_-) \right] \cos\theta \, \mathrm{d}\theta \quad (20)$$

$$= \frac{3\Phi_0}{4\pi R_\odot^2} \left(\cos\theta_+ - \cos\theta_- \right) \quad (21)$$

$$= \frac{3\Phi_0}{2\pi R_\odot^2} \sin\left(\frac{\theta_- - \theta_+}{2} \right) \sin\left(\frac{\theta_+ + \theta_-}{2} \right) \quad (22)$$

$$\approx -\frac{3\Phi_0}{4\pi R_\odot^2} \rho_0 \sin\gamma_0 \sin\theta_0. \quad (23)$$

Here we have defined the central colatitude $\theta_0 = (\theta_+ + \theta_-)/2$ and recognized that for tilt angle γ_0 and heliocentric angle ρ_0 between the poles, their latitudinal separation is $(\theta_+ - \theta_-) = \rho_0 \sin\gamma_0$ (assuming $\theta_+ > \theta_-$). Thus, as noted by Wang and Sheeley, Jr. (1991), the axial dipole strength of a newly-emerged BMR depends on its flux, its latitudinal pole separation, and the cosine of its emergence latitude.
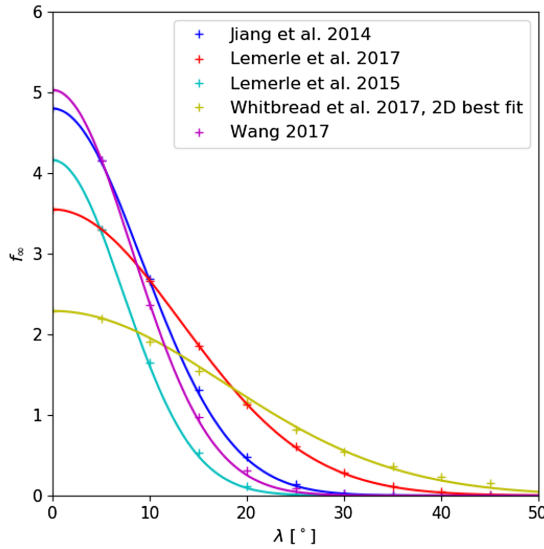
Fig. 7 Latitude dependence of the dipole amplification factor for BMRs in different published SFT models. The solid lines show Gaussian fits. Reproduced from Petrovay et al (2020).

Importantly, the axial dipole strength of a BMR can change under the ensuing SFT evolution: it will be amplified if the BMR emerged near the equator, or will decay if the BMR emerged far from the equator. It was first recognized by Jiang et al (2014a) that the "dipole amplification factor"

$$f_\infty = \lim_{t \to \infty} \frac{b_{1,0}(t)}{b_{1,0}(t_{\mathrm{em}})} \tag{24}$$

is well approximated by a Gaussian function of latitude, of the form

$$f_\infty(\lambda_0) = A \, \exp\left(-\frac{\lambda_0^2}{2\lambda_R^2}\right), \tag{25}$$

where it is convenient to work in terms of latitude λ rather than colatitude θ (the two are related by $\lambda = \pi/2 - \theta$). Figure 7 shows the functional form measured in several different numerical SFT models, where we note that both the amplitude A and width λ_R depend on the model. Once these parameters are known, equation (25) – coupled with the linearity of the SFT evolution equation (5) or (7) – allows the net axial dipole strength at the end of a solar cycle to be determined algebraically just by adding up the contributions of the individual BMRs, without the need to solve the evolution equation (Petrovay et al, 2020).

The interpretation of Figure 7 is that only BMRs that emerge with latitude $|\lambda| < \lambda_R$ will contribute to the global dipole moment at the end of the solar cycle. Petrovay et al (2020) call λ_R the "dynamo effectivity range", and give

the following simple physical derivation. To give a lasting contribution, a BMR must be close enough to the equator that some of its leading-polarity flux is able to cross the equator by diffusion, in opposition to the meridional flow. The timescale for advective separation at the equator is Δ_u^{-1}, where $\Delta_u = R_\odot^{-1} u'_\theta(\pi/2)$ is the equatorial divergence of $u_\theta(\theta)$. Equating this to the diffusion timescale $(\lambda R_\odot)^2/\eta$ from latitude λ to the equator suggests that

$$\lambda_R \approx \sqrt{\frac{\eta}{R_\odot^2 \Delta_u}} = \mathrm{Rm}_0^{-1/2}. \tag{26}$$

Note the reappearance of the magnetic Reynolds number from Section 3.2. Petrovay et al (2020) computed $f_\infty(\lambda_0)$ for numerical solutions with several different u_θ profiles, and in most cases found that the Gaussian width λ_R was indeed well approximated by $\mathrm{Rm}_0^{-1/2}$, the exception being a flow where u_θ peaks at a very low latitude compared to observations.

Petrovay et al (2020) went further and derived (25) analytically. The trick is to recognize that the final dipole moment – once the B_r distribution has become (near) axisymmetric – will be proportional to the remaining net magnetic flux in each hemisphere. (There will also be a coefficient depending on the latitudinal profile of the near-steady state as in (11).) Because it is determined purely by flux crossing the equator, the evolution of the net hemispheric flux can be quite well approximated by a Cartesian SFT model near the equator, which has the advantage of being analytically tractable. Thus Petrovay et al (2020) consider the "low-latitude limit" of (7),

$$\frac{\partial \langle B_r \rangle}{\partial t} + \frac{1}{R_\odot} \frac{\partial}{\partial \lambda} \left(u_\lambda \langle B_r \rangle \right) = \frac{\eta}{R_\odot^2} \frac{\partial^2 \langle B_r \rangle}{\partial \lambda^2}. \tag{27}$$

By choosing the linearised meridional flow $u_\lambda = R_\odot \Delta_u \lambda$, we can define the Lagrangian coordinate $\ell = \mathrm{e}^{-\Delta_u t} \lambda$ and new time variable $\tau = \left(1 - \mathrm{e}^{-2\Delta_u t}\right)/(2\mathrm{Rm}_0)$ to reduce Equation (27) to a standard diffusion equation

$$\frac{\mathrm{D}}{\mathrm{D}\tau} \left(\mathrm{e}^{\Delta_u t} \langle B_r \rangle \right) = \frac{\partial^2}{\partial \ell^2} \left(\mathrm{e}^{\Delta_u t} \langle B_r \rangle \right), \tag{28}$$

where $\mathrm{D}/\mathrm{D}t$ denotes the partial derivative with ℓ kept constant rather than λ. Equation (28) may be solved for a variety of initial conditions using standard techniques.

If the initial condition consists of a single (monopole) point source,

$$\langle B_r \rangle(\lambda, 0) = \frac{\Phi_0}{2\pi R_\odot^2} \delta(\lambda - \lambda_0), \tag{29}$$

then solving (28) gives

$$\mathrm{e}^{\Delta_u t}\langle B_r\rangle = \frac{\Phi_0}{2\pi R_\odot^2 \sqrt{4\pi\tau}} \exp\left(-\frac{(\ell-\lambda_0)^2}{4\tau}\right), \tag{30}$$

which for large t is approximately

$$\langle B_r\rangle(\lambda,t) \sim \frac{\Phi_0\sqrt{\mathrm{Rm}_0}\mathrm{e}^{-\Delta_u t}}{2\pi R_\odot^2\sqrt{2\pi}} \exp\left(\frac{-\mathrm{Rm}_0\left(\mathrm{e}^{-\Delta_u t}\lambda-\lambda_0\right)^2}{2}\right). \tag{31}$$

In the approximation (31), the flux difference between the hemispheres is

$$\Phi_\mathrm{N} - \Phi_\mathrm{S} = 2\pi R_\odot^2 \left(\int_0^\infty \langle B_r\rangle\,\mathrm{d}\lambda - \int_{-\infty}^0 \langle B_r\rangle\,\mathrm{d}\lambda\right) = \Phi_0\,\mathrm{erf}\left(\sqrt{\frac{\mathrm{Rm}_0}{2}}\lambda_0\right), \tag{32}$$

valid for either sign of λ_0.

For a BMR we must combine two point sources as in (17), each contributing half of the flux Φ_0, so

$$\Phi_\mathrm{N} - \Phi_\mathrm{S} = \frac{\Phi_0}{2}\left[\mathrm{erf}\left(\sqrt{\frac{\mathrm{Rm}_0}{2}}\lambda_+\right) - \mathrm{erf}\left(\sqrt{\frac{\mathrm{Rm}_0}{2}}\lambda_-\right)\right] \tag{33}$$

$$\approx \frac{\Phi_0\sqrt{\mathrm{Rm}_0}(\lambda_+-\lambda_-)}{\sqrt{2\pi}} \exp\left(-\frac{\mathrm{Rm}_0\lambda_0^2}{2}\right), \tag{34}$$

where we recognize the finite difference as an approximation of the derivative at $\lambda_0 = (\lambda_- + \lambda_+)/2$. We therefore expect that, to a good approximation, $b_{1,0}(t) \to a(\Phi_\mathrm{N} - \Phi_\mathrm{S})/R_\odot^2$ as $t \to \infty$, for some constant a that depends on the (normalized) shape of the steady B_r profile (thus only on u_θ and D). At the initial time, Equation (23) gives $b_{1,0}(0) = 3\Phi_0(\lambda_+ - \lambda_-)\cos\lambda_0/(4\pi R_\odot^2)$. Approximating $\cos\lambda_0 \approx 1$, the ratio is therefore

$$f_\infty \approx \frac{a\sqrt{8\pi\mathrm{Rm}_0}}{3} \exp\left(-\frac{\mathrm{Rm}_0\lambda_0^2}{2}\right). \tag{35}$$

Thus we recover (25) with $\lambda_R = \mathrm{Rm}_0^{-1/2}$ as claimed. Moreover, for a known asymptotic profile of $B_r(\theta)$, we can determine a and hence also predict the amplitude A.

4.2 Non-Bipolar Source Regions

Real solar active regions cannot always be represented as simple, symmetric BMRs. Even a region with two polarities will be effectively "multipolar" if the polarities are asymmetric in shape, and this will modify the evolution of $b_{1,0}$
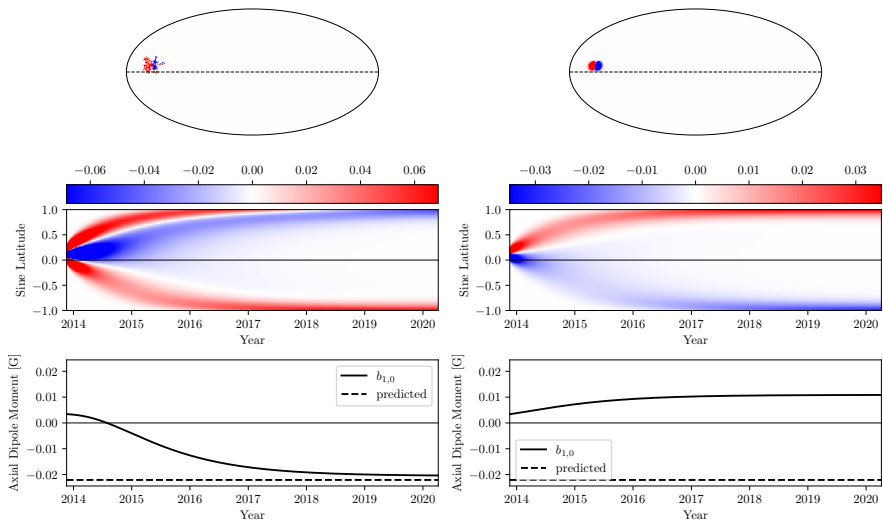
Fig. 8 Evolution of an active region with asymmetric bipolar shape, taken from the simulation in Figure 1. The left column shows the region with its original shape, with second row showing $\langle B_r \rangle$ and third row the axial dipole strength $b_{1,0}$. The right column shows the evolution of an "equivalent" symmetric BMR having the same initial flux and $b_{1,0}$. The dashed line shows the final $b_{1,0}$ predicted by equation (36) using the observed magnetogram.

compared to a symmetric BMR. This was investigated by Iijima et al (2019), who ran SFT simulations with Gaussian BMRs of the form (14), but where the leading polarity has a narrower width than the following polarity (controlled by the b parameter in (14)). When calibrated to the observed level of sunspot area asymmetry, their SFT simulation gave a more realistic evolution of both $b_{1,0}$ and the magnetic butterfly diagram, as compared to a reference simulation with equally-sized polarities. In particular, they noted that a wider following polarity leads to more following polarity flux crossing the equator, cancelling some of the trans-equatorial leading polarity flux and weakening the asymptotic contribution of the region. Similarly, Wang et al (2021) found for asymmetric BMRs with more diffuse following polarity, that f_∞ is systematically reduced (see their Figure 4). As an illustration, Figure 8 shows an example of the SFT evolution for an asymmetric region; in this case, the effect is sufficiently extreme to reverse the sign of $b_{1,0}$ altogether compared to a symmetric BMR.

Jiang et al (2019) considered the SFT evolution of a more complex "δ-type" flux distribution. They showed that $b_{1,0}$ changed sign during the SFT evolution, ending up with a completely different end-of-cycle contribution than would be expected for a BMR emerging at the same latitude with the same flux and same initial $b_{1,0}$. Wang et al (2021) showed further that the dipole amplification f_∞ is no longer a simple function of emergence latitude for such complex regions. However, the net effect of all of the real complex and asymmetric regions seems to be a reduction in the net end-of-cycle dipole, at least for
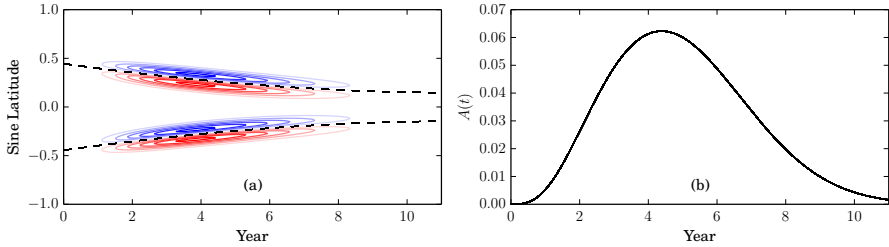
Fig. 9 The smooth source term $\langle S \rangle(\lambda, t)$ used by Petrovay and Talafha (2019). In (a), red/blue contours show $\langle S \rangle$, and dashed lines indicate $\pm \sin[\lambda_0(t)]$ from (39). Panel (b) shows the overall cycle shape $A(t)$ from (38) with $a = 0.00185$, $b = 4.058$, $c = 0.71$.

Cycle 24. Evidence for this comes from Yeates (2020), who compared an SFT simulation of that cycle where all active regions emerged with their observed flux distributions to a simulation where they were all approximated by symmetric BMRs with the same flux and initial $b_{1,0}$. The net $b_{1,0}$ at the end of the cycle was overestimated by 24% when the regions were modelled with BMRs.

For predicting the dipole contributions of more complex regions, Wang et al (2021) showed that (35) can be generalized to regions with non-bipolar shapes, by treating them as a superposition of point sources. In particular, for an active region with initial flux distribution $B_r(\theta, \phi, 0)$, combining the hemispheric flux differences (32) predicts that the axial dipole strength at the end of the cycle would be

$$\lim_{t \to \infty} b_{1,0}(t) \approx \frac{a}{R_\odot^2} \int_S B_r(\theta, \phi, 0) \operatorname{erf}\left[\sqrt{\frac{\mathrm{Rm}_0}{2}}\left(\frac{\pi}{2} - \theta\right)\right] \sin\theta \, \mathrm{d}\theta\mathrm{d}\phi, \qquad (36)$$

where a is the coefficient in the relation $b_{1,0} \approx a(\Phi_\mathrm{N} - \Phi_\mathrm{S})/R_\odot^2$. Wang et al (2021) verified this prediction against SFT simulations for 84 regions during Cycle 24. It gives an accurate prediction for the region in Figure 8.

4.3 Modelling the Source Term

It is not always viable to use observations of real individual active regions to construct the source term. This situation arises when working with historical data, when running SFT models into the future for forecasting purposes, or just in conceptual simulations studying the underlying physics. In such cases the source term $S(\theta, \phi, t)$ needs to be modelled, either as a smooth function (*e.g.*, Cameron and Schüssler, 2007; Petrovay and Talafha, 2019) or as random realizations of active regions drawn from a statistical distribution (*e.g.*, Schrijver, 2001; Mackay and Lockwood, 2002; Baumann et al, 2004; Jiang et al, 2018; Wang and Lean, 2021).

The smooth function approach has primarily been used in the 1D SFT model, (7), for example by Petrovay and Talafha (2019), who use a pair of flux

rings in each hemisphere, shown in Figure 9(a) and given by

$$\langle S\rangle(\lambda,t) = (-1)^n A(t) \left\{ \exp\left(-\frac{[\lambda-\lambda_+(t)]^2}{2\delta_\lambda^2}\right) - \exp\left(-\frac{[\lambda-\lambda_-(t)]^2}{2\delta_\lambda^2}\right) \right.$$
$$\left. + \exp\left(-\frac{[\lambda+\lambda_+(t)]^2}{2\delta_\lambda^2}\right) - \exp\left(-\frac{[\lambda+\lambda_-(t)]^2}{2\delta_\lambda^2}\right) \right\}. \tag{37}$$

This model incorporates a number of observed solar cycle features:

(i) All polarities alternate according to the solar cycle number, n.
(ii) The cycle has an asymmetrical shape in time, shown in Figure 9(b) and given by the Hathaway et al (1994) observed fit

$$A(t) = a(t-t_{\min}) \left(\exp\left[\frac{(t-t_{\min})^2}{b^2}\right] - c \right)^{-1}, \tag{38}$$

where $t_{\min}$ is the start of the cycle.
(iii) The centres $\pm\lambda_0$ of each pair of flux rings, *i.e.* $\lambda_0 = (\lambda_+ + \lambda_-)/2$ shown by dashed lines in Figure 9(a), migrate equatorward at the rate

$$\lambda_0(t) = 26.4 - 34.2\left(\frac{t}{T}\right) + 16.1\left(\frac{t}{T}\right)^2 \quad [°] \tag{39}$$

fitted empirically by Jiang et al (2011), where T is the cycle length (11 years).
(iv) The separation $\Delta_\lambda = \lambda_- - \lambda_+$ decreases as λ_0 approaches the equator, according to

$$\Delta_\lambda(t) = 0.5\frac{\sin\lambda_0(t)}{\sin 20°} \quad [°]. \tag{40}$$

This models the longitude-averaged effect of the well-established Joy's Law (van Driel-Gesztelyi and Green, 2015), whereby BMRs emerging at lower latitude have (on average) smaller tilt angle $|\gamma_0|$, defined in (18).

The statistical BMR approach is similar, except the functions above are treated as overall distributions from which discrete BMRs are chosen at random. For the longest historical simulations, which date back to 1700 (Jiang et al, 2018; Wang et al, 2021), the only observational input is the sunspot number time series – equivalent to emergence rate, $A(t)$. For 20th Century simulations, data on the areas and locations of individual sunspot groups can be used (*e.g.*, Cameron et al, 2010). However, even here the magnetic flux and tilt angle (equivalently axial dipole strength) must be chosen at random as they are not available observationally before the onset of routine magnetograms in the 1970s.

The tilt angle is problematic as Joy's Law, as modelled in (40), holds only for the mean, and there is known to be very significant scatter (*e.g.*, Wang and Sheeley, Jr., 1989; Yeates, 2020). Recent studies have shown that individual

"rogue" active regions – defined as those with dipole moments significantly different from Joy's Law expectation at their latitude – can have a significant effect on the overall polar field at the end of the cycle (Jiang et al, 2015; Nagy et al, 2017). In light of (35), such rogue regions must typically emerge near to the equator, although their relative contribution depends on Rm_0 and would be reduced if Rm_0 were large. Nevertheless, simulations based on statistical source terms without individual dipole moment data should be treated with caution, particularly for prediction.

The widely-accepted $\alpha\Omega$ paradigm for the solar dynamo suggests a "self-consistent" way to build a fully synthetic SFT model: set the amount flux emerging through the source term in cycle n proportional to the axial dipole strength at the end of cycle $n - 1$. Talafha et al (2022) modified the one-dimensional model of Petrovay and Talafha (2019) to use such an approach. They used this model to systematically study the impact of two possible non-linearities in the source term: tilt quenching (where BMRs are less tilted in strong cycles) and latitude quenching (where BMRs emerge at higher latitudes in strong cycles). SFT simulations show that both effects act to reduce the axial dipole produced in strong cycles (Cameron et al, 2010; Jiang, 2020). They are both therefore possible saturation mechanisms to explain why the solar dynamo doesn't exhibit runaway exponential growth. Talafha et al (2022) showed that the relative impact of tilt versus latitude quenching on the end-of-cycle axial dipole depends primarily on the dynamo effectivity range λ_R in equation (26). In particular, for small λ_R, latitude quenching reduces the end-of-cycle dipole more than tilt quenching, and *vice versa* for large λ_R. However, the amount of tilt and/or latitude quenching present on the real Sun remains under debate.

5 Physical Justification

As introduced by Leighton (1964), the SFT model is purely phenomenological. But can equation (1) be derived from known physical laws? The relevant law governing the evolution of the large-scale magnetic field is the mean-field MHD (magnetohydrodynamic) induction equation,

$$\frac{\partial B_r}{\partial t} = \mathbf{e}_r \cdot \nabla \times \left(\mathbf{u} \times \mathbf{B} - \eta \nabla \times \mathbf{B} \right), \tag{41}$$

where $\mathbf{u}$ is the plasma velocity and – anticipating the form of (1) – we have made a simple approximation for the turbulent electromotive force of the form $-\eta \nabla \times \mathbf{B}$ (cf. McCloughan and Durrant, 2002). Thus η represents turbulent diffusivity, not ohmic resistivity (which is negligible in the highly conducting photosphere). This assumption of a turbulent diffusivity is discussed further in Section 5.3 below.

Consider the first term in (41). Decomposing $\mathbf{u} = \mathbf{u}_h + u_r \mathbf{e}_r$, where $\mathbf{e}_r \cdot \mathbf{u}_h = 0$, and similarly $\mathbf{B} = \mathbf{B}_h + B_r \mathbf{e}_r$, we can write

$$\mathbf{e}_r \cdot \nabla \times (\mathbf{u} \times \mathbf{B}) = \nabla \cdot (u_r \mathbf{B}_h) - \nabla \cdot (\mathbf{u}_h B_r). \tag{42}$$

The last term is precisely the advection term in the SFT equation (1), while the term $\nabla \cdot (u_r \mathbf{B}_h)$ represents flux emergence, so corresponds to the source term S in (1). Thus the SFT model is incorporating the correct advection terms.

Now consider the diffusion term in (41). For simplicity, we will assume that $\eta = \eta(r)$ only, in which case

$$-\mathbf{e}_r \cdot \nabla \times (\eta \nabla \times \mathbf{B}) = \eta \nabla_h^2 B_r + R_\eta. \tag{43}$$

This has the diffusion term from (1) plus an additional remainder term

$$R_\eta = -\frac{\eta}{R_\odot} \nabla_h \cdot \mathbf{B} - \frac{\eta}{R_\odot \sin\theta} \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial B_\theta}{\partial r}\right) - \frac{\eta}{R_\odot \sin\theta} \frac{\partial}{\partial\phi}\left(\frac{\partial B_\phi}{\partial r}\right). \tag{44}$$

Using $\nabla \cdot \mathbf{B} = 0$, this may be rewritten entirely in terms of B_r, simplifying to

$$R_\eta = \frac{\eta}{R_\odot^2} \frac{\partial^2}{\partial r^2}(r^2 B_r). \tag{45}$$

Thus in mean-field MHD there is an additional term representing the radial diffusion of magnetic flux that is missing from the original SFT equation (1). Physically, this incorporates the fact that the surface magnetic field is connected to the interior; for example, the decay of active regions can be slowed if they remain connected to deeper layers of the convection zone where the diffusivity is lower (Wilson et al, 1990; Whitbread et al, 2019).

Self-consistent computation of the radial diffusion term R_η would require simulation of the three-dimensional magnetic field in the solar convection zone. However, two approaches have been used to parametrize (45) in SFT models without the need for three-dimensional simulations, and these will be considered next.

5.1 Exponential Decay Term

The most common parametrization for the radial diffusion term (45) is to assume that $R_\eta \approx -B_r/\tau$, so that (1) becomes

$$\frac{\partial B_r}{\partial t} + \nabla_h \cdot (\mathbf{u}_h B_r) = \eta \nabla_h^2 B_r - \frac{B_r}{\tau} + S. \tag{46}$$

Multiplying by $e^{t/\tau}$ shows that

$$\frac{\partial}{\partial t}(e^{t/\tau} B_r) + \nabla_h \cdot (\mathbf{u}_h e^{t/\tau} B_r) = \eta \nabla_h^2 (e^{t/\tau} B_r) + e^{t/\tau} S. \tag{47}$$
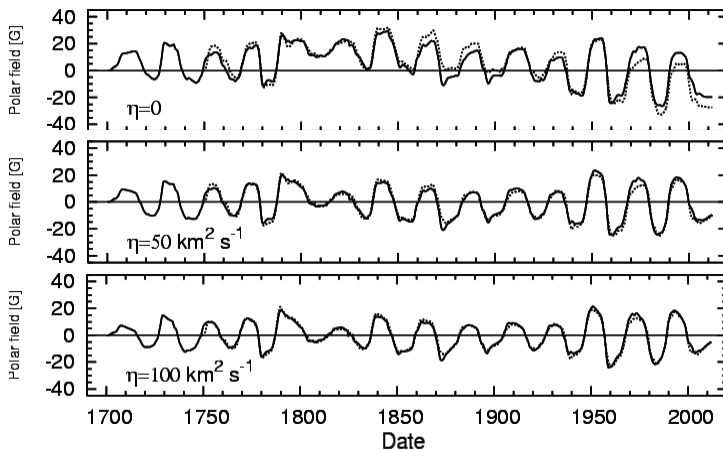
Fig. 10 Application of the radial diffusion term to reduce spurious cycle-to-cycle memory in the SFT model, from Baumann et al (2006). The top row shows the north polar field (above 75° latitude) in a simulation with no radial diffusion, while the middle and bottom rows show the same simulation with $\eta_0 = 50\,\mathrm{km^2 s^{-1}}$ and $100\,\mathrm{km^2 s^{-1}}$ according to the prescription in Section 5.2. The simulation uses random emerging BMRs proportional to the observed sunspot numbers. The dashed line shows a simulation started in 1750, illustrating how the memory of the initial conditions persists. (Credit: Baumann, I., Schmitt, D. and Schüssler, M., A&A, 446, 307-314, 2006, reproduced with permission © ESO.)

Thus if B_r^∞ denotes the solution to the original equation (1), corresponding to $\tau \to \infty$, then the solution with finite τ but all other parameters the same is $B_r = \mathrm{e}^{-t/\tau} B_r^\infty$. In other words, the solution decays exponentially at uniform rate τ^{-1}. For example, the dipole amplification factor (35) for a BMR would become

$$f_\infty \approx \frac{\sqrt{8\pi \mathrm{Rm}_0}}{3} \exp\left(-\frac{\mathrm{Rm}_0 \lambda_0^2}{2}\right) \exp\left(-\frac{t}{\tau}\right), \qquad (48)$$

reflecting continuing decay of the magnetic field due to the new term.

The first application of such a decay term was by Schrijver et al (2002), who motivated it not by consideration of radial diffusion but purely as a necessary addition to reduce the "memory" of the polar field (equivalently $b_{1,0}$) over multiple solar cycles. Without it, the varying amount of polar field production caused by the differing sunspot numbers in different cycles led to an unrealistic drift in the polar field over time, rather than the regular reversals that are observed. This drift is illustrated (for another SFT model) in the top panel of Figure 10.

The optimization studies discussed in Section 3 have also looked for the optimum τ in shorter simulations where long-term memory is not an issue. With their simplified source term, Petrovay and Talafha (2019) found that a decay term (with τ in the range 5–10 yr) was essential, otherwise $b_{1,0}$ reversed too late for all of the flow profiles and parameters tried. And in simulations of Cycle 23 driven by idealised BMRs, Whitbread et al (2017) found that a decay term with $\tau < 5\,\mathrm{yr}$ helped to reduce unrealistically high values of $b_{1,0}$. However, they found that emerging active regions with observed shapes

reduced $b_{1,0}$ in itself (as did Yeates, 2020, for Cycle 24), and the optimization did not strongly select for a particular τ. Moreover, the fit of the optimum model did not improve significantly when the decay term was included in the model compared to when it was not. Lemerle et al (2015) also found that τ was not strongly constrained by the optimization process, with acceptable solutions found for suitable parameter combinations with τ in the range from 7–32 yr. In summary, the presence of a decay term as required by Schrijver et al (2002) does not seem to be ruled out by observations.

It should be noted that, in principle, an additional decay term is not the only way to reduce the cycle-to-cycle memory of $b_{1,0}$ in the model. Alternatives that have been adopted include imposed cycle-to-cycle variations in either the meridional flow speed (Wang et al, 2002) or the tilt angles of emerging BMRs (Cameron et al, 2010). It is difficult to choose definitively between these options with only about four solar cycles of full magnetogram observations.

A final remark is that the simplified decay term B_r/τ is still compatible with Faraday's Law, in the sense that it can be written as the curl of an electric field. In particular, we would need $\mathbf{e}_r \cdot \nabla \times \mathbf{E} = B_r/\tau$. For example, writing $\mathbf{E} = -\nabla \times (\Psi \mathbf{e}_r)$, we could determine Ψ and hence $\mathbf{E}$ by solving the Poisson equation

$$\nabla_h^2 \Psi = \frac{B_r}{\tau}, \tag{49}$$

which has a unique solution on the sphere since $\int_S B_r \, dS = 0$. Of course, this does not mean that this approximation is a good representation of the real radial diffusion term (45); for example, this particular $\mathbf{E}$ will be not be localized to the active region itself.

5.2 Diffusive Interior Model

An improved parametrization for (45) was suggested by Baumann et al (2006). They observed that if one assumes a purely diffusive evolution with uniform diffusivity $\eta = \eta_0$ throughout the convection zone, then the term R_η may be approximated using only B_r on the solar surface.

Specifically, Baumann et al (2006) consider a purely poloidal field $\mathbf{B} = \nabla \times \nabla \times (\mathbf{r}P)$ inside the convection zone $R_b < r < R_\odot$, with boundary conditions $B_r(R_b, \theta, \phi) = 0$ and $B_\theta(R_\odot, \theta, \phi) = B_\phi(R_\odot, \theta, \phi) = 0$. Under a purely diffusive decay

$$\frac{\partial \mathbf{B}}{\partial t} = -\eta_0 \nabla \times (\nabla \times \mathbf{B}) \tag{50}$$

with η_0 constant, and a suitable choice of P (which choice doesn't affect $\mathbf{B}$), this reduces to the scalar problem

$$\frac{\partial P}{\partial t} = \eta_0 \nabla^2 P, \quad \frac{\partial}{\partial r}(rP)\Big|_{r=R_\odot}, \quad P\Big|_{r=R_b} = 0. \tag{51}$$

The solution, omitting the monopole term, may be written as an expansion

$$P(r,\theta,\phi,t) = \sum_{n=0}^{\infty}\sum_{l=1}^{\infty}\sum_{m=-l}^{l}\left[a_{l,n}j_l(k_{l,n}r) + c_{l,n}y_l(k_{l,n}r)\right]Y_l^m(\theta,\phi)e^{-\eta_0 k_{l,n}^2 t},$$

(52)

where Y_l^m are spherical harmonics and j_l, y_l are spherical Bessel functions of the first and second kinds. Linearity of (51) allows us to set $a_{l,n} = 1$ without loss of generality, so the inner boundary condition fixes the other coefficient

$$c_{l,n} = -\frac{j_l(k_{l,n}R_{\rm b})}{y_l(k_{l,n}R_{\rm b})}.$$

(53)

The upper boundary condition then gives

$$l\left[j_l(k_{l,n}R_{\odot})y_l(k_{l,n}R_{\rm b}) - y_l(k_{l,n}R_{\odot})j_l(k_{l,n}R_{\rm b})\right] =$$
$$k_{l,n}R_{\odot}\left[j_{l-1}(k_{l,n}R_{\odot})y_l(k_{l,n}R_{\rm b}) - y_{l-1}(k_{l,n}R_{\odot})j_l(k_{l,n}R_{\rm b})\right].$$

(54)

This equation must be solved numerically for each l and n to determine the eigenvalues $k_{l,n}$, which give the decay times $\tau_{l,n} = (\eta_0 k_{l,n}^2)^{-1}$ for each component, where l is the spherical harmonic degree and n is the radial mode number. Since the SFT model does not give the subsurface radial structure, Baumann et al (2006) propose to keep only the modes with $n = 0$, which are the slowest decaying modes for each l. They modify the SFT equation (1) to

$$\frac{\partial B_r}{\partial t} + \nabla_h \cdot (\mathbf{u}_h B_r) = \eta \nabla_h^2 B_r - \sum_{l=1}^{\infty}\sum_{m=-l}^{l}\frac{b_{l,m}(t)}{\tau_{l,0}}Y_l^m(\theta,\phi) + S,$$

(55)

where $b_{l,m}(t)$ are the spherical harmonic coefficients in the expansion of B_r,

$$B_r(\theta,\phi,t) = \sum_{l=1}^{\infty}\sum_{m=-l}^{l} b_{l,m}(t)Y_l^m(\theta,\phi).$$

(56)

The interior diffusivity η_0 that determines $\tau_{l,0}$ is taken to be different from the coefficent η of the classical diffusion term.

Note that, since radial modes with $n > 0$ are neglected, the effect on $b_{1,0}$ is identical to the simple exponential decay term, with $\tau = \tau_{1,0} = (\eta_0 k_{1,0}^2)^{-1}$. Accordingly, Baumann et al (2006) showed that their alternative form of the decay term can also reduce the spurious long-term memory of the SFT model, as illustrated in the middle and bottom rows of Figure 10. They found that diffusivity values in the range $\eta_0 = 50-100\,\mathrm{km}^2\mathrm{s}^{-1}$ gave polar field evolutions consistent with recent observations. For $R_{\rm b} = 0.7R_{\odot}$, and since $k_{1,0} \approx 5.46$, this corresponds to decay times for $b_{1,0}$ in the range $\tau_{1,0} \approx 5-10\,\mathrm{yr}$. In their model driven by idealized BMRs, Whitbread et al (2017) found an optimum

$\eta_0 = 190\,\mathrm{km^2 s^{-1}}$, giving a decay time $\tau_{1,0} = 2.7\,\mathrm{yr}$, in agreement with the τ found by optimizing the simple exponential decay term. Virtanen et al (2017) also adopted the Baumann et al (2006) model, but in a simulation where active regions had observed shapes; they found a value $\eta_0 = 100\,\mathrm{km^2 s^{-1}}$ to give reasonable results.

It is worth remarking that these implementations of (55) have used different diffusivities for η (the classical horizontal diffusion) and η_0 (which determines $\tau_{l,0}$). Moreover, the extra term in (55) includes both radial and horizontal diffusion due to the interior diffusivity η_0. If one evaluates the radial diffusion term (45) for a single mode of the interior solution (52), one obtains

$$\frac{\eta_0}{R_\odot^2}\frac{\partial^2}{\partial r^2}\left(r^2 B_r\right) = -\eta_0\left(k_{l,n}^2 - \frac{l(l+1)}{R_\odot^2}\right)B_r, \qquad (57)$$

giving a decay time $\tau'_{l,n} = \eta_0^{-1}[k_{l,n}^2 - l(l+1)/R_\odot^2]^{-1}$ for radial diffusion alone. However, for small l the difference from $\tau_{l,n}$ is negligible.

5.3 Other Turbulent Transport Effects

If we drop the simple assumption of a turbulent diffusion in the mean-field induction equation (41), then there are a wealth of possible transport effects that could be explored in SFT models. One such effect – expected to be present from numerical convection simulations – is turbulent pumping (Petrovay, 1994), which adds $-\boldsymbol{\gamma}\times\mathbf{B}$ to the turbulent electromotive force (mathematically equivalent to $\mathbf{u}$). Downward pumping ($\gamma_r < 0$) in a region near the surface could reduce the aforementioned diffusive link of active regions to deeper layers (Cameron et al, 2012; Karak and Cameron, 2016). This is because it will tend to make the magnetic field lines radial, and if $B_\theta, B_\phi \approx 0$ in some region near the surface then it follows from (44) that $R_\eta = 0$, so no additional radial diffusion term should be included in the SFT model. Latitudinal pumping ($\gamma_\theta \neq 0$) is also found to be very strong in convection simulations. However, this relies on a significant influence of rotation on the turbulence, which is weaker nearer the surface than in deeper layers.

6 Beyond the Classical Model

Several have sought to improve on the classical SFT model described in the previous sections. We therefore conclude this review by outlining some of these developments.

6.1 Improved Small-Scale Flows

The approximation of small-scale flows by a uniform supergranular diffusivity, D, is perhaps the greatest simplification in the classical model. Three main approaches for improving the fidelity of the small-scale flow model have been applied.
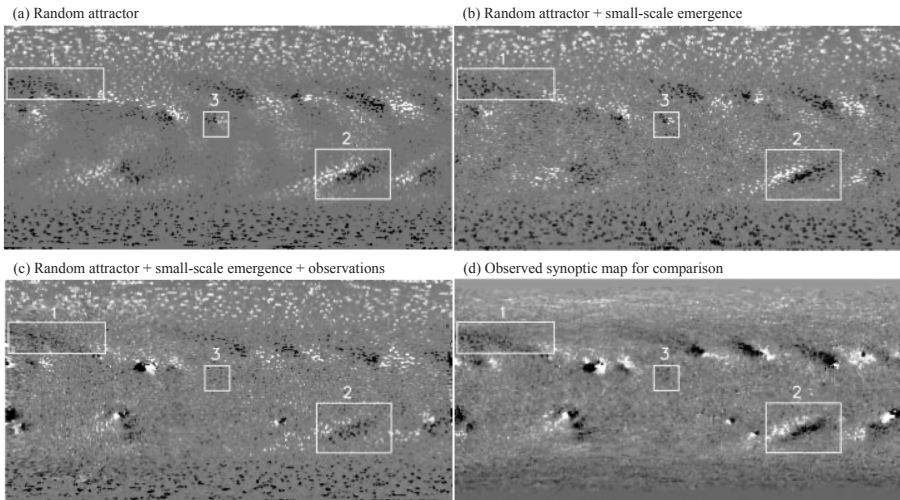
Fig. 11 Illustration of the "random attractor" model for flux dispersal, taken from several figures of Worden and Harvey (2000). Panels (a)-(c) show simulated maps after evolving for 27 days, all starting from a synoptic map for CR1928 but with successively more model components included. (Differential rotation and meridional flow were included in all three cases.) Panel (d) shows the "ground truth": an observed synoptic map for CR1929. In all cases B_r is shown in greyscale (white positive, black negative). (Reproduced with permission from Springer Nature. Original article: https://doi.org/10.1023/A:1005272502885)

Computationally cheapest is the method of Worden and Harvey (2000), whose primary aim was to improve the unobserved or poorly observed regions of synoptic maps. For this application, the classical diffusion model is not ideal because it does not reproduce the "clumping" of magnetic flux on supergranular network boundaries that is clearly evident in observed portions of the map. To better reproduce this, Worden and Harvey (2000) replaced the diffusion with a "random attractor" term added to each pixel in the map (without increasing the resolution compared to the classical SFT model). This is shown in Figure 11(a). They also added a random emergence term to each pixel to sustain the small-scale background field. This background field was found not to affect the diffusion of large-scale flux patterns, but it gives a more accurate net flux in quiet regions (Figure 11b). The technique was successful in improving the appearance of simulated maps, and continues to be used in the Air Force Data-Assimilative Photospheric flux Transport model (ADAPT; Arge et al, 2010; Hickmann et al, 2015).

A second approach is to dispense completely with parametrization of the small-scale flows, and model them directly through the advection term. This requires higher spatial and temporal resolution so as to resolve individual convective cells on the computational grid. Nevertheless, it has been applied successfully in the Advective Flux Transport (AFT) model (Upton and Hathaway, 2014, 2018). In this model, the small-scale flows are randomly imposed,

based on a vector spherical harmonic decomposition of the form

$$u_\theta(\theta, \phi) = \sum_{l=1}^{l_{\max}} \sum_{m=0}^{l} \left(S_l^m \frac{\partial Y_l^m(\theta, \phi)}{\partial \theta} + T_l^m \frac{1}{\sin \theta} \frac{\partial Y_l^m(\theta, \phi)}{\partial \phi} \right), \qquad (58)$$

$$u_\phi(\theta, \phi) = \sum_{l=1}^{l_{\max}} \sum_{m=0}^{l} \left(S_l^m \frac{1}{\sin \theta} \frac{\partial Y_l^m(\theta, \phi)}{\partial \phi} - T_l^m \frac{\partial Y_l^m(\theta, \phi)}{\partial \theta} \right), \qquad (59)$$

where the complex amplitudes S_l^m and T_l^m determine the curl-free and divergence-free components of $\mathbf{u}_h$ and are chosen to match the spectrum to observations. Hathaway et al (2000) found that observed Doppler flows could be well matched by a two-component spectrum, comprising a supergranular component centred on $l = 110$ and a granular component centered on $l = 4000$.

The third approach is to dispense with a computational grid altogether and model the magnetic flux by a discrete ensemble of individual flux "concentrations". This was implemented by Schrijver (2001) whose main aim was to simulate cool stars other than the Sun, and who therefore wanted to include the mixed-polarity network of small-scale magnetic flux because of its contribution to chromospheric emission. The discrete model of Schrijver (2001) includes (i) emergence of both active regions and ephemeral regions as BMRs, (ii) a large-scale random walk dispersal as well as differential rotation and meridional flow, (iii) a model for fragmentation and coalescence of flux concentrations, and (iv) cancellation of flux between opposite polarity fragments. The model has been successfully applied over all latitudes (Schrijver and Title, 2001) and over a full 11-year cycle (Schrijver and Liu, 2008). A similar model in Cartesian geometry was applied by Martin-Belda and Cameron (2016) to study the dispersion of a single active region.

One notable new feature that all three of these models have in common is nonlinearity: the rate of magnetic flux dispersal is chosen to depend on the local magnetic field strength, $|B_r|$. In particular, dispersal is suppressed in strong-field regions, compared to the classical diffusion model. This better represents real active regions which suppress shedding of the magnetic flux by supergranulation (Schrijver, 1989). The effect is particularly important for more active stars (Schrijver, 2001) but is still clearly observed on the Sun.

6.2 Fluctuating Large-Scale Flows

The classical model neglects fluctuations in the meridional flow and differential rotation, keeping them steady for periods of a solar cycle or longer. However, observations do suggest variations over the course of the cycle, particularly in the meridional flow. For example, Hathaway and Rightmire (2010) estimated the flow from cross-correlating latitudinal strips in magnetograms over Solar Cycle 23, and found that the dominant Legendre component, $P_2^1 \sim \sin(2\theta)$, reduced in amplitude from $11.5 - 13\,\mathrm{ms}^{-1}$ at cycle minimum to only $8.5\,\mathrm{ms}^{-1}$ at cycle maximum.

A plausible cause of meridional flow variations is the observed inflow toward active regions determined by helioseismology (Gizon et al, 2001). In SFT simulations, Jiang et al (2010) showed that an axisymmetric meridional inflow toward the activity belts leads to a significant decrease of the polar field, suggesting that such meridional flow variations could be a significant ingredient in the SFT model. And Cameron et al (2010) pointed out that the variations in P_2^1 found by Hathaway and Rightmire (2010) could be explained by this inflow, without the need for an overall modulation of meridional flow speed.

Other studies have accounted for the observed dependence of inflow speed on the active region magnetic flux, through applying a nonlinear velocity that depends on $|B_r|$. De Rosa and Schrijver (2006) added a velocity of the form

$$\delta\mathbf{u}(\theta, \phi, t) = \alpha\nabla|\overline{B_r}|^\beta \tag{60}$$

to the discrete SFT model – where $\overline{B_r}$ denotes a Gaussian smoothing of the original B_r with width $15°$ – but found that the observed flow speeds ($50\,\mathrm{m\,s^{-1}}$) prevented altogether the dispersal of active regions. However, Martin-Belda and Cameron (2016) did not find this problem and proposed that the original calculations of De Rosa and Schrijver (2006) were underestimating the flux dispersal because they continued to apply the nonlinear damping of dispersal within the active region, while the inflows alone could themselves account for the damping effect. Cameron and Schüssler (2012) proposed an axisymmetric parametrization

$$\delta u_\theta(\theta, t) = c_0 \int_0^\pi \frac{\sin(\theta')}{\sin(30°)} \frac{\mathrm{d}\langle|B_r|\rangle}{\mathrm{d}\theta'} e^{-(\theta-\theta')^2/\sigma}\,\mathrm{d}\theta', \tag{61}$$

which corresponds to a Gaussian smoothing of the derivative in latitude (with σ chosen to give width $20°$). The $\sin(\theta')$ factor suppresses unrealistically strong fluctuations at high latitudes, and an amplitude $c_0 = 9.2\,\mathrm{m\,s^{-1}G^{-1}}$ gives comparable inflow speeds to Gizon et al (2001). Again, the presence of inflows reduces the axial dipole at the end of the solar cycle, by about 30% in a moderate cycle (Martin-Belda and Cameron, 2017), with about a 9% variation between cycles suggesting that this nonlinearity could conceivably help to saturate the Babcock-Leighton dynamo. Nagy et al (2020) coupled an SFT model with flux-dependent inflows to such a dynamo model. They confirmed that inflows do indeed tend to have a stabilizing effect on cycle amplitudes, although they also greatly increase the probability of the dynamo entering a grand minimum of reduced activity – a nonlinear effect which is not apparent from SFT alone. On the other hand, Yeates (2014) found that the inflows in a BMR-driven SFT model for Cycle 23 gave poorer matches to the observed butterfly diagram and dipole reversal time.

A more pragmatic approach is to impose the observed flow variations directly, as in the AFT model (Upton and Hathaway, 2014), where the best-fit Legendre coefficients are extracted from 27-day averaged velocity fields

derived from magnetogram cross-correlation. These then determine $\mathbf{u}_h(\theta, t)$ in the model, allowing variations in both meridional flow and differential rotation.

6.3 Observational Data Assimilation

In applications where the aim is to recreate as accurately as possible the real Sun at an observed time, it makes sense to construct magnetic maps that combine SFT model results with real observations. The role of the SFT model is then to fill in unobserved (or poorly observed) parts of the solar surface, such as high latitudes or the far side of the Sun. This approach is central to the model of Worden and Harvey (2000), as illustrated in Figure 11(c) which shows the result of combining daily magnetogram observations with the simulation. The observations are weighted more highly near disk-centre and also eastward of Central Meridian (where the time since previous observation is greatest). Similar assimilation of observed magnetograms has been applied in the discrete SFT model (Schrijver and DeRosa, 2003) and in the AFT model of Upton and Hathaway (2014).

A more sophisticated approach to data assimilation has been implemented in the ADAPT model, which includes several different sequential data-assimilation methods such as ensemble Kalman filtering (Hickmann et al, 2015). The concept is to perform an ensemble of model runs. Each is adjusted at intervals using the observed magnetogram data, with observations being given greater weight in areas where the model runs disagree with one another.

Unfortunately, difficulties arise in driving time-dependent coronal magnetic field simulations from SFT models with data assimilation. In such simulations, the required photospheric boundary condition is the tangential electric field $\mathbf{E}_h$, not simply B_r. In the classical SFT model, the natural electric field would be

$$\mathbf{E}_h = -\mathbf{u} \times \mathbf{B} + \eta \nabla \times \mathbf{B} + \mathbf{E}_S, \tag{62}$$

where $\mathbf{E}_S$ accounts for the source term (*i.e.*, $-\mathbf{e}_r \cdot \nabla \times \mathbf{E}_S = S$). When S comprises individual active regions that have no net magnetic flux, a well-behaved electric field can be determined (*e.g.*, Yeates and Bhowmik, 2022). But if the magnetic flux is unbalanced over a larger region then it is impossible to find a localized $\mathbf{E}_S$ as would be expected from Ohm's Law (Yeates, 2017). This can be a problem when observed magnetograms are incorporated directly, particularly when active regions straddle the edge of the assimilation region so that only one polarity is included. If the flux imbalance is corrected by spreading it over the full Sun, the resulting spurious electric fields lead to generation of significant spurious electric currents in time-dependent coronal simulations (Weinzierl et al, 2016).

In practice the simplest way to ensure flux balance is to rephrase the right-hand side of equation (5) as $-\mathbf{e}_r \cdot \nabla \times \mathbf{E}_h$, then apply a "constrained transport" discretization with a staggered mesh (Yee, 1966). Here E_θ and E_ϕ are defined at cell edges, and B_r at cell centres. Such a numerical scheme is used, for example, by Yeates (2014). When assimilating magnetograms into the SFT model in this framework, one would estimate $\mathbf{E}_h$ from the observed front-side

evolution. In the case of a flux imbalance, this would automatically create a balancing polarity just outside the observed region, minimizing disruption to the global topology of the coronal magnetic field. However, it remains the case that systematic errors in observed magnetograms, especially centre-to-limb variations of the errors, are not well understood. A better understanding of these errors will require forward modelling with radiative MHD and Stokes polarimetric inversions.

Supplementary information.   None.

Declarations

Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

Arge CN, Henney CJ, Koller J, et al (2010) Air Force Data Assimilative Photospheric Flux Transport (ADAPT) Model. In: Maksimovic M, Issautier K, Meyer-Vernet N, et al (eds) Twelfth International Solar Wind Conference, pp 343–346, https://doi.org/10.1063/1.3395870

Baumann I (2005) Magnetic flux transport on the sun. PhD thesis, Göttingen, URL https://www.sidc.be/users/evarob/Literature/PhDs/Baumann_magnetic%20flux%20transport%20on%20the%20sun.pdf

Baumann I, Schmitt D, Schüssler M, et al (2004) Evolution of the large-scale magnetic field on the solar surface: A parameter study. Astron Astrophys 426:1075–1091

Baumann I, Schmitt D, Schüssler M (2006) A necessary extension of the surface flux transport model. Astron Astrophys 446:307–314

Bhowmik P, Nandy D (2018) Prediction of the strength and timing of sunspot cycle 25 reveal decadal-scale space environmental conditions. Nature Comms 9:5209

Cameron RH, Schüssler M (2007) Solar cycle prediction using precursors and flux transport models. Astrophys J 659:801–811

Cameron RH, Schüssler M (2012) Are the strengths of solar cycles determined by converging flows towards the activity belts? Astron Astrophys 548:A57

Cameron RH, Schüssler M (2015) The crucial role of surface magnetic fields for the solar dynamo. Science 347:1333–1335

Cameron RH, Jiang J, Schmitt D, et al (2010) Surface flux transport modeling for solar cycles 15–21: Effects of cycle-dependent tilt angles of sunspot groups. Astrophys J 719:264–270

Cameron RH, Schmitt D, Jiang J, et al (2012) Surface flux evolution constraints for flux transport dynamos. Astron Astrophys 542:A127

Cameron RH, Jiang J, Schüssler M (2016) Solar cycle 25: Another moderate cycle? Astrophys J Lett 823:L22

Charbonneau P, Knapp B (1995) Genetic Algorithms in Astronomy and Astrophysics. Astrophys J Supp Ser 101:309

De Rosa ML, Schrijver CJ (2006) Consequences of large-scale flows around active regions on the dispersal of magnetic field across the solar surface. In: Fletcher K, Thompson M (eds) Proceedings of SOHO 18/GONG 2006/HELAS I, Beyond the spherical Sun, p 12

DeVore CR (1987) The decay of the large-scale solar magnetic field. Solar Phys 112:17–35

DeVore CR, Sheeley Jr. NR, Boris JP (1984) The concentration of the large-scale solar magnetic field by a meridional surface flow. Solar Phys 92:1–14

Dikpati M, Gilman PA, Ulrich RK (2010) Physical Origin of Differences Among Various Measures of Solar Meridional Circulation. Astrophys J 722:774–778

Gizon L, Duvall JT. L., Larsen RM (2001) Probing Surface Flows and Magnetic Activity with Time-Distance Helioseismology. In: Brekke P, Fleck B, Gurman JB (eds) Recent Insights into the Physics of the Sun and Heliosphere: Highlights from SOHO and Other Space Missions, p 189

Hanasoge SM (2022) Surface and interior meridional circulation in the sun. Living Rev Solar Phys 19:3

Hathaway DH, Rightmire L (2010) Variations in the sun's meridional flow over a solar cycle. Sci 327:1350–1352

Hathaway DH, Wilson RM, Reichmann EJ (1994) The shape of the sunspot cycle. Solar Phys 151:177–190

Hathaway DH, Beck JG, Bogart RS, et al (2000) The photospheric convection spectrum. Solar Phys 193:299

Hazra G (2021) Recent advances in the 3d kinematic babcock-leighton solar dynamo modeling. J Astrophys Astron 42:22

Hickmann KS, Godinez HC, Henney CJ, et al (2015) Data assimilation in the adapt photospheric flux transport model. Solar Phys 290:1105–1118

Hung CP, Brun AS, Fournier A, et al (2017) Variational Estimation of the Large-scale Time-dependent Meridional Circulation in the Sun: Proofs of Concept with a Solar Mean Field Dynamo Model. Astrophys J 849:160

Iijima H, Hotta H, Imada S, et al (2017) Improvement of solar-cycle prediction: Plateau of solar axial dipole moment. Astron Astrophys 607:L2

Iijima H, Hotta H, Imada S (2019) Effect of morphological asymmetry between leading and following sunspots on the prediction of solar cycle activity. Astrophys J 883:24

Jiang J (2020) Nonlinear Mechanisms that Regulate the Solar Cycle Amplitude. Astrophys J 900:19

Jiang J, Işik E, Cameron RH, et al (2010) The effect of activity-related meridional flow modulation on the strength of the solar polar magnetic field. Astrophys J 717:597–602

Jiang J, Cameron RH, , et al (2011) The solar magnetic field since 1700. i. characteristics of sunspot group emergence and reconstruction of the butterfly diagram. Astron Astrophys 528:A82

Jiang J, Cameron RH, Schüssler M (2014a) Effects of the scatter in sunspot group tilt angles on the large-scale magnetic field at the solar surface. Astrophys J 791:5

Jiang J, Hathaway DH, Cameron RH, et al (2014b) Magnetic flux transport at the solar surface. Space Sci Rev 186:491–523

Jiang J, Cameron RH, Schüssler M (2015) The cause of the weak solar cycle 24. Astrophys J Lett 808:L28

Jiang J, Wang JX, Jiao QR, et al (2018) Predictability of the solar cycle over one cycle. Astrophys J 863:159

Jiang J, Song Q, Wang JX, et al (2019) Different contributions to space weather and space climate from different big solar active regions. Astrophys J 871:16

Jiang J, Zhang Z, Petrovay K (2022) Comparison of physics-based prediction models of solar cycle 25. (submitted)

Karak BB, Cameron R (2016) Babcock-Leighton Solar Dynamo: The Role of Downward Pumping and the Equatorward Propagation of Activity. Astrophys J 832:94

Leighton RB (1964) Transport of magnetic fields on the sun. Astrophys J 140:1547–1562

Lemerle A, Charbonneau P (2017) A coupled 2 × 2d babcock-leighton solar dynamo model. ii. reference dynamo solutions. Astrophys J 834:133

Lemerle A, Charbonneau P, Carignan-Dugas A (2015) A coupled 2 × 2d babcock-leighton solar dynamo model. i. surface magnetic flux evolution. Astrophys J 810:78

Mackay DH, Lockwood M (2002) The evolution of the sun's open magnetic flux. ii. full solar cycle simulations. Solar Phys 209:287–309

Mackay DH, Yeates AR (2012) The sun's global photospheric and coronal magnetic fields: Observations and models. Living Rev Solar Phys 9:6

Mackay DH, Priest ER, Lockwood M (2002) The evolution of the sun's open magnetic flux. i. a single bipole. Solar Phys 207:291–308

Martin-Belda D, Cameron RH (2016) Surface flux transport simulations: Effect of inflows toward active regions and random velocities on the evolution of the sun's large-scale magnetic field. Astron Astrophys 586:A73

Martin-Belda D, Cameron RH (2017) Inflows towards active regions and the modulation of the solar cycle: A parameter study. Astron Astrophys 597:A21

McCloughan J, Durrant CJ (2002) A method of evolving synoptic maps of the solar magnetic field. Solar Phys 211:53–76

Nagy M, Lemerle A, Labonville F, et al (2017) The Effect of "Rogue" Active Regions on the Solar Cycle. Solar Phys 292:167

Nagy M, Lemerle A, Charbonneau P (2020) Impact of nonlinear surface inflows into activity belts on the solar dynamo. J Space Weather Space Climate 10:62

Petrovay K (1994) Theory of passive magnetic field transport. In: Rutten RJ, Schrijver CJ (eds) Solar Surface Magnetism, p 415

Petrovay K, Talafha M (2019) Optimization of surface flux transport models for the solar polar magnetic field. Astron Astrophys 632:A87

Petrovay K, Nagy M, Yeates AR (2020) Towards an algebraic method of solar cycle prediction. i. calculating the ultimate dipole contributions of individual active regions. J Space Weather Space Clim 10:50

Schrijver CJ (1989) The effect of an interaction of magnetic flux and supergranulation on the decay of magnetic plages. Solar Phys 122:193–208

Schrijver CJ (2001) Simulations of the photospheric magnetic activity and outer atmospheric radiative losses of cool stars based on characteristics of the solar magnetic field. Astrophys J 547:475–490

Schrijver CJ, DeRosa ML (2003) Photospheric and heliospheric magnetic fields. Solar Phys 212:165–200

Schrijver CJ, Liu Y (2008) The global solar magnetic field through a full sunspot cycle: Observations and model results. Solar Phys 252:19–31

Schrijver CJ, Title AM (2001) On the formation of polar spots in sun-like stars. Astrophys J 551:1099–1106

Schrijver CJ, DeRosa ML, Title AM (2002) What is missing from our understanding of long-term solar and heliospheric activity? Astrophys J 577:1006–1012

Sheeley Jr. NR (2005) Surface evolution of the sun's magnetic field: A historical review of the flux-transport mechanism. Living Rev Solar Phys 2:5

Sheeley Jr. NR, DeVore CR (1986) The decay of the mean solar magnetic field. Solar Phys 103:203–224

Snodgrass HB, Ulrich RK (1990) Rotation of doppler features in the solar photosphere. Astrophys J 351:309

Sun X (2018) Polar field correction for hmi line-of-sight synoptic data. arXiv e-prints URL https://arxiv.org/abs/1801.04265

Talafha M, Nagy M, Lemerle A, et al (2022) Role of observable nonlinearities in solar cycle modulation. Astron Astrophys 660:A92

Upton LA, Hathaway DH (2014) Predicting the sun's polar magnetic fields with a surface flux transport model. Astrophys J 780:5

Upton LA, Hathaway DH (2018) An updated solar cycle 25 prediction with adt: The modern minimum. Geophys Res Lett 45:8091–8095

Van Ballegooijen AA, Cartledge NP, Priest ER (1998) Magnetic flux transport and the formation of filament channels on the sun. Astrophys J 501:866–881

van Driel-Gesztelyi L, Green LM (2015) Evolution of Active Regions. Living Rev Solar Phys 12:1

Virtanen IOI, Virtanen II, Pevtsov AA, et al (2017) Reconstructing solar magnetic fields from historical observations. ii. testing the surface flux transport model. Astron Astrophys 604:A8

Wang YM (2017) Surface flux transport and the evolution of the sun's polar fields. Space Sci Rev 210:351–365

Wang YM, Lean JL (2021) A new reconstruction of the sun's magnetic field and total irradiance since 1700. Astrophys J 920:100

Wang YM, Sheeley, Jr. NR (1989) Average properties of bipolar magnetic regions during sunspot cycle 21. Solar Phys 124:81–100

Wang YM, Sheeley, Jr. NR (1991) Magnetic flux transport and the sun's dipole moment: New twists to the babcock-leighton model. Astrophys J 375:761–770

Wang YM, Nash AG, Sheeley, Jr. NR (1989) Magnetic flux transport on the sun. Science 245:712–718

Wang YM, Lean J, Sheeley, Jr. NR (2002) Role of a variable meridional flow in the secular evolution of the sun's polar fields and open flux. Astrophys J 577:L53–L57

Wang ZF, Jiang J, Wang JX (2021) Algebraic quantification of an active region contribution to the solar cycle. Astron Astrophys 650:A87

Weinzierl M, Yeates AR, Mackay DH, et al (2016) A new technique for the photospheric driving of non-potential solar coronal magnetic field simulations. Astrophys J 823:55

Whitbread T, Yeates AR, Muñoz Jaramillo A, et al (2017) Parameter optimization for surface flux transport models. Astron Astrophys 607:A76

Whitbread T, Yeates AR, Muñoz Jaramillo A (2019) The need for active region disconnection in 3d kinematic dynamo simulations. Astron Astrophys 627:A168

Wilson PR, McIntosh P, Snodgrass HB (1990) The reversal of the solar polar magnetic fields. i. the surface transport of magnetic flux. Solar Phys 127:1–9

Worden J, Harvey J (2000) An evolving synoptic magnetic flux map and implications for the distribution of photospheric magnetic flux. Solar Phys 195:247–268

Yeates AR (2014) Coronal magnetic field evolution from 1996 to 2012: Continuous non-potential simulations. Solar Phys 289:631–648

Yeates AR (2017) Sparse reconstruction of electric fields from radial magnetic data. Astrophys J 836:131

Yeates AR (2020) How good is the bipolar approximation of active regions for surface flux transport? Solar Phys 295:119

Yeates AR, Bhowmik P (2022) Automated driving for global nonpotential simulations of the solar corona. Astrophys J 935:13

Yee K (1966) Numerical solution of inital boundary value problems involving maxwell's equations in isotropic media. IEEE Trans Antennas Propagation 14:302–307