Solar activity simulation and forecast with a flux-transport dynamo

DOI:

[10.1093/mnras/sty1625](https://doi.org/10.1093/mnras/sty1625)

Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Macario Rojas, A., Smith, K., & Roberts, P. (2018). Solar activity simulation and forecast with a flux-transport dynamo. *Royal Astronomical Society. Monthly Notices*, 479(3), 3791-3803. <https://doi.org/10.1093/mnras/sty1625>

Published in:

Royal Astronomical Society. Monthly Notices

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

Solar activity simulation and forecast with a flux-transport dynamo

Alejandro Macario-Rojas,^{*} Katharine L. Smith, Peter C. E. Roberts

School of Mechanical Aerospace and Civil Engineering, The University of Manchester, M13 9PL, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present the assessment of a diffusion-dominated mean field axisymmetric dynamo model in reproducing historical solar activity and forecast for solar cycle 25. Previous studies point to the Sun’s polar magnetic field as an important proxy for solar activity prediction. Extended research using this proxy has been impeded by reduced observational data record only available from 1976. However, there is a recognised need for a solar dynamo model with ample verification over various activity scenarios to improve theoretical standards. The present study aims to explore the use of helioseismology data and reconstructed solar polar magnetic field, to foster the development of robust solar activity forecasts. The research is based on observationally inferred differential rotation morphology, as well as observed and reconstructed polar field using artificial neural network methods via the hemispheric sunspot areas record. Results show consistent reproduction of historical solar activity trends with enhanced results by introducing a precursor rise time coefficient. A weak solar cycle 25, with slow rise time and maximum activity -14.4% ($\pm 19.5\%$) with respect to the current cycle 24 is predicted.

Key words: Sun: activity – Sun: magnetic field – sunspots

1 INTRODUCTION

The application of theoretical dynamo models that aim to reproduce aspects of solar activity is based upon the current understanding of observed and inferred solar phenomena. Despite the significant efforts to produce self-consistent models that can faithfully represent solar activity, the highly complex underlying mechanisms and voids in solar physics knowledge have to date prevented theorists from doing so. Current wisdom points to the application of varied nonlinear three-dimensional processes converging on observed phenomena, whose large-scale overall evolution is remarkably systematic; this allows, under specific circumstances discussed below, the axisymmetrisation of leading phenomena simplifying the modelling problem.

Axisymmetric solar dynamo models used for forecasts are related by magnetohydrodynamic (MHD) theory, the main difference between these models are the parameterisations that result in differing dynamo characteristics. In general terms, in order to reproduce the main temporo-spatial characteristics of the observed Sun, dynamo model calibration is necessary. The identification of plausible combinations of parameters yielding acceptable reproduction of observed activity is an arduous endeavour that does not necessarily guarantee forecast reliability. The main limitation in this respect, is the narrow time record of relevant solar activity features that can provide suitable input parameters for MHD models in addition to

validation data for the outputs of these models. For instance, the approach presented by Jiang et al. (2007) using observed photospheric magnetic field data as the main parameter for prediction, correctly anticipated the magnitude of the solar cycle 24 and satisfactorily reproduced the three precedent cycles, that is for the cycles where photospheric magnetic field data was available. Some attempts to overcome limitations of solar activity data use other proxies such as the relatively large sunspot number (SSN) record to calibrate solar dynamo models. In this regard, Dikpati (2008) and Karak (2010) used SSN cycle periods to adjust the dynamo model meridional flow speed, and Kitiashvili & Kosovichev (2008) fitted simplified dynamo equations to observed SSN activity.

The solar activity forecast approach presented in this investigation is centred on reproducing mean large-scale solar activity features using a dynamo model. The preferred solar activity leading indicator is the solar poloidal magnetic field strength in a similar approach to Jiang et al. (2007). However, in this investigation we propose the use of polar magnetic field, and the selection of suitable dynamo parameterisations based on the application of a detailed axial solar rotation profile inferred from helioseismology. Additionally, in order to enrich the typical dynamo calibration process, the short record of solar polar magnetic field strength is extended by means of neural network reconstruction. The investigation is organised into two main sections. The following section 2 describes the mean field axisymmetric dynamo model and its parameterisations, as well as characteristic model assumptions and caveats. Section 3 presents the characteristics of the resulting solar-like dynamo, the

^{*} E-mail: alejandromacario.rojas@manchester.ac.uk

proposed solar polar field strength reconstruction, and finally the performance of the overall approach in reproducing solar activity is discussed.

2 THE MEAN FIELD AXISYMMETRIC DYNAMO MODEL

A key characteristic found in mean field axisymmetric dynamo models is the solution of the dynamo problem equations in a meridional slab with suitable boundary conditions. In this investigation a finite differences computational code is developed incorporating the openly-available Fortran SURYA code (Choudhuri 2005) to solve the system of coupled differential equations. The MATLAB® code used in this investigation employs separated poloidal and toroidal magnetic diffusivity models as in the SURYA code, and by extension by Jiang et al. (2007), as well as the same approach for magnetic growth suppression. However, the solar model parameterisations approach is distinct in this investigation. Since the solar plasma's angular velocity differential rotation profile is well constrained by helioseismology, a realistic parameterisation model is chosen over the typical analytical fit. This in turn underpins the adjustment of the loosely defined solar magnetic diffusivity and poloidal field source parameterisation models enabling the successful reproduction of major solar activity features. The impact of the proposed modifications of the parameterisations on the reproduction of solar activity are discussed after a detailed presentation of the models. In first instance the bulk field dynamics are discussed in subsections 2.1 and 2.2; subsequently the polar field source and magnetic diffusivity models are presented in subsection 2.3.

2.1 Angular Velocity

The dominant angular velocity, hereinafter referred to as Ω , and weaker meridional circulation, define the bulk velocity field vector ($\mathbf{v}_T$) used in mean field axisymmetric dynamo models. Eq. (1) summarises the velocity field vector in spherical coordinates in terms of the axial rotation (azimuthal) and the meridional circulation component $\mathbf{v}$, where r and θ are the local radius and colatitude respectively.

$$\mathbf{v}_T = r \sin \theta \Omega(r, \theta) \hat{\mathbf{e}}_\phi + \mathbf{v} \quad (1)$$

The main enigma in solar dynamo theory is how the magnetic field is sustained against diffusion and recycled periodically. A natural explanation for the generation of the azimuthal (toroidal) field from the meridional (poloidal) field is that axial rotation stretches out the local polar field lines frozen to the moving plasma; the resulting wound magnetic field builds the toroidal field (ω -effect). Thus, regions of large shear concentrate and strengthen the toroidal field.

Figure 1 shows three representations of the longitudinal averaged rotation map as function of solar latitude and depth. The most common rotation map (Charbonneau et al. 1999) used in solar dynamo modelling is shown in the analytical envelope in Figure 1a. In this subfigure, the comparison of depth profiles for various latitudes between the analytical fit model and inferred data via helioseismology, highlights a rough angular velocity approximation. On the other hand, the profile in Figure 1b is derived from insightful mathematical laws (Balbus et al. 2012), wherein differential rotation is explained through angular entropy gradients and inertial angular velocity. Finally, Figure 1c shows an O-grid mesh fitted to data from the Global Oscillation Network Group (GONG) (Muñoz-Jaramillo et al. 2009). This realistic model captures most of the characteristics

of the average differential rotation morphology which, according to the isorotation contour patterns in Figure 1c, can be divided into three main concentric rings.

Starting from $0.5R_\odot$ towards the surface, an innermost ring shows near constant angular velocity, i.e. solid body rotation, which is abruptly wrenched at $\sim 0.7R_\odot$ (the tachocline) with various magnitudes, depending on latitude. A mid ring is bounded by the tachocline as inner radius up to $\sim 0.95R_\odot$. Within this ring isorotation contours diverge from radial spokes as latitudes approach the solar equator, this is in contrast to the observed trend in Figure 1a, where no divergence with latitude is observed. In the outer ring extending from $0.95R_\odot$ to the surface, the isorotation contours shown in Figure 1c gradually match the observed surface rotation profile; it is suspected that this outer ring is dominated by combinations of vigorous turbulent stresses and magnetic braking (Balbus et al. 2012) reducing angular rotation speed. Due to the fundamental importance of the plasma flow morphology to the current understanding of the solar dynamo, this investigation departs from the typical analytical model in Figure 1a and employs the model shown in Figure 1c. Implications of the application of this differential rotation model for other dynamo parameterisations are reported accordingly in the following subsections.

2.2 Meridional Circulation

Vigorous rotation acting on anisotropic and compressible plasma in the convection zone generates a second important macroscopic flow. It is believed that the combined effect of plasma in differential axial rotation, pressure gradients, and turbulent buoyant convective plasma generate cellular circulation that collectively may be the cause of the observed meridional circulation on the solar surface. Analysis of the effect of multiple cellular circulation zones on the reproduction of observed solar activity is studied by Hazra et al. (2014). They concluded that as long as equatorward flow exists at the bottom of the convection zone at low solar latitudes, where azimuthal magnetic field is possibly generated, solar activity can be reproduced effectively in solar dynamo models. In addition, Hazra et al. (2014) assessed the effect of meridional flow speed variation on the solar cycle period in an array of radially stacked cells. The study indicates a significant influence of the innermost cell speed in determining the length of the solar cycle.

Observed solar surface meridional circulation indicates that this may vary with the progress of the solar cycle, e.g. at mid latitudes $\sim 10 \text{ m s}^{-1}$ during the 23rd solar maximum and $\sim 14 \text{ m s}^{-1}$ by the end of that cycle (Hathaway & Upton 2014). In this respect, the Karak (2010) study suggests that meridional flow fluctuation may be a fundamental phenomenon influencing the duration and magnitude of solar cycles, or at least an indicator of inner activity. The current definition of reduced meridional flow parameters from observational data, makes the use of the stable single hemispheric cell approach, which is the customary choice in coarse reproduction of solar-like activity. In this investigation we adopt this standard approach acknowledging that this would imply some inherent degree of uncertainty in solar activity reproduction and forecasting. However this is justified by the fact that a single hemispheric cell encompasses the essential meridional flow component of the most up to date representation of the solar dynamo.

The two-dimensional meridional circulation with velocity $\mathbf{v}$ can be expressed through a stream function ψ as shown in equation (2) assuming polytropic density stratification given by equation (3). In this regard, Cardoso & Lopes (2012) show that a more realistic description of density stratification may lead to noticeable

Figure 1. Solar longitudinal averaged isorotation contours (left column) and depth profiles (right column) for selected solar latitudes. The subfigures show examples of axial solar rotation models: a) Analytical fit commonly used in dynamo modelling, b) Mathematical model based on the thermal wind equation, and c) Mapping from helioseismology (GONG). Vertical lines in the depth profiles show the core location of the tachocline, solid curves are the profiles from the models, and dashed curves correspond to inferred observational data.

changes in plasma flow evolution. However, considering the existing uncertainty produced by the meridional circulation structure, the standard polytropic stratification is deemed sufficient for the axisymmetric dynamo model used in this investigation.

$$\rho \mathbf{v} = \nabla \times [\psi(r, \theta) \hat{\mathbf{e}}_\phi] \quad (2)$$

$$\rho = \left(\frac{R_\odot}{r} - \gamma \right)^m \quad (3)$$

$$\psi r \sin \theta = \psi_0 (r - R_p) \sin \left[\frac{\pi (r - R_p)}{(R_\odot - R_p)} \right] \left[1 - e^{-\beta_1 \theta^\epsilon} \right] \times \left[1 - e^{\beta_2 (\theta - \pi/2)} \right] e^{[(r - r_0)/r]^2} \quad (4)$$

The hemispheric stream function model in equation (4) (Chatterjee et al. 2004) employs the parameters reported in Table 1. With the exception of β values and poleward flow velocity near the surface at

Table 1. Meridional circulation parameters.

Parameter	Reference model	This model	Unit
β_1	1.36×10^{-8}	1.50×10^{-8}	m^{-1}
β_2	1.63×10^{-8}	1.80×10^{-8}	m^{-1}
γ	0.95	0.95	–
Γ	3.47×10^8	3.47×10^8	m
ϵ	2.0000001	2.0000001	–
m	1.5	1.5	–
R_p	$0.61 R_\odot$	$0.61 R_\odot$	m
v_0	-29^a	-25	m s^{-1}

^a Errata reported in Jiang et al. (2007)

mid-latitudes, v_0 , the parameters used in this investigation are equal to those reported by Chatterjee et al. (2004), hereinafter referred to as the reference model. It is important to bear in mind that variation in parameter values between this investigation and the reference model result from the use of different differential rotation maps as discussed in subsection 2.1. As the reproduction of solar-like activity depends on the characteristics of the merged parameterisations, similar variations are expected to result for all parameterisations of the solar dynamo model.

The meridional circulation model used in this investigation generates a maximum equatorward counterflow of 2.29 m s^{-1} at the tachocline region, which is consistent with the observed trend in equatorward sunspot drift near midlatitudes (Hathaway et al. 2003). In addition, the model yields a reference Schwabe cycle of 10.87 years in agreement to the average cycle of interest in this investigation and similar to that reported by Jiang et al. (2007) (10.80 years).

2.3 Poloidal Field Source and Magnetic Diffusivity Models

The generation of the poloidal magnetic field is of significance to the current theoretical explanation of the solar dynamo. The origin of this magnetic field component can be ascribed to small-scale toroidal flux tube rotation under the action of Coriolis forces (decay of active regions by cyclonic turbulence termed the α -effect), the body-axial flux tube twist contributing to the poloidal component (Priest 2014), possibly by magnetic pumping at the base of the convection zone, and predominantly by diffusion and reconnection of near-surface tilted dipolar regions viz. the Babcock-Leighton process. These chiefly near-surface phenomena are argued to be the source of the poloidal field regeneration. The stochastic nature of the poloidal field generation dominates the part of the Schwabe cycle, from solar maximum to the end of the cycle, and is believed to be one of the most important causes of solar cycle irregularities (Jiang et al. 2007). This theory is supported by the observed regularity and predictability of the rising phase of SSN cycles vis-à-vis the declining (Babcock-Leighton dominated) unstable phase. In continuation with the magnetic cycle, the generated poloidal magnetic field is transported to the poles and solar interior by meridional flow and magnetic diffusion around evolving photospheric convection cells (supergranule cells).

Another identified implication derived from the Ω model selection and subsequent cumulative modifications, is the required explicit suppression of the poloidal field at the dynamo poles. The Babcock-Leighton process in this investigation is phenomenologically represented by α in equation (5), modified¹ from Hotta &

¹ In Hotta & Yokoyama (2010), the argument of the exponential function in the poloidal source equation includes the term $-\pi/4$.

Yokoyama (2010). In this equation the generation of the poloidal field is reasonably concentrated above $0.95R_{\odot}$, and suppressed at the poles. This α model differs from that used in the reference model, and the SURYA code, where only the latitudinal dependence of the Coriolis effect (represented by the factor $\cos \theta$ in equation (5)) is considered. The suppression of α at the poles and the heuristic value of $\alpha_0 = 15 \text{ m s}^{-1}$ ensures the relaxation of the dynamo model in this investigation to a stable periodic solution from an initial arbitrary state (super-critical α).

$$\alpha = \frac{\alpha_0}{4} \cos \theta \left[1 + \operatorname{erf} \left(\frac{r - 0.95R_{\odot}}{0.05R_{\odot}} \right) \right] \left[1 - \operatorname{erf} \left(\frac{r - R_{\odot}}{0.01R_{\odot}} \right) \right] \times \underbrace{\left(\frac{1}{1 + e^{-30\theta}} \right)}_{\text{Suppression at poles}} \sin \theta \quad (5)$$

The final fundamental ingredient in the reproduction of macroscopic solar fluctuations, is the turbulent magnetic diffusivity in the convection zone. The importance of the solar turbulent diffusivity resides in the interconnecting role between the polar and toroidal fields, which are believed to be generated in separate specific regions of the Sun. Unfortunately, the solar turbulent diffusivity process is currently poorly understood and therefore poorly constrained in solar dynamo models. This uncertainty has led to two main schools of thought considering the diffusivity behaviour in the Sun's interior and the representation of this in mean-field dynamo models. With the exception of its near-surface value, where there is apparent consensus supported by mixing-length theory, the adoption of either high or low turbulent magnetic diffusivity values in solar dynamo models is used Muñoz-Jaramillo et al. (2010). Because there is little evidence of the phenomena taking place in the Sun's interior, flux transport dynamo parameters are adjusted within characteristic ranges to match surface phenomena that in both diffusivity cases may reproduce observed patterns. However, the implications for solar activity forecasting resulting from each approach are profound. For example, in the low diffusivity scenario² where advection is the main mechanism communicating the magnetic field in the majority of the convection zone, a cycle memory effect exists in dynamo models due to high latency in the magnetic field communication. On the other hand, high diffusivity favours faster communication and interaction between magnetic field components, facilitating the establishment of hemispherical dipolar parity and restraining hemispheric asymmetry.

At present the high turbulent diffusivity approach has shown reasonable results in reproducing solar cycles behaviours as discussed above implying that long-term solar activity memory, may be unnecessary to explain the behaviour of solar cycles. Furthermore, the observed lack of correlation between solar activity during a given Schwabe cycle and the polar magnetic field observed by its end (Jiang et al. 2007), highlight the randomness involved in the creation of the polar field and thereby the relevance of solar polar magnetic precursor forecasts. The high magnetic diffusivity approach used in this investigation is similar to that reported by various authors (Chatterjee et al. 2004; Hazra et al. 2014; Jiang et al. 2007; Karak 2010) with separated poloidal and toroidal diffusivity components in equation (6) and equation (7)³.

² Typically $\eta = 10^6 - 10^{12} \text{ cm}^2 \text{ s}^{-1}$ from bottom to top of the convection zone (Dikpati & Gilman 2006), produce diffusion time across the convection zone of 200 years.

³ This equation differs in $0.95R_{\odot}$ reported in the reference model and uses $0.975R_{\odot}$ instead, errata reported in (Choudhuri 2005)

Suitable diffusivity model parameter values, in combination with aforementioned solar dynamo parameterisations, are obtained by solar dynamo calibration with observed solar activity patterns, e.g. sunspot butterfly diagrams. In this investigation this has led to the use of $\eta_{SCZ1} = 3 \times 10^{10} \text{ cm}^2 \text{ s}^{-1}$, a slightly lower value than the reference model $\eta_{SCZ1} = 4 \times 10^{10} \text{ cm}^2 \text{ s}^{-1}$, and similar $\eta_{RZ} = 2.2 \times 10^8 \text{ cm}^2 \text{ s}^{-1}$ and $\eta_{SCZ} = 2.4 \times 10^{12} \text{ cm}^2 \text{ s}^{-1}$ to the reference model.

$$\eta_P = \eta_{RZ} + \frac{\eta_{SCZ}}{2} \left[1 + \operatorname{erf} \left(\frac{r - 0.70R_{\odot}}{0.025R_{\odot}} \right) \right] \quad (6)$$

$$\eta_t = \eta_{RZ} + \frac{\eta_{SCZ1}}{2} \left[1 + \operatorname{erf} \left(\frac{r - 0.72R_{\odot}}{0.025R_{\odot}} \right) \right] + \frac{\eta_{SCZ}}{2} \left[1 + \operatorname{erf} \left(\frac{r - 0.975R_{\odot}}{0.025R_{\odot}} \right) \right] \quad (7)$$

It is expected that the strong toroidal field at the bottom of the convection zone, where it is supposedly generated by high differential rotation, presents suppressed diffusivity. This is a plausible assumption because dynamically stable plasma is found in the near-radiative zone and because strong flux tubes may contribute to reduce turbulent diffusivity. Additionally, the extended suppressed toroidal diffusivity in this region, helps to preserve the former high-latitude toroidal magnetic field whilst it is transported frozen to the moving plasma by the equatorward counter flow (meridional circulation near the bottom of the convection zone). As the toroidal field is stronger than the poloidal field within the convection zone, lower diffusivity in the toroidal field is expected until it matches the estimated common diffusivity at the surface. Unlike the double step profile of η_t , the weak poloidal field may be easily diffused throughout the convection zone suggesting a single step profile.

2.3.1 Solar Activity Simulation

Sunspots, regions subject to magnetic buoyancy, are considered as a suitable proxy of solar activity for comparison between observational and simulated data. In this respect, the magnetic growth suppression approach implemented in the SURYA algorithm is used. In this approach the condition of the toroidal field at $0.71R_{\odot}$ is monitored every 10 simulation days (Choudhuri 2005). Wherever the field exceeds the threshold of 10^5 G in this layer, half of this field is transferred to upper grid points to emulate a magnetic buoyancy-like effect. Thereby these grid points represent active regions where sunspots may appear. Figure 2 shows the mean toroidal field component evolution in a steady dynamo solution (super-critical condition) from the code used in this investigation. The figure shows the characteristic toroidal field evolution during a simulated Schwabe cycle wherein weak field regions have been removed for illustration purposes. The complimentary Video A1 shows in detail the evolution of the toroidal field strengths and highlights the ultimate differences between the standard parameterisations, and those from this investigation. Note that the solar dynamo simulation using the aforementioned settings corresponds to an average solar cycle and that the number of active regions is defined by the resolution of the simulation, the implications of these factors are discussed in section 3; in this investigation a meridional slab with 256 radial and 256 angular partitions is used.

Figure 3 presents the probability density functions (PDF) of emerging active regions during a simulated Schwabe cycle from both standard parameterisations and those used in this investigation. The superimposed shaded areas correspond to the PDFs from

Figure 2. Theoretical toroidal evolution of strong magnetic field. Weak magnetic field regions, of less than approximately 10^3 G, have been removed for illustration purposes. Light solid and dark dashed lines represent magnetic field of opposite polarity. Active regions correspond to magnetic field strengths higher than 10^5 G, which are believed to be the source of buoyant rise of magnetic flux tubes generating sunspots. In this illustration, emerging active regions are projected on the surface plane constituting a theoretical butterfly diagram.

Figure 3. Probability density functions from emerging active regions and regular sunspot areas. The curves represent the simulated average solar cycle activity, superimposed shaded areas correspond to observed solar cycles. The darkest shaded area corresponds to the solar cycle 23 which importance for this investigation is discussed in section 3.

observed regular sunspot area activity (greater than 100 millionths of a hemisphere) during the cycles 12–23. The set of parameterisations presented in this investigation favour the growth of azimuthal flux in a more even fashion at mid-latitudes as seen from the comparison of the right tail of the PDFs. In addition to this, the skewness of the PDF of this investigation shows better resemblance to the PDF from observational data. Assuming that magnetic tension generated by a strong toroidal field at the bottom of the convection zone is able to restrain or even momentarily stop its advection by the equatorward counterflow, the incorporation of this phenomenon in the dynamo model could favour the release of solar activity at higher latitudes further improving the PDF profile. However, further work is required to verify this hypothesis, as such in this investigation we use the dynamo parameterisations described above. The PDFs correlation coefficients between observed and simulated solar activity are 0.4921 and 0.6374 for the standard parameterisations and those used in this investigation respectively. Thus the parameterisations

in the current study represent an improvement over the standard model.

3 SOLAR CYCLES FORECASTING

The designed solar-like dynamo model reaches a steady state activity response⁴ from an arbitrary initial magnetic state after a small number of simulated Schwabe cycles. It is worth mentioning that the stable characteristic response and the duration of the transient response during the initialisation of the dynamo model depend on the parameterisations employed. Since parameterisations remain invariant with time in the dynamo model, and these represent well-known average data in the case of the differential rotation and to a lesser extent in the meridional circulation, the simulated solar activity corresponds to an average solar cycle. In reality, fluctuations in all solar parameters confer uniqueness to each observed solar cycle. For this reason, current axisymmetrisation of the solar problem implemented in the solar dynamo, and poorly understood internal phenomena such as the possible occurrence of velocity quenching in the equatorward counterflow at the base of the convection zone, lead to coarse albeit representative approximations of observed solar activity. In addition to this, as discussed in the previous subsection 2.3.1, the number of active regions is defined by the resolution of the simulation, that is the denser the meridional slab grid the higher the number of active regions although the activity pattern evolution is preserved. For these reasons, the following effort is directed towards the identification of an observed average solar cycle to suitably scale and fine tune the dynamo model output.

⁴ The initial magnetic state used in this investigation is zero polar field in both hemispheres and an arbitrary toroidal magnetic field strength of equal but opposite polarity for hemispheres to establish dipolar parity. Although dipolar parity is observed in the recent record of solar activity, it is argued that quadrupolar parity may appear under specific conditions, e.g. superimposed beating magnetic configuration components. Analysis of J.C. Staudacher's sunspot drawings covering the period 1749 – 1796 made by Arlt (2009) and further studied by Sokoloff et al. (2009) in the context of solar dynamo theory, suggest that cycles 0 and 1 may have shown quadrupolar parity.

In order to identify the average solar cycle, in this investigation the total sunspot area record reported in the butterfly diagrams⁵ and the monthly average SSN record⁶ are assessed. The SSN record encompasses the last 23 full cycles and the butterfly diagrams, containing hemispheric activity details shown Figure 4a, the last 12 full cycles. From the SSN record, cycle periods and maximum activity magnitudes are evaluated, whilst from the hemispheric sunspot area record latitudinal activity is evaluated. In this investigation hemispheric sunspot area data (in units of millionths of a hemisphere MSH) is grouped into small sunspots, transition sunspots, and regular sunspots according to the classification presented by Tlatov & Pevtsov (2014). By processing the complete record of hemispheric sunspot areas, it is possible to identify average sunspot activity as function of solar latitude and solar cycle evolution as shown in Figure 4b. The resulting average butterfly map is then used to quantify differences in hemispheric activity for each solar cycle. The cycle with closest hemispheric activity similarity to the average butterfly map is assigned with rank 1, up to the least similar with rank 13 in this case, i.e. each cycle in the butterfly record is related to a rank value. In a similar way, the identification of the average Schwabe period and SSN magnitude from the SSN record assists the ranking process of the solar cycles. Figure 4c summarises the data assessment for the common cycles and identifies the solar cycle 23 as the average cycle according to the cumulative ranking from the three indicators, i.e. butterfly diagrams, SSN maxima, and SSN cycle periods.

Having identified the average solar cycle it is possible to establish conversion factors to relate observed solar characteristics to the dynamo model output. This fundamental procedure enables the identification of the scaling relationship between the model output, i.e. representing an average solar cycle, and the observed average solar cycle 23. Firstly, in this investigation we are interested in a conversion factor to relate the theoretical polar magnetic field strength at SSN minimum, viz. instantaneous initial state, to the corresponding observed value⁷; this is because the solar polar magnetic field strength is used as the main parameter for solar activity forecast in a similar approach to Jiang et al. (2007). The smoothed polar field, solid lines in Figure 5, represent the average observed field in the polar caps with effective latitudes at $\pm 70^\circ$ (Svalgaard et al. 1978). This observable surface indicator of the solar state is of central importance in the fine adjustment of the dynamo model presented in this investigation, since it is hypothesised that the magnitude of the polar magnetic field is related to the poloidal magnetic structures underneath. In this respect, the method of meridional slab modification within the region enclosed by $0.8R_\odot$ and the external radius as a means to modulate the poloidal potential is implemented (Jiang et al. 2007). This method allows the even modulation of the poloidal magnetic field in the dynamo model slab. However, note in Figure 5 that the smoothed northern and southern hemispheric polar field curves differ slightly at time zero (at SSN minimum). In order to include these observed hemispheric asymmetries, in this investigation we use a sectioned modification meridional slab divided by the solar equator. In this divided meridional slab, the magnitude of the smoothed south hemisphere polar field is chosen arbitrarily as the reference value to compute the model's scale factors; in other words the characteristic southern dynamo polar magnetic field strength at a point near the surface and solar latitude 70° is proportional to the

respective observed value wherein the constant of proportionality is the scale factor. Finally, scaling the theoretical emerging active regions profile resulting from the poloidal potential modulation, hereinafter referred to as pseudo-SSN, to observed solar activity is necessary; the maximum smoothed SSN is used for this purpose. Attention of subsequent analyses should be centred only on approximately the first lustrum of solar activity, e.g. the shaded area in Figure 5, corresponding to the relatively stable SSN rising cycle phase as discussed in subsection 2.3.

The calibrated solar dynamo model for cycle 23, that is the fully scaled dynamo output from the selected solar activity indicators, is shown Figure 6. The root-mean-square error (RMSE) of the model's pseudo-SSN and the observed SSN for the rising and declining phases are reported in the figure. The effect of forced initial hemispheric poloidal potential asymmetry in the dynamo model is clearly observed in the polar magnetic strength magnitude evolution. From the plots in Figure 5 and Figure 6 sufficient resemblance during the rising phase of the SSN cycle is observed. Conversely, the most remarkable incompatibility is observed in the declining phase of the sunspot cycle as expected due to the highly random processes dominating the poloidal field generation as previously discussed. The main factor limiting conclusions about the adequacy of the solar dynamo with polar magnetic field as precursor for forecasting purposes, is the reduced record of polar solar magnetic activity encompassing only the last full three solar cycles. In order to overcome this limitation, we use extended polar field strengths reconstructed with the longer record of hemispheric sunspot areas. The reconstruction method and criteria is discussed in the following subsection 3.1.

3.1 Solar Polar Field Strength Reconstruction

With the advance of the solar cycle and as the generation of poloidal flux from toroidal flux develops near the solar surface, part of the generated flux is advected by meridional flow circulation to the poles where these gradually accumulate building up a distinguishable polar field of maximum strength near the end of the Schwabe cycle. Additionally, remnants of sunspot activity, i.e. remaining flux after cancellation and submergence, is diffused locally generating unipolar patches. According to the sunspot magnetic flux level and helicity (McMaken & Petrie 2017), a small amount may withstand flux fragmentation and cancellation to form long-lasting diffuse unipolar patches. These patches gradually form distinguishable trails of unipolar poleward drift with respect to the surrounding average weak diffuse magnetic field; these characteristic features can be observed in the averaged photospheric magnetic field reported in synoptic magnetic butterfly diagrams, Figure 7. However, the surface of the Sun is intrinsically active where magnetic field lines are in continuous efficient interaction. Spontaneous magnetic field lines emerging throughout the surface are fragmented, redistributed, and diffused in the magnetic carpet renewing itself in a few tens of hours (10 – 40 hours (Priest 2014)). This inherent surface activity affects trails of unipolar regions enhancing their decline. Thus, regular sunspots (flux density above $\sim 10^{21}$ Mx (Tlatov & Pevtsov 2014)) are more likely to contribute to long-lasting unipolar patches than small sunspots or pores ($\lesssim 10^{20}$ Mx). For this reason, in this investigation it is assumed that the information about past solar activity conveyed to the pole caps is plausibly related to chiefly regular sunspot activity. The overall aforementioned process generates a delay in the evolution of the polar magnetic field with respect to the sunspot activity. In this investigation, the phase shift between both proxies is identified at 1500 days (4.1 years) via cross-correlation

⁵ <https://solarscience.msfc.nasa.gov/greenwch.shtml>

⁶ <http://www.sidc.be/silso/datafiles>

⁷ <http://wso.stanford.edu/>

Figure 4. The average solar cycle. Subfigure *a*) presents the record of hemispheric sunspot areas as function of solar latitude and time. Subfigure *b*) shows the average butterfly map computed with data from the observed record. This benchmark map assist the ranking of the solar cycles, wherein rank 1 corresponds to closest similarity to this benchmark. Data from the SSN record provides two ranking benchmarks, the cycle period and SSN magnitude. The cycle number intersection of the three ranking sets is shown in Subfigure *c*). According to the cumulative ranking from the three indicators, solar cycle 23 is identified as the average cycle.

Figure 5. Selected indicators for observed solar cycle 23. The dark solid curves represent smoothed polar magnetic strength magnitudes, left scale, and the solid curve in light colour is the normalised SSN, right scale. In this investigation, the observed polar magnetic field strength values at time zero are used as modification inputs to hemispheric poloidal potential modulation in the dynamo model for each studied solar cycle. On the other hand, the observed maximum SSN value in the solar cycle 23 is used to scale the resulting theoretical dynamo activity output. In this way, observed cycle 23 SSN and pseudo-SSN maxima from the calibrated dynamo model are equivalent.

Figure 6. Same as Figure 5 for the calibrated dynamo to the observed solar cycle 23. Polar field magnitudes correspond to the solar latitudes $\pm 70^\circ$. Starting from the forced initial poloidal potential modulation to match the observed initial polar field magnitudes at time zero, the dynamo model evolves to polar field inversion around the fourth year in close resemblance to the observed trend in Figure 5 within the relatively stable SSN rising phase. The maximum value of the pseudo-SSN profile, obtained from the historical number of emerging active regions, is equivalent to the maximum value of observed SSN during the identified average solar cycle 23. Bear in mind that SSN equivalence is ensured only for the solar cycle 23 whereas in other cycles it is not necessarily true as expected from the diverse polar magnetic field strengths initial inputs used in the solar dynamo.

analysis of the SSN⁶ record and the Wilcox Solar Observatory polar magnetic field record (WSO)⁷. On the grounds of this reasoning, regular sunspots in the sunspot areas record (Figure 4) may contain sufficient information to reconstruct unrecorded polar field cycles. Artificial neural network methods can be suitably used to address

this task owing to the complex nature of the relationships involved as discussed below.

An artificial neural network (NN) is a computational model built with a collection of interconnected simple processing instances

Figure 7. Longitudinally averaged magnetic field (magnetic butterfly) diagram for the solar cycle 23. Reddish colour represents positive flux, and highly active sunspot areas are marked in black. Trails of unipolar poleward drift are observed in both solar hemispheres. In this investigation is argued that regular sunspot activity is likely to be captured in the polar magnetic field pole caps as this is transported by meridional circulation. If the information reaching the pole caps contain sufficient detail about past sunspot activity, then this may open the possibility to reconstruct unrecorded historical polar field from hemispheric sunspot areas. Image adapted from [Petric \(2013\)](#)

called neurones, resembling in essence a biological neural network. Each neurone consists of adjustable scalar weights to scale inputs, a fixed transfer function to model the neurone shape response, and an adjustable scalar bias to shift the function. Interconnection ensures that every neurone contributes to the collective neural network response. The objective of the NN implementation is to learn from and respond to given data in a desired way, wherein adaptability and capability of generalisation of the NN to new data is the ultimate objective. To this aim, a fundamental design stage in NN implementation is the *training* phase in which the NN learns in either an unsupervised or supervised manner ([Graupe 2007](#); [Demuth et al. 2014](#)). During the *training* phase in unsupervised NNs, the adjustable scalars in each neurone are adjusted aiming at the discrimination of data lying within specific criteria. Unsupervised NNs use input data to auto adjust their parameters dynamically. This type of NN is valuable in classification and pattern recognition, e.g. cluster analysis and image processing for object detection. On the other hand, a supervised NN uses data with labeled responses from input and expected response (target) data⁸. In this case, an iterative process modifies the adjustable scalars in each neurone during the *training* phase aiming at the faithful reproduction of target reference data. Supervised NNs find application in automated classification such as face/voice recognition, and in the modelling of continuous system response as a function of input predictor variables (regression). This later capability of supervised NNs is typically applied in time series predictions such as in financial analyses or even proposed for SSN trend prediction ([Li et al. 1990](#)). For the problem at hand, supervised NN training under the regression category was selected because the training dataset, i.e. predictor and target data, encompass sequential vectors.

When dealing with sequential vectors as NN inputs, adding feedback and delays typically improves the NN performance because these establish memory capability valuable for learning se-

quential patterns. This type of NN model is called a dynamic network, and relates the history of the input sequence and feedback sequence to establish the NN response. A Nonlinear AutoRegressive network with eXogenous inputs (NARX) is a dynamic model that can use several layers of neurones and independent input series, namely *exogenous series*, which are known to influence the target series. The NARX model output $y(t)$ is thus a function of past values of $y(t - n_y)$ (target or feedback data), and exogenous series $x(t - n_x)$, Eq. (8). The powerful NARX model may be applied in a myriad of applications. In this investigation the NARX model is used to represent a nonlinear dynamic system relating present and past data of the driving hemispheric sunspot area series (exogenous), and the record of solar polar field (target). The selection of this input dataset for the NARX model is justified due to the known influence of the exogenous series on the target series as discussed earlier in this section.

$$y(t) = f[y(t-1), \dots, y(t-n_y), x(t-1), \dots, x(t-n_x)] \quad (8)$$

The NARX model used for the reconstruction of the polar field in this investigation is presented in Figure 8. The two exogenous input series ($x_{1,2}(t)$) are the north and south hemispheric regular sunspot areas, and the target series ($y(t)$) is an observed polar field. It should be noted that the two hemispheric exogenous input series are used in the proposed reconstruction of each solar polar field. This is a reasonable approach assuming that sufficient magnetic communication exists between both hemispheres as suggested by the nearly mirrored polar magnetic strength history in Figure 5. Hence, possible loss of past solar activity information during the creation of unipolar patches and the subsequent hemispheric poleward drift, may be plausibly recovered from the regular sunspot areas record of both hemispheres. Resuming the description of the model, a layer of six interconnected sigmoid neurones in the so-called hidden layer, are fed with the exogenous and target series delayed 1 – 10 states (10 Julian days per state). This delay range is heuristically established according to the ultimate performance of the model. Finally, the output layer consists of a linear neurone used as a function approximator.

Specifically, Figure 8 illustrates a parallel NARX architecture (closed loop) useful in the practical application of the model. In a previous step, the model is trained to establish suitable weights and bias values. During the training phase the target series, in this case a polar field series, is available requiring no $\hat{y}(t)$ feedback. The training algorithm adopted is the standard Levenberg-Marquardt back-propagation, with internal data division of 70% training, 15% validation, and 15% testing. During the training phase, the Levenberg-Marquardt optimisation looks for adequate weights and bias values to reduce the error between the target signal and the model's output. Since internal data division and sampled data for training is stochastic, mainly to prevent target memorisation, computed weights and biases may differ amongst various trained models leading to various performance levels. This means that a trained model would have a unique set of bias and weight values related to the *randomly* sampled input data required by the training algorithm. Furthermore, depending on the overall sampled data during the training phase, the NARX may be more or less adaptable to interpret *new* input data according to its success in building a representative nonlinear dynamic system. Aiming at improving NN generalisation, multiple models may be trained to select the one able to generalise best. In this investigation 200 models for each polar field are trained, and in order to identify the most accurate one, the available target data is divided in two parts: 95% is used for the model training phase, and the remaining 5% is saved for a subsequent best model identification

⁸ One of the main differences between supervised and unsupervised NNs is that the later adjusts itself using the network inputs only since no target outputs are available.

Figure 8. Layer diagram of the Nonlinear AutoRegressive network with eXogenous inputs (NARX) in parallel configuration (closed loop). The network uses 6 hidden layer neurones and the input and feedback delays range 1 – 10. The two exogenous inputs correspond to the hemispheric regular sunspot area records. During the training phase, that is the procedure to establish weights and biases values, the target series (one observed hemispheric polar field series at a time) is available requiring a series-parallel architecture; after the training phase, model output feedback is used in replacement of the target series in the parallel architecture shown. Architecture change calls for further generalisation performance assessment to discard possible target memorisation during the training phase.

phase. The optimum trained NARX model for each hemisphere is used to reconstruct the respective polar field. In this implementation phase, the target input $y(t)$ is substituted by $\hat{y}(t)$ transforming the series-parallel architecture (open loop) into the parallel architecture shown in Figure 8. The regular sunspot area record provides sufficient exogenous input series to reconstruct polar field from the Schwabe cycle 12 to 20, and to forecast their values up to the next sunspot minimum. As mentioned before, this last feature is possible due to the existing phase shift between the sunspot area record and polar field. In Figure 9 the exogenous input series, target series, estimated series, and data segmentation for the training and best model identification phases are identified.

The reconstructed and forecasted polar field time series are used to establish the poloidal potential modulation in the dynamo model to reproduce observed solar activity. It is convenient at this point to bear in mind that the poloidal potential modulation is effective only at the start of each simulated solar cycle, which implies natural poloidal potential evolution along the simulated solar cycle. In this way the polar field time series are sampled at SSN minima to identify the initial poloidal potential modulation for each cycle. With the exception of the starting date of the forthcoming solar cycle 25, all other dates are established. Thus a short-term prediction to estimate the start of the solar cycle 25 is required. To this end, a standard autoregressive model of degree 9 (Werner 2012) is used on the monthly smoothed SSN as shown in Figure 10. The model short-term prediction sets the start of the solar cycle 25 around the second half of 2018⁹. In Figure 11 the magnitude of the full solar polar field series from the NARX reconstruction and forecasts, as well as from the Wilcox Solar Observatory record are presented. Squares in the figure indicate the solar polar field values considered for subsequent solar dynamo input (at the beginning of each solar cycle). From Figure 11 it is observed that the NARX model methodology succeeded in obtaining a nonlinear dynamic system embodying long-term characteristics of the complex processes relating sunspot areas and polar field strengths. However cumulative short-term variations reducing polar field strength representativeness of past solar activity, such as active regions helicity hemispheric compatibility and irregular unipolar poleward migra-

tion velocity (McMaken & Petrie 2017), are arguably an important source of degraded NARX model performance. An immediate apparent discrepancy observed in the NARX curves is the expected maximum of polar field magnitude at the start of the Schwabe cycles, which is true during the observed magnetic field window in the figure (WSO data). It is essential to acknowledge that the NARX curves are by no means accurate representations of polar field strength history but coarse estimations derived from a small portion of the observed Sun. However, the NARX reconstruction provides interesting clues of possible past behaviour challenging the generalisation of the aforementioned expected trends. In this regard, the consolidated polar faculae (PF) database reported by Muñoz-Jaramillo et al. (2013) provides supporting evidence to the validity of the NARX reconstruction. For instance, the times of polar faculae maxima (marked as triangles in Figure 11) although similar to the times of WSO polar field maxima and in agreement to those of SSN minima, show a local maximum near 1960 differing to the reported start of the solar cycle 20 that took place in 1964; this time of polar faculae maximum and the respective maximum from the NARX reconstruction are similar as shown in Figure 11. Acknowledging the possible existence of this behaviour, all sampled reference magnetic states used in the solar dynamo correspond to the start of each Schwabe cycle, independently of the shape of the polar field curves. Hence, the true representativeness of the sampled estimated polar field strength should be assessed by the solar dynamo-NARX combined efficacy in reproducing solar activity.

A possible way to improve the presented NARX model is by incorporating information from other recorded solar activity proxies, e.g. the record of polar faculae, to provide a broader and richer set of scenarios during the development of the model. However, in this investigation the solar cycles simulation presented in the following subsection 3.2, is carried out with the models developed in this subsection using the hemispheric polar magnetic field time series alone.

3.2 Solar Cycles Simulation and Forecast

Solar activity dynamo simulations of the cycles 12–25 are presented in Figure 12. The theoretical magnetic and sunspot butterfly diagrams are shown in Figure 12a, and the pseudo-SSN (solid curves) are superimposed on the observed 13-month smoothed monthly SSN (light areas) in Figure 12b. Pseudo-SSN are obtained from

⁹ In close agreement to the SILSO⁶ forecasts (Kalman-filter optimised forecasts ML and SC).

Figure 9. Time series employed for the southern polar field short-term forecast. In this example all relevant time series are superimposed for illustrative purposes only. Phase shift between observed hemispheric sunspot areas (exogenous inputs) and polar field strength (target) is cancelled during computations. Multiple NARX models are trained to finally select one able to generalise best during the evaluation phase. The magnified area shows the target series segment $y(t)$ and the estimated value of the polar field $\hat{y}(t)$. The best NARX model is expected to represent a nonlinear dynamic system relating exogenous input series and target series. In total, a polar magnetic strength forecast and a reconstruction is obtained for each hemisphere.

Figure 10. Start of the solar cycle 25 prediction using an autoregressive model of degree 9 on the monthly smooth SSN.

the normalised and scaled chronological sum of emerging active regions in Figure 12a, with respect to the reference observed solar cycle 23. Note that during the declining phase, the theoretical curves overestimate the observed SSN. This is an expected behaviour as the solar dynamo model gradually adjusts itself from the imposed initial poloidal potential magnitudes and asymmetries to bring back a subsequent average solar cycle. The theoretical curves, dashed lines in Figure 12b, show the direct output (raw) from the solar dynamo model, whereas the solid curves show the corrected theoretical curves with the process detailed below.

Raw data in Figure 12b shows important characteristics of the dynamo model and its potential use for forecast application. Firstly, theoretical and observed SSN maxima of cycle 23 must match since this cycle serves as data fulcrum to relate theory with observations as discussed earlier in this section 3 (refer to Figures 5-6). Other theoretical and observed maxima show close match such as in cycles 12, 17, 20, and 22; the rest show various levels of differences

with the highest found in cycle 19. From these results it is evident that the solar dynamo model produces results with limited sensitivity to the simulation initial state, i.e. approximately ± 50 SSN with respect to the initial observed reference solar cycle. For example, dynamo input data (solar polar field) from the observed solar cycle 20 produced underestimated pseudo-SSN activity (theoretical curve in the figure) of the cycle 21, albeit with the correct increasing trend. Due to this model limitation simulation results are assessed in pairs, that is in trend transitions from the initial reference observed cycle to the subsequent simulated cycle, rather than direct SSN cycles comparison. Figure 13 summarises maxima forward percentage trends between the theoretical and observed SSN. From the theoretical points in Figure 13 is seen that the transitions $17 \rightarrow 18$, and $21 \rightarrow 22$ are discordant to the observed trends. Or in other words, if forecast had been the case, simulation of cycles 18 and 22 would have lead to incorrect predicted trends with respect to cycles 17 and 21 respectively.

External investigations of solar cycle activity reproduction using high diffusivity mean field axisymmetric dynamo models, show similar performances to those presented in this investigation. In a first example, in Karak's (Karak 2010) study the meridional circulation is adjusted to match observed SSN periods. This parameter variation modified the average dynamo model response producing various levels of solar activity. That approach loosely approximated the magnitude of some solar cycles, with maximum difference found in the solar cycle 19. Regarding the trends between solar cycles and their precursors, cycles 15, 18, and 20 are discordant to observations in Karak's study. In another example, Jiang et al. (2007) report simulations using the observed photospheric magnetic field as cycles' adjusting parameter, thereby encompassing cycle 21 onwards. In Jiang et al. (2007) study, theoretical curves of the cycles 21, 22, 23 show close match to the observed SSN. Additionally a forecast for the cycle 24 is reported that is now known to have successfully anticipated peak observed activity. However, regarding the trends between solar cycles and their precursors, cycle 21 is discordant in that study. At this point is important to bear in mind that the photo-

Figure 11. Full solar polar field series from the NARX reconstruction and forecasts, and from the Wilcox Solar Observatory record. Triangles show dates of reported north-south maxima average polar field strength magnitude according to Muñoz-Jaramillo et al. (2012), derived from a consolidated polar faculae database. Dates of peak values from the NARX reconstruction are similar to those from the polar faculae database supporting the reconstruction validity in this respect.

spheric magnetic field precursor approach by Jiang et al. (2007), is used as reference study in this investigation as mentioned in section 1. Similarities in the approaches of both investigations allow direct comparison of the raw results for the cycles 21 – 24 in Figure 12b. Interestingly, the discordant cycle 21 → 22 trend and correct reproduction of the solar cycle 24 are observed in both cases. The ability to compare further solar cycles in this investigation, is enabled by the proposed polar field reconstruction of past solar cycles. The validity of the reconstructed polar field is tested in their capability to furnish adequate initial states for the solar dynamo to provide a suitable representation of cycle features. Nonetheless, the solar dynamo model as well as the reconstructed polar field add undifferentiated uncertainty components to the final solar activity simulation. Thus, theoretical cycles in this investigation report combined limitations and strengths of the solar dynamo model, the reconstructed polar field, and simulation assumptions and methodology. Comparison amongst observed and theoretical cycles in the range 12 – 20 in Figure 12b, highlights the potential of the proposed approach.

The Karak (2010) study provides an additional example of solar dynamo model performance. The range of interest for comparison in this case are cycles 12 – 20. Comparison of results between Karak’s study and the raw results in Figure 12b show that reconstructed polar field yield better solar cycles reproduction than those generated using observed SSN periods as the adjusting dynamo parameter. Karak’s study shows that meridional flow variation is a fundamental ingredient in solar activity simulation. This is a reasonable argument considering that the main relationship between solar activity and meridional flow may be explained by the interaction of equatorward meridional flow at the base of the convection zone. In that possible explanation, plasma flow may encounter motion restriction imposed by magnetic tension found in highly active zones; meridional flow is quenched by solar activity. This suggests that a more robust solar dynamo model should include meridional circulation fluctuations along the solar cycles simulation. In this regard, a deeper understanding about the meridional flow structure would provide crucial information to improve current solar dynamo models.

Reproduction of past solar cycles is fundamental to the solar

dynamo validation process, which in turn improves its reliability for adequate forecasts. The analysis presented above shows that high diffusivity mean field axisymmetric dynamo models do not capture completely the whole range of observed solar activity, at least with the methodologies discussed above. A plausible explanation to the theoretical mismatch found in this investigation, is that the initial adjustment of the dynamo model includes information from only one component of the magnetic field, i.e. poloidal modulation via observed solar polar magnetic field. In this regard, simulations point to the possible need to include additional information about the state of the toroidal component. Expanding on this possibility, take the theoretical solar cycle 19 in Figure 12b as an example. The initial adjustment of the poloidal potential in the dynamo model introduces partial information of the previous cycle 18 because, without modification, the toroidal field strength corresponds to that of an average solar cycle. With the progress of the simulated cycle, diffusion and the meridional flow gradually transports the modified poloidal field to the tachocline region where it is stretched in the azimuthal direction by the differential rotation. The preexisting and newly produced toroidal field generate emerging active regions resulting in a small increase of simulated solar activity, i.e. the solar cycle 19. If on the other hand, both fields are adjusted, the newly created toroidal field would find a suitable magnitude of the preexisting toroidal field to effectively produce the observed activity. The challenge in this case, however, is the identification of an observed proxy of the toroidal field state.

The rise times of solar cycles may give qualitative information of the toroidal field state. This hypothesis is based on an extreme case wherein the preexisting saturated toroidal field at the beginning of the solar cycle, triggers large flux tube buoyant sections that rise rapidly to the solar surface. In contrast to the poloidal field, this toroidal field variation is slow, possibly sufficient to retain its bulk characteristics even after maximum solar activity. If this is true, then a given solar cycle rise time and the polar magnetic field at the end of that cycle, may contain sufficient information to adjust the solar dynamo model for the simulation of the subsequent solar cycle. In order to capture the qualitative state of the toroidal field into a quantitative correction factor, Eq. 9 is used wherein t_i , t_f ,

Figure 12. Solar activity simulation results. The theoretical magnetic and sunspot butterfly diagrams are shown in panel (a). Panel (b) shows the raw (dashed curves) and corrected pseudo-SSN (solid curves), superimposed on the observed 13-month smoothed monthly SSN⁶ (light areas). Of interest in this investigation are sunspot cycles *activity trends* between the simulated and the precursory cycle.

Figure 13. Solar maxima forward change percentage trends from the raw theoretical and observed SSN. Discordant pairs for a given solar cycle have opposite signs, whereas concordant pairs exhibit different levels of agreement. The highest level of agreement (overlapping pair) is observed in the solar cycle 23 because the theoretical response is calibrated on this cycle.

and t_{max} are solar cycles times at the start, end, and activity peak respectively.

$$C = \frac{t_f - t_{max}}{t_f - t_i} \quad (9)$$

The correction is carried out in cycle pairs wherein the cycle $(n-1)$ provides the factor C for the cycle n . For example, to preserve the magnitude of the reference solar cycle 23, the correction vector is normalised to the cycle 22 as this sets the initial conditions of the cycle 23. The normalised correction vector is then multiplied by the raw theoretical curves shown in Figure 12b. The corrected theoretical cycles are presented in the same figure as solid curves. This modification results in a dramatic improvement on the magnitudes

and trends of simulated cycles with respect to observed activity. Figure 14, shows the corrected trends of the simulated cycles. SSN percentage change error in solar cycle 22 is reduced in comparison to the same cycle in Figure 13, and the discordant trend in cycle 18 disappears. Nonetheless, two new discordant trends appear in cycles 13 and 14. Interestingly, the erroneous cycle transitions, i.e. $12 \rightarrow 13$ and $13 \rightarrow 14$, coincide with the Maunder's original butterfly diagrams published in 1904 covering the period 1876 – 1902 (Maunder 1904). In Maunder's original work, butterfly diagrams were built with data of sunspots permanence at a given latitude. Wherever a sunspot centre appeared on the solar surface, and remained for more than one day during a given solar rotation, the latitude was marked in a latitude-time chart. Sunspot area information was added later with the assembly of sunspot observations as formats changed along the years. Recent evaluation of early datasets has identified non-uniform sunspot area data⁵. In an effort to regularise datasets, these were compared with the Mount Wilson photographic plate collection dating from 1917 to 1941⁵. All this may suggest that the anomalous transition trends between theoretical and observed data for cycles 13 and 14 in Figure 14, may be a result of non-uniform or erroneously formatted sunspot area data.

The solar dynamo model's adequate reproduction of solar peak activity and solar maxima forward change percentage trends of past solar cycles, allows a forecast for the solar cycle 25 to be proposed. A weak solar cycle 25, with slow rise time and maximum activity -14.4% ($\pm 19.5\%$) with respect to the current cycle 24 is predicted.

4 CONCLUSIONS

In this paper we have analysed the capability of a mean field axisymmetric dynamo model with a manipulated magnetic field to

Figure 14. Same as Figure 13 with corrected theoretical SSN. Each cycle correction factor improves the level of agreement observed in the cycle pairs. In addition, possible non-uniform or faulty formatted data in the early cycles record is disclosed.

reproduce observed solar activity. In general, current solar activity reproduction using solar dynamo models with a poloidal field precursor is limited to a few solar cycles. Aiming at extending the verification of the solar dynamo model with a broader number of solar activity scenarios, unrecorded solar polar magnetic field was reconstructed. Reconstruction of the solar polar magnetic field was carried out using the hemispheric sunspot areas record in conjunction with artificial neural network methods. As a result, the extended solar dynamo model testing encompassed thirteen solar cycles, rather than the conventional observed four cycles.

Simulations of 13 solar cycles with the proposed diffusion-dominated solar dynamo approach, showed satisfactory reproduction of SSN peak trends for 9 of 12 solar cycles. In addition, the extended verification permitted the identification of a correction factor to improve solar dynamo simulations. Further, possible anomalous data in the historical hemispheric sunspot area are identified in early solar cycle reports. Corrected simulated activity for the forthcoming solar cycle 25 predicts weaker peak activity than observed in the current cycle.

Solar activity simulations presented in this investigation should be treated as rough approximations as many poorly understood physical processes controlling the actual solar cycle are represented in the model. Nonetheless, the reported combination of the consolidated solar polar magnetic field strengths, mean field axisymmetric dynamo model parameterisation approach, and calibration methodology, proved to be sufficiently effective in reproducing various historical solar activity scenarios.

REFERENCES

- Arlt R., 2009, *SoPh*, 255
 Balbus S. A., Latter H., Weiss N., 2012, *MNRAS*, 420, 2457
 Cardoso E., Lopes I., 2012, *ApJ*, 757, 71
 Charbonneau P., Christensen-Dalsgaard J., Henning R., Larsen R., Schou J., Thompson M., Tomczyk S., 1999, *ApJ*, 527, 445
 Chatterjee P., Nandy D., Choudhuri A. R., 2004, *A&A*, 427, 1019
 Choudhuri A., 2005, *The User's Guide to the Solar Dynamo Code Surya*
 Demuth H. B., Beale M. H., De Jess O., Hagan M. T., 2014, *Neural network design*. Martin Hagan
 Dikpati M., 2008, *AnGeo*, 26, 259
 Dikpati M., Gilman P. A., 2006, *ApJ*, 649, 498
 Graupe D., 2007, *Principles of artificial neural networks*. World Scientific
 Hathaway D. H., Upton L., 2014, *JGRA*, 119, 3316

Figure A1. A snapshot of the supplementary video. Comparison of simulated toroidal field morphology evolution using standard and the proposed solar parameterisations.

- Hathaway D. H., Nandy D., Wilson R. M., Reichmann E. J., 2003, *ApJ*, 589, 665
 Hazra G., Karak B. B., Choudhuri A. R., 2014, *ApJ*, 782, 93
 Hotta H., Yokoyama T., 2010, *ApJ*, 714, L308
 Jiang J., Chatterjee P., Choudhuri A. R., 2007, *MNRAS*, 381, 1527
 Karak B. B., 2010, *ApJ*, 724, 1021
 Kitiashvili I., Kosovichev A., 2008, *ApJ*, 688, L49
 Li M., Mehrotra K., Mohan C. K., Ranka S., 1990
 Maunder E. W., 1904, *MNRAS*, 64, 747
 McMaken T. C., Petrie G. J. D., 2017, *ApJ*, 840, 100
 Muñoz-Jaramillo A., Nandy D., Martens P. C., 2009, *ApJ*, 698, 461
 Muñoz-Jaramillo A., Nandy D., Martens P. C., 2010, *ApJ*, 727, L23
 Muñoz-Jaramillo A., Sheeley N. R., Zhang J., DeLuca E. E., 2012, *ApJ*, 753, 146
 Muñoz-Jaramillo A., Dasi-Espuig M., Balmaceda L. A., DeLuca E. E., 2013, *ApJ*, 767, L25
 Petrie G., 2013, *ApJ*, 768, 162
 Priest E., 2014, *Magnetohydrodynamics of the Sun*. Cambridge University Press
 Sokoloff D., Arlt R., Moss D., Saar S. H., Usoskin I., 2009, *Proceedings of the International Astronomical Union*, 5, 111
 Svalgaard L., Duvall T. L., Scherrer P. H., 1978, *SoPh*, 58, 225
 Tlatov A. G., Pevtsov A. A., 2014, *SoPh*, 289, 1143
 Werner R., 2012, *SunGe*, 7, 75

APPENDIX A: SUPPLEMENTARY MATERIAL

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.