



A solar cycle 25 prediction based on 4D-var data assimilation approach

Allan Sacha Brun¹ , Ching Pui Hung^{1,2}, Alexandre Fournier²,
Laurène Jouve³, Olivier Talagrand⁴,
Antoine Strugarek¹  and Soumitra Hazra¹

¹DAP/AIM, CEA Paris-Saclay, 91191 Gif-sur-Yvette, France

²IPGP, Université de Paris, UMR 7154 CNRS, F-75005 Paris, France

³Université de Toulouse, UPS-OMP, IRAP, 31028 Toulouse Cedex 4, France

⁴LMD, UMR 8539, Ecole Normale Supérieure, Paris Cedex 05, France

Abstract. Based on our modern 4D-var data assimilation pipeline *Solar Predict* we present in this short proceeding paper our prediction for the next solar cycle 25. As requested by the Solar Cycle 25 panel call issued on January 2019 by NOAA/SWPC and NASA, we predict the timing of next minimum and maximum as well as their amplitude. Our results are the following: the minimum should have occurred within the first semester of year 2019. The maximum should occur in year 2024.4 ± 6 months, with a value of the sunspot number equal to 92 ± 10 . This is in agreement with the NOAA/NASA consensus published in April 2019. Note that our prediction errors are based on $1-\sigma$ measure and do not consider all the systematics, so they are likely underestimated. We will update our prediction and error analysis regularly as more data becomes available and we improve our prediction pipeline.

Keywords. Sun, dynamo, 11-yr cycle, solar cycle prediction, data analysis, data assimilation

1. Introduction

The Sun possesses an intense surface activity modulated by its 11-yr magnetic dynamo cycle (Brun & Browning 2017). Over the last four centuries it has become clear that the period of the so-called Hale solar cycle is not perfectly stable, varying between 9 and 13 years typically (Clette & Lefèvre 2012). Likewise its amplitude has varied significantly, from being weak or even null (grand minima phase) to being very strong as in cycle 19. Hence it has become crucial to be able to predict the solar activity in order to anticipate and ideally mitigate the impact of our fierce Sun and its highly variable activity. It is of course very difficult to predict the solar activity cycle given the high degree of nonlinearity of the solar dynamo. Still there seems to be some order in this otherwise chaotic behavior and we can attempt to capture it to the best we can. To this end several groups have developed various ways of predicting the strength and timing of the next solar cycle, see for instance (Hathaway 2015) and (Petrovay 2019) for recent updates and summaries. Here we briefly present our own solar cycle predicting pipeline based on a novel 4-D var data assimilation method (Talagrand 2010), coupled to a 2.5D mean field dynamo model (see details in Jouve *et al.* 2011; Hung *et al.* 2015, 2017). In this short paper we present a summary of our answer to the Solar Cycle 25 - Call for Predictions issued in January 2019 by the NOAA Space Weather Prediction Center and NASA.

In §2 we briefly present the methodology behind our solar cycle prediction 4D-var tool *Solar Predict*. In §3 we present the solar data used to perform our prediction for solar cycle 25 and the *hindcasting* of cycles 22, 23 and 24. In §4 we validate our procedure using

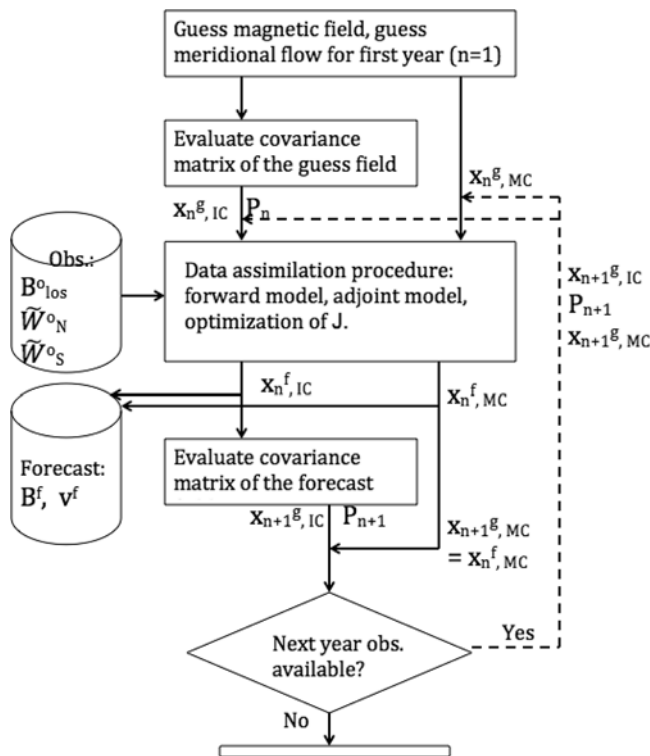


Figure 1. A schematic diagram illustrating data assimilation procedure. Integer subscripts refer to discrete time indices.

solar data from past solar cycles in order to test our prediction skills before performing our data assimilation (DA) analysis for cycle 25 and discussing our results in §5.

2. Brief presentation of the *Solar Predict* Data Assimilation pipeline

Our *Solar Predict* tool developed and used to perform inversion and prediction of the solar dynamics and cycle is based on a 4-D Var data assimilation (DA) method (Talagrand 2010), coupling sunspot number (SSN) and butterfly diagram (e.g. temporally stacked $B_{los}(\theta, t)$ time series data within a 2.5-D mean field axisymmetric dynamo model (depending on radius r and colatitude θ , but solving for all 3 components (B_r, B_θ, B_ϕ) of the magnetic field). This versatile tool can assimilate various length of solar data time series and invert different physical ingredients using our DA algorithm. Here, we focus on predicting solar cycle 25, using solar cycle 24 data (so-called "recent climatology"). An extensive description of *Solar Predict* can be found in (Jouve et al. 2011; Hung et al. 2015, 2017) and an image summarizing the pipeline is shown in Figure 1.

2.1. Observational Proxy

In order to make our prediction we need a model to generate time series of physical variables that can be directly compared to the observations. To do so, we use a mean field Babcock-Leighton dynamo model as described in (Jouve & Brun 2007; Sanchez et al. 2014; Hung et al. 2017), with small modifications to the parameters and with a slightly more complex resistivity profile, with a 2-step profile in radial direction.

Since the model does not produce sunspots per se, we introduce a proxy for the total sunspot number (SSN^f) of the model, in the form of a pseudo-Wolf number \tilde{W}^f defined by

$$\tilde{W}^f(t) = \int_0^\pi \int_{0.7}^{0.71} [B_\phi^f(r, \theta, t)]^2 r^2 \sin \theta \, dr \, d\theta. \tag{2.1}$$

In radius, the integral is restricted to a thin layer (between 0.70 and 0.71 solar radius) where toroidal flux tubes are thought to originate. We further multiply \tilde{W}^f by a constant c_{SSN} that allows to adjust the amplitude of \tilde{W}^f within the range of values of real solar sunspot number record, e.g. $SSN^f = c_{SSN} * \tilde{W}^f$ (cf. Fig. 2).

If we wish to capture the north-south asymmetry, we can further decompose \tilde{W}^f into its north and south components

$$\tilde{W}^f(t) = \tilde{W}_N^f(t) + \tilde{W}_S^f(t), \tag{2.2}$$

in which the north (resp. south) component \tilde{W}_N^f (resp. \tilde{W}_S^f) is computed by restricting the integration in Eq. 2.1 to the northern (resp. southern) hemisphere.

The other class of data will consist of time series of the line-of-sight component of the magnetic field at the model surface for all co-latitudes θ , $B(\theta, t)_{los}^f$, defined as

$$B_{los}^f(\theta, t) = B_r^f(r = 1, \theta, t) \sin \theta = (\cos \theta + \sin \theta \partial_\theta) A_\phi^f(r = 1, \theta, t). \tag{2.3}$$

assuming $r = 1$ at the surface. Such data will be directly compared to solar butterfly-like data. The solar data used to perform our prediction is discussed in the next section.

2.2. Objective function \mathcal{J}

To successfully perform our prediction, we aim to minimize an objective function defined in terms of the differences between the observations and our dynamo model trajectory,

$$\mathcal{J} = \sum_{t_i}^{N_t} \left\{ \sum_{\theta_j}^{N_\theta} \frac{(B_{i,j}^f - B_{i,j}^o)^2}{x / \sin^2 \theta} + \frac{(SSN_i^f - SSN_i^o)^2}{\sigma_{SSN,i}^2} \right\}, \tag{2.4}$$

where \mathcal{J} is the objective function, B is the surface radial field, SSN is the sunspot number (see equations (2.2) and (2.3) for a description of how we compute these quantities from our dynamo model). Superscripts o and f denote observed and forecast (model-based) values respectively. The misfit of the surface radial field is normalized with $1/\sin^2 \theta$, as the uncertainty of observation increases with latitude. The relative weighting of misfit in the surface field and SSN is controlled by the factor x . The misfit in SSN is normalized with the variance of SSN^o . The misfits are summed over the number of observations in time (N_t) and latitudes (N_θ , for the surface field).

We then assimilate the magnetic observations for the n^{th} sunspot cycle, and get an estimate of the average flow and the magnetic configuration on the meridional plane, at the end of the cycle. Finally, we can obtain a preliminary guess of the maximum of the $n + 1^{th}$ cycle by extrapolation of the model beyond the assimilation window, based on the forecast flow and magnetic configuration estimated from the observations of the n^{th} cycle. The first maximum of the extrapolated dynamic trajectory of the modeled SSN is our guess of the $n + 1^{th}$ maximum.

As mentioned above, the observations used for the experiment are compared to our magnetic sunspot proxy (Equation 2.1) and with the surface line of sight magnetic field B_{los}^o (Equation 2.3).

Table 1. Gussed maxima of the $(n + 1)^{th}$ cycle by extrapolation of the forecast model, from the results of assimilation of the observations of the n^{th} cycle.

x	$n = 21$		$n = 22$		$n = 23$	
	$t_{max,n+1}$	$SSN_{max,n+1}^f$	$t_{max,n+1}$	$SSN_{max,n+1}^f$	$t_{max,n+1}$	$SSN_{max,n+1}^f$
0.1	1990.8	190	2000.7	208	2012.5	135
1.0	1991.1	222	2000.7	239	2013.0	115
5.0	1991.4	202	2001.1	211	2013.5	90.4
10	1991.5	192	2001.2	196	2013.8	76.2
50	1991.7	174	2001.7	152	2014.1	68.9
observed	1990.5	212 ± 11	2001.1	177 ± 9	2013.7	106 ± 8

Table 2. Predicted minima between cycle 24 and 25 by extrapolation of the forecast model, from the results of assimilation of the observations of cycle 24.

minimum $n = 24/25$ forecast	
x	$t_{min,24-25}$
0.1	2019.09 ± 0.10
0.3	2019.16 ± 0.10
0.5	2019.1 ± 0.10
1	2019.08 ± 0.10
10	2019.12 ± 0.10

Table 3. Predicted maxima of the cycle 25 by extrapolation of the forecast model, from the results of assimilation of the observations of cycle 24.

maximum cycle $n = 25$ forecast		
x	$t_{max,25}$	$SSN_{max,25}^f$
0.1	2024.3 ± 0.40	97 ± 6
0.3	2024.3 ± 0.40	97 ± 6
0.5	2024.3 ± 0.40	96 ± 6
1.0	2024.3 ± 0.35	94 ± 6
10	2024.5 ± 0.40	88 ± 6

2.3. Overall Data assimilation procedure

In Figure 1 we represent the overall data assimilation procedure. We use both direct and adjoint (tangent linear) 2.5D mean field dynamo models to assess the variation (gradients) of the objective function \mathcal{J} (again see details in: Jouve *et al.* (2011); Hung *et al.* (2015, 2017)). Thanks to the coupling of the direct and adjoint dynamo models in our DA pipeline we are able to efficiently minimize \mathcal{J} . This results on the misfit between the observations and the model observation proxies to tend to zero. We perform this over about one solar cycle prior to letting the forecast (e.g. the model with observationally constrained magnetic field configurations for $A_\phi(\vec{r}, t)$ and $B_\phi(\vec{r}, t)$ and the meridional circulation) evolve freely. We repeat this procedure for as many x control parameters as needed (typically 5 or 6 values). We then systematically perform stochastic perturbations of each average trajectories to further assess the error bars of the predictions (see Tables 2 & 3). So overall our procedure is divided in two steps: a) data assimilation of the existing solar data over about one solar cycle to obtain good initial magnetic and meridional states to be used in step b) as initial conditions of the dynamo model that we let go unconstrained. The trajectory then obtained in step b) constitutes a prediction of the magnetic state of the Sun.

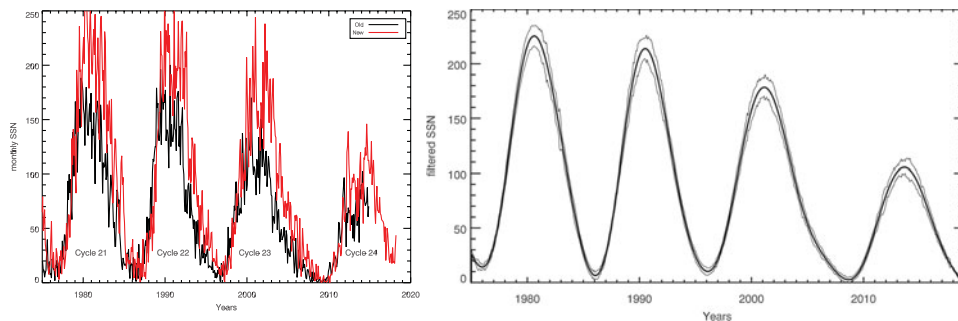


Figure 2. (Left) Monthly averaged SSN time series (old and new in red Clette & Lefèvre (2018)) from SIDC/SILSO data base and (Right) filtered version SSN^o time series used to perform our data assimilation procedure. On the filtered curve the $\pm 1\text{-}\sigma$ curve are being shown as well.

3. Solar data set used

We base our prediction on two types of solar data a) sunspot number time series and b) butterfly diagram of solar activity (see for instance Hathaway (2015)), from which we derive the line of sight magnetic field B_{los}^o magnetic field. Historically, sunspot series have started in 1749 and for instance are available at the Solar Influences Data Analysis Center (SIDC) <http://www.sidc.be/silso/datafiles>. As indicated in the NOAA/NASA call of January 2019, there are 2 time sunspot number series (SSN), the old and the revised ones (Clette & Lefèvre 2018). For the predictions presented in this note we will make use of the new SSN time series up to December 2018. We start from the monthly smoothed one as shown on Figure 2 and apply a filter on that data (see below).

Daily magnetograms of the surface magnetic field of the Sun have been available since the 2nd half of the 20th century thanks to facilities such as Kitt Peak and Wilcox observatories or more recently from space probes such as SoHO or SDO or the ground network GONG. There are now easily accessible via the NSO web site: <https://www.nso.edu>.

To create the time-latitude butterfly diagram used in our assimilation and prediction pipeline, we start from synoptic maps also provided as a “by-product” on NSO website. After applying an azimuthal average on the synoptic maps for every single Carrington Rotation (CR) available in the data bases, we generate one time snapshot “slice” as a function of latitude (or sine of latitude) and stack them in time. This procedure is similar and inspired by the one used by Dr. David Hathaway to generate his blue-yellow butterfly diagram (<http://solarcyclescience.com/solarcycle.html>). We can generate the butterfly diagram with any source of synoptic maps data, including synchronous maps if necessary.

The magnetic observations used to make the butterfly diagram shown on Figure 3 are obtained from Kitt Peak (KPVT), SOHO (MDI), GONG and SOLIS 1-degree synoptic maps data from 1976 up to December 2018. For the most recent data we use the NISP (NSO Integrated Synoptic Program) web server and use SOLIS maps when available and replace them by MDI or GONG maps when necessary. We use Carrington Rotation synoptic maps from CR 1625 until CR 2211, hence covering the last 4 solar cycles or so.

We further processed the observations with a low-pass Butterworth filter of order 4 with the cut-off frequency $1/T$ set at $T = 10$ years for the surface radial field and project the latitudinal spatial structures into Legendre polynomial $P_\ell(\cos \theta)$, using a cut-off of $\ell_{max} = 10$. For the SSN time series we also apply a filter (using again a 4th order Butterworth filter) whose cut-off is at $T = 5$ years. The latest data point includes observations up to December 2018. We display the original and filtered solar data used in this study in Figures 2 and 3.

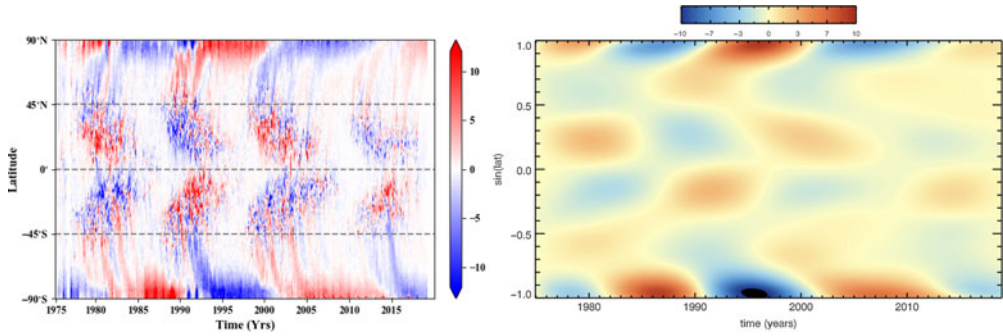


Figure 3. (Left) Full resolution butterfly diagram using Carrington Rotation synoptic maps (see text for details). The color contour plots are scaled between +/- 10 Gauss. (Right) Filtered version $B_{los}^o(\theta, t)$ used in the DA pipeline.

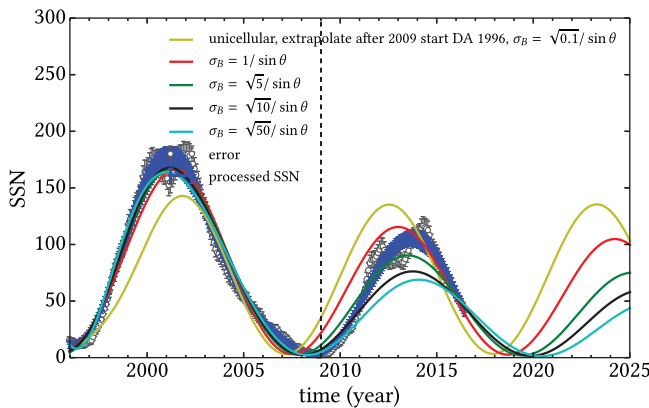


Figure 4. An example of hindcasting for cycle 23: Data assimilation of the magnetic observations for the whole cycle 23, and extrapolation of the modeled SSN based on the estimate of the average flow (recall that one of our control parameter is the meridional circulation amplitude and profile). We conduct the tests with different weighting factor x , namely, $x = 0.1$ (brown), 1 (red), 5 (green), 10, (black) and 50 (light blue) (see objective function \mathcal{J} in eq. 2.4 for the definition of x). The 13-month mean and (low pass) filtered observations are shown with black and blue error bars, respectively.

4. Hindcasting of the previous solar cycles 22, 23 and 24

In order to validate our prediction of the next solar cycle 25, we have used our pipeline on cycles 22, 23 and 24, using real solar data from the previous cycles 21, 22, and 23 (see section 3 for details and Figures 2 and 3). This has allowed us to assess the best range of parameters in our 4D-Var assimilation procedure and the accuracy of our prediction and its associated 1- σ error bars. These parameters and error bar information are then used in the next section 5 to perform our solar cycle 25 prediction. We display in Figure 4 one realization of our *hindcasting* validation procedure on cycle 23 and provide in table 1 a summary for all three past cycles. On Figure 4, we note that the model tracks the solar data very well in the data assimilation part, here from 1996 until 2009. Various trajectories of the *Solar Predict* pipeline are being plotted against real solar data. The parameter x has been varied such that we give more or less weight to the SSN time series over the butterfly diagram. All models except the one represented with a mustard curve track are within the observation error bars. In the extrapolation part beyond 2009, for which the *Solar Predict* pipeline was run in its prediction mode as if there were no solar

data available, we can assess how the various trajectories evolve in time. Since solar data for cycle 24 is mostly available, we can directly compare these extrapolated trajectories with reality. We note that our set of trajectories contain the trajectory that the Sun really evolved to after 2009. This gives us confidence that our *Solar Predict* pipeline is able to guess educate the near future of the Sun's magnetic state. We further see on table 1 that the range of the x parameter between 0.1 and 10 generally makes a good job at predicting the next solar cycle timing and amplitude. For each of the three test cycles (22, 23, 24) we are able to predict the timing and the amplitude of the cycle within the observations error bars. Hence, in the next section we will use this range of value of x to bracket the future magnetic state of the Sun.

5. Equatorially Symmetric Prediction of Solar Cycle 25

In this section we present in detail our prediction for the next solar cycle 25. We first consider the equatorially symmetric case.

5.1. Next Minimum

On Table 2, we provide our prediction for the current minimum between cycle 24 and 25. As we can see our model predicts that we were close to reaching the minimum in the late part of the first semester of 2019. This minima seems to have extended into fall 2019 but did not deepen further. Recall that this prediction used data up to December 2018, not November 2019, 11 more months of data could of course modify our prediction.

5.2. Next Maximum

On Table 3, we provide our prediction (as of January 2019) for the next solar cycle 25 maximum. As we can see our model predicts that we will reach the next maximum in the middle of 2024 between the first trimester and the 3rd and that cycle 25 should be comparable in amplitude with cycle 24. We provide 1- σ error bar and hence our timing for the next cycle seems quite precise with only a 6 months window. This may be a bit too optimistic with respect to the real precision of our pipeline and future work is needed to assess further the precision and accuracy of our solar cycle prediction.

Our SSN proxy predicts mean values for the sunspot number ranging between 88 and 97. If we further take into account ensemble forecasting (based on stochastic perturbations of the trajectories) and systematic error bars, we predict a sunspot number for cycle 25 as low as 82 or as high as 103.

We also display on Figure 5 our various realizations and associated time series, for which we have changed the parameter x that controls the relative weighting of the SSN vs the butterfly diagram data set while assimilating the data and minimizing the objective function \mathcal{J} . We see that during the time period when we have data of solar cycle 24 available, the model curves follow closely the SSN time series and that even if we vary x to have less weight on the SSN data series and instead favor the butterfly diagram one, this does not make a large difference. Indeed all curves track the data well, as expected from advanced DA procedure as soon as the objective function \mathcal{J} is kept small.

We also note that our DA pipeline indicates a maximum horizon of predictability of one solar cycle. Making a prediction from minimum state is empirically more favorable than near the maximum. This feature is well known in the solar forecasting community (Cameron & Schüssler 2008). By starting from a state close to the minimum of cycle 24, our horizon of predictability for the next cycle reaches about one cycle period. However, near the maximum this horizon is shortened to less than a cycle period. The initial state and the control of the growth of error is key in performing predictions (Sanchez *et al.* 2014).

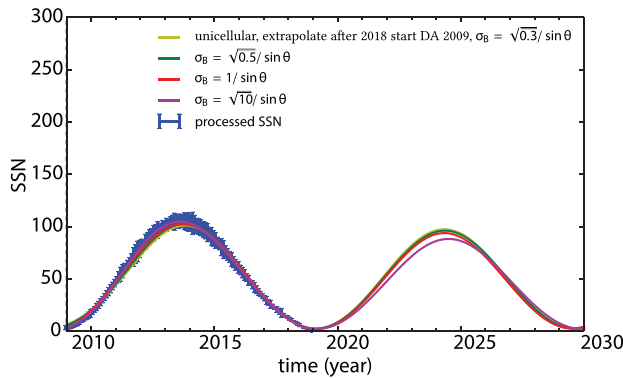


Figure 5. An example of forecasting for cycle 25: Extrapolation beyond December 2018 from the estimated average unicellular flow and magnetic fields, from the forecast based on assimilation of magnetic observations in cycle 24. The magnitude of scaling factor of the misfit in surface radial field is the parameters to be studied, with brown, green, red and violet curves for $x = 0.3; 0.5, 1.0$ and 10.0 , respectively (see objective function \mathcal{J} in eq. 2.4). The observed monthly mean SSN values (to which we applied a 5-yr filtered) are shown with blue error bars.

5.3. Discussion of our predictions and caveats

Since our prediction was submitted to the NOAA/NASA cycle 25 panel in early February 2019, the panel has published a consensus prediction (see for instance [Weber et al. 2019](#), contribution in this conference proceedings). The NOAA/NASA consensus (version 2 as of April 2019) is: a SSN between 95 and 130 and a maximum reached between 2023 and 2026. So our prediction falls well into the consensus and we can note that our own error bars are certainly too optimistic (small). There are some caveats in analyzing the raw results of our solar cycle 25 predictions. For instance, our current dynamo model does not yet include large asymmetries between rising and declining phases of the cycle nor a time derivative of the meridional circulation state (we assume constant flow for the time being but can perform time dependent inversions as demonstrated in ([Hung et al. 2017](#))). Further the Waldmeier effect is only captured by adapting the diffusivity profile at the base of the convection zone as done by ([Karak & Choudhuri 2011](#)). Such a modelling shortcoming induces biases that are hard to quantify, and they do not appear in our uncertainty analysis, which should therefore be considered with caution, as it likely underestimates the systematic errors impacting our prediction. Our pipeline can also perform north and south hemispheres specific predictions and will report on them in a future paper.

Acknowledgments

We acknowledge financial support through IRS SpaceObs project of Paris-Saclay, SolarGeoMag project of Labex UnivEarths, the ERC Synergy grant 810218 WHOLE SUN and ERC Proof of Concept grant 640997 SolarPredict, CNES Solar Orbiter and Space Weather grants and INSU/PNST funding.

References

- Brun, A. S. 2007. Towards using modern data assimilation and weather forecasting methods in solar physics. *Astronomische Nachrichten*, 328, 329
- Brun, A. S. & Browning, M. K. 2017, Magnetism, Dynamo Action and the Solar-Stellar Connection, *Living Reviews in Solar Physics*, 14, 4

- Cameron, R. & Schüssler, M. 2008 A Robust Correlation between Growth Rate and Amplitude of Solar Cycles: Consequences for Prediction Methods. *The Astrophysical Journal*, 685, 1291
- Clette, F. & Lefèvre, L. 2012, Are the sunspots really vanishing?. Anomalies in solar cycle 23 and implications for long-term models and proxies. *Journal of Space Weather and Space Climate*, 2, A06
- Clette, F. & Lefèvre, L. 2018, The new Sunspot Number: continuing upgrades and possible impacts, IAU Symposium, 17
- Hathaway, D. H. 2015. The Solar Cycle. *Living Reviews in Solar Physics*, 12, 4
- Hung, C. P., Jouve, L., Brun, A. S., Fournier, A., Talagrand, O., *et al.* 2015. Estimating the Deep Solar Meridional Circulation Using Magnetic Observations and a Dynamo Model: A Variational Approach. *The Astrophysical Journal*, 814, 151
- Hung, C. P., Brun, A. S., Fournier, A., Jouve, L., Talagrand, O., Zakari, M., *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. *The Astrophysical Journal*, 849, 160
- Jouve, L. & Brun, A. S. 2007. On the role of meridional flows in flux transport dynamo models. *Astronomy and Astrophysics*, 474, 239–250
- Jouve, L., Brun, A. S., Talagrand, O., *et al.* 2011. Assimilating Data into an $\alpha\Omega$ Dynamo Model of the Sun: A Variational Approach. *The Astrophysical Journal*, 735, 31
- Karak, B. B. & Choudhuri, A. R. 2011. The Waldmeier effect and the flux transport solar dynamo. *Monthly Notices of the Royal Astronomical Society*, 410, 1503–1512
- Petrovay, K. 2019, Solar Cycle prediction, arXiv e-prints, [arXiv:1907.02107](https://arxiv.org/abs/1907.02107)
- Sanchez, S., Fournier, A., & Aubert, J. 2014, The Predictability of Advection-dominated Flux-transport Solar Dynamo Models. *The Astrophysical Journal*, 781, 8
- Talagrand, O. 2010, in *Data Assimilation: Making Sense Of Observations*, ed. W. Lahoz, B. Khattatov & R. Menard (Berlin: Springer)
- Weber, M., Upton, L. Biesecker, D., *et al.* 2019, Solar Cycle 25 Prediction, <https://www.swpc.noaa.gov/news/solar-cycle-25-forecast-update>