Modelling Persistent Cycles in Solar Activity

Federico Maddanu[1*] **Tommaso Proietti**[*]

[*] University of Rome Tor Vergata

[1] federico.maddanu@gmail.com

Abstract

Solar activity at decadal time scales is characterized by persistent periodic patterns with global effects on the Earth's climate. The paper deals with the analysis and prediction of monthly sunspots numbers, adopting a recently proposed time series model for long range dependent cycles. Learning is based on maximum likelihood and optimal signal extraction filters are available for cycle estimation and prediction. The analysis confirms the presence of stationary long memory in the sunspots generating process. Moreover, our formulation provides a reliable method for solar cycles predictions, yielding forecasts of the oncoming 25th cycle. In particular, we claim a main peak in early 2024 with amplitude 114 and the ending of the cycle in early 2030.

Introduction

Understanding solar magnetic activity is a crucial and practical issue, since it affects climate change and technology on the Earth. The sunspots series constitutes the longest homogeneous global indicator of solar activity determined by direct solar observations. It is characterized by a quasi-periodic behaviour with period length averaging 10-11 years, which is known in the literature as the *Schwabe* cycle. However, the 11-years sunspots oscillation is not the only periodic variation in solar activity. A long term change in its amplitude with the average period of about 80-90 years is referred to as the *Gleissberg* cycle, which has troughs at the so called "Dalton Minimum" (around 1800-1820) and "Gleissberg Minimum" (around 1890-1910) and at a peak at the "Modern Maximum" (around 1950).

The paper deals with the analysis and prediction of monthly sunspots numbers, adopting a recently proposed time series model for long range dependent cycles. Learning is based on maximum likelihood and optimal signal extraction filters are available for cycle estimation and prediction.

The analysis on the predictions of the past solar cycles suggests that our approach may lead to accurate forecasts of the sunspots series. On the other hand, our formulation shares the main advantage of time series methods, such that cycle predictions have equal chance of success in the decaying or rising phase of the cycle, across the maximum or, in particular, across the minimum. This is in contrast with polar precursors methods, where prediction usually becomes available after the minimum (see [3]).

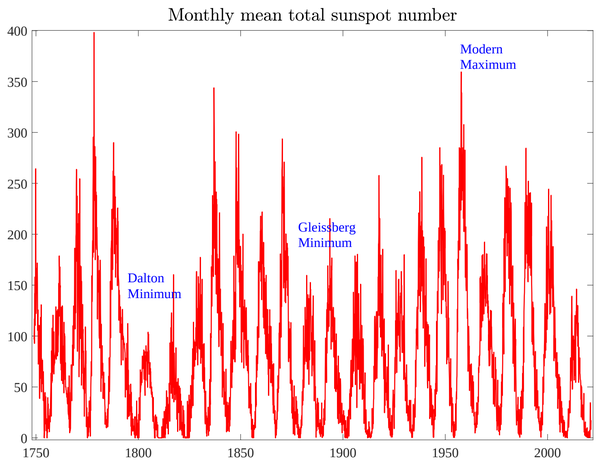

Figure 1: Monthly average sunspot numbers. Source: Royal Observatory of Belgium.

Methodology

The following linear additive model is considered:

$$y_t = \psi_{0t} + \sum_{j=1}^{M} \psi_{jt}, \tag{1}$$

where ψ_{0t} is the red noise component, capturing the variability of the process at the long run frequency,

$$\psi_{0t} = \mu + \phi\psi_{0,t-1} + \eta_{0t}, \quad \eta_{0t} \sim \text{i.i.d. } \mathrm{N}(0, \sigma^2_{\eta 0}). \tag{2}$$

The process is stationary if $|\phi| < 1$, in which case, $\mathrm{Var}(\psi_{0t}) = \sigma^2_{\psi 0}$, $\sigma^2_{\psi 0} = \frac{\sigma^2_{\eta 0}}{1-\phi^2}$ and $\mathrm{Cov}(\psi_{0t}, \psi_{0,t-k}) = \phi^k \frac{\sigma^2_{\eta 0}}{1-\phi^2}$. The components ψ_{jt} are fractional Sinusoidal Waveform (fSW) processes (c.f. [4]) letting λ_j, $j = 1, \cdots, M$, denote a frequency in $(0, \pi)$, the following specification provides a model for a stochastic long memory cycle with periods $P_j = 2\pi/\lambda_j$:

$$\begin{aligned} \psi_{jt} &= \alpha_{jt}\cos(\lambda_j t) + \alpha^*_{jt}\sin(\lambda_j t) \\ (1-L)^{d_j}\alpha_{jt} &= \eta_{jt}, && \eta_{jt} \sim \text{i.i.d. } \mathrm{N}(0, \sigma^2_{\eta j}), \\ (1-L)^{d_j}\alpha^*_{jt} &= \eta^*_{jt}, && \eta_{jt} \sim \text{i.i.d. } \mathrm{N}(0, \sigma^2_{\eta j}), \end{aligned} \tag{3}$$

where $E(\eta_{jt}\eta^*_{j,t-s}) = 0$ for all $(t, s) \in (Z)$. The properties of the process depend on the memory parameter d_j, the frequency λ_j, and the disturbance error variance $\sigma^2_{\eta j}$. If $d_j < 0.5$, ψ_{jt} is a stationary process with unconditional variance $\mathrm{Var}(\psi_{jt}) = \sigma^2_{\psi j}$, $\sigma^2_{\psi j} = \sigma^2_{\eta j}\frac{\Gamma(1-2d_j)}{\Gamma^2(1-d_j)}$. The model parameters in (1)-(3) are estimated by maximum likelihood, with the support of the Durbin-Levinson (DL) algorithm (see [2]). Finally, signal extraction for the k-th cycle component in (1), $\mathbf{s} = [s_1, \cdots, s_t, \cdots, s_{n+h}]'$, where, for instance, $s_t = \alpha_{jt}\cos(\lambda_j t) + \alpha^*_{jt}\sin(\lambda_j t)$ with $h \geq 0$, is available in closed form, as

$$\hat{\mathbf{s}} = \mathbf{\Gamma}_{\mathbf{s},\mathbf{y}}\tilde{\mathbf{\Phi}}'_n\tilde{\mathbf{D}}_n\tilde{\mathbf{\Phi}}_n(\mathbf{y} - \boldsymbol{\mu}) \tag{4}$$

where $\mathbf{y} = [y_1, \cdots, y_n]'$, $\tilde{\mathbf{\Phi}}'_n\tilde{\mathbf{D}}_n\tilde{\mathbf{\Phi}}_n$ is the matrix decomposition of the inverse covariance matrix of $\mathbf{y}$ obtained via the DL algorithm, while $\mathbf{\Gamma}_{\mathbf{s},\mathbf{y}} = \mathrm{Cov}(\mathbf{s}, \mathbf{y})$ has (i, j) elements equal to $\tilde{\sigma}^2_{\eta k}\frac{\Gamma(1-2\hat{d}_k)\Gamma(\hat{d}_k+|i-j|)}{\Gamma(1-\hat{d}_k+|i-j|)\Gamma(\hat{d}_k)\Gamma(1-\hat{d}_k)}\cos(\lambda_k(|i-j|))$. The extraction of the red noise component and the prediction of y_t follow straightforwardly.

Main analysis and results

Data. The series analysed in the current application is the monthly mean total sunspot number available at the World Data Center SILSO (`http://sidc.oma.be/silso/datafiles`) and displayed in figure 1. We focus the analysis not directly on the sunspots series itself, but on the Box-Cox transformation (c.f. [1]) thereof, aiming at stabilization of the variance and at the reduction. The optimal values of the power and location parameters in the Box-Cox transform have been obtained via a grid search procedure.

Model Selection. The periodogram of the transformed series indicates the highest peak at frequency 0.0481, implying a periodicity of about 10.89 years. Therefore, we set model (1)-(3) with $\lambda_1 = 0.0481$, while the remaining $M - 1$ frequencies correspond to the successive harmonics $\lambda_j = \lambda_1 j$, for $j = 2, \cdots, M$, such that the Schwabe cycle will be fitted by the sum of fSW components $\sum_{j=1}^{M}\psi_{jt}$, whereas the red noise component ψ_{0t} determines the long term Gleissberg variation. As far as the choice of the number M of fSW components is concerned, we set $M = 4$. The reason for contemplating 4 harmonics is to avoid the presence of non stationary cycles, indeed, via the implementation of a Lagrange multiplier test of the null $H_0 : d_j = 0.5$, we found that the hypothesis of non stationary cycles cannot be rejected if $M < 4$.

Model Estimation. Table 1 presents the maximum likelihood estimates of the parameters. The estimates of the memory parameters for the first two fSW components are close to the non stationary threshold of 0.5, denoting high persistence and long memory, while the remaining d_3 and d_4 estimates show negative memory, even so they represent altogether only the 0.35% of the total variability of the series, so that dropping them from the main analysis does not change the results remarkably. The variance decomposition of the series is presented in the last column of table 1 in terms of the $\sigma^2_{\psi j}$ parameters. Observe that the first fSW component represents the 93.73% of the total variance of the series, while the second one results to be very less significant, contributing just for the 0.68%. Finally, the autoregressive coefficient of the red noise component is positive and close to 1, such that the noise term actually tends to be brown in the sense of a disturbance generated by a Brownian motion. This component accounts for 5.24% of the total variance of the series.

Table 1: Maximum likelihood estimates of the parameters of the model (1)-(3).

	λ_j	$\hat{d}_j$	$\hat{\sigma}^2_{\eta,j}$	$\hat{\sigma}^2_{\psi j}$
fSW component	0.0481	0.4958	2.8235	109.3177
	0.0962	0.4632	0.1661	0.7933
	0.1443	-0.2736	0.0379	0.0414
	0.1924	-0.2255	0.3438	0.3662
	$\hat{\mu}$	$\hat{\phi}$	$\hat{\sigma}^2_{\varepsilon}$	$\hat{\sigma}^2_{\psi 0}$
Red noise component	10.4266	0.9939	0.0745	6.1134

Model (1)-(3) performs quite well in capturing the dynamic structure of the series. Residuals appear very weakly autocorrelated, and the residual periodogram does not display any residual periodicity. Finally, the comparison with the quantiles of the Gaussian distribution and the histogram do not show any systematic departure from a Gaussian distribution.

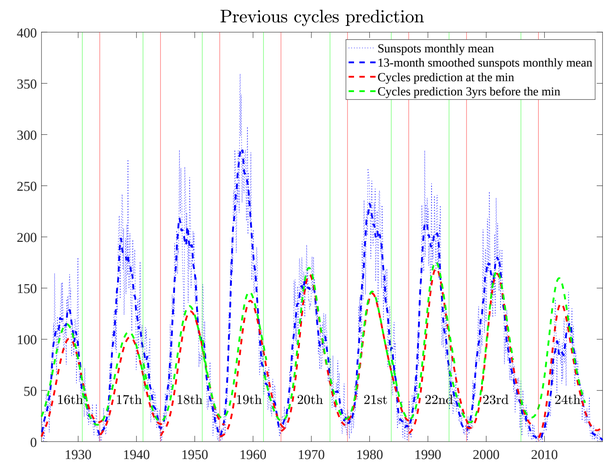

Figure 2: 11-years ahead out-of-sample predictions for each of the last nine solar cycles, with increasing windows starting both from the cycles minima (vertical red lines) and three years before the minima (vertical green lines).

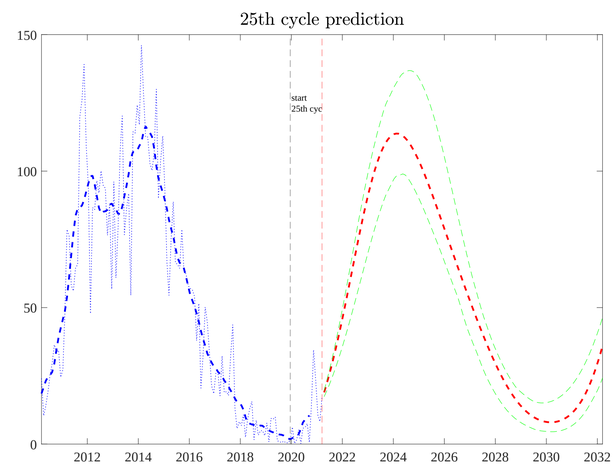

Figure 3: 11-years ahead out-of-sample predictions for the 25th cycles starting from March 2021, the dashed green lines are the confidence bands obtained via a bootstrap procedure. The blue dashed line in the background of both panels is the 13-month smoothed monthly sunspots series.

Solar cycles prediction. Typical issues in solar cycles prediction regard the determination of the amplitude and duration of a cycle. The amplitude is defined as the local maximum value reached by the sunspots series in the inter-cycle interval. The latter is delimited by the local minimum value observed before the local maximum (coinciding with the end of the previous cycle) and the local minimum value reached after the local maximum (which establishes the beginning of the next cycle). The difference between these two minima determines the duration of the cycle.

As we already said, the main advantage of our time series method is that predictions are available at each time point with equal chance of success. This is in contrast with most of the precursors methods, which provides more accurate forecasts of the next solar maximum in the proximity of the previous minimum. Figure 2 strongly highlights this peculiarity in our approach by showing the 11-years ahead out-of-sample predictions for each of the last nine cycles obtained according to model (1)-(3), with an increasing window starting both from the minima (dashed red line) and three years before the minima (dashed green line). Indeed, there is not a huge difference between the two predictors.

Finally, figure 3 shows forecasts for the oncoming 25th cycle starting from March 2021. The confidence bands (dashed green lines) are obtained via the bootstrap procedure described in [4]. We predict the 25th cycle's maximum in February 2024 with amplitude 114 in the range $(103, 126)$, and we expect the end in February 2030. Notice that the predicted amplitude is close to that claimed by polar precursors methods in recent years (see [3]). This is not irrelevant since polar precursor methods have consistently proven their skills in solar cycles prediction and are highly regarded in the solar activity literature (see [3]).

Conclusions

The paper has proposed a novel time series model for analysing and predicting solar activity, as measured by the monthly sunspots numbers. The analysis confirms the presence of long memory in the sunspots series and suggests that the latter are driven by stationary stochastic cycles.

Furthermore, the 11-years ahead out-of-sample predictability is a testbed for the past nine solar cycles, showing that the model provides an easy and accurate method for the prediction of the solar cycle maximum amplitude and duration. Finally, we provide forecasts of the oncoming 25th cycle, predicting a main peak in early 2024 with amplitude 114 and the ending in early 2030. The predicted maximum is close to that claimed by polar precursors methods in recent years (see [3]), which rank very high for solar cycles prediction in the solar activity literature.

References

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252.

Brockwell, P. J., & Davis, R. A. (1991). *Time series: theory and methods*. Springer Science & Business Media.

Petrovay, K. (2020). Solar cycle prediction. *Living Rev Sol Phys*, *17*(2).

Proietti, T., & Maddanu, F. (2021). Modelling cycles in climate series: the fractional sinusoidal waveform process. *Submitted in the Journal of Econometrics.*.