Solar Physics
DOI: 10.1007/•••••-•••-•••-••••-•

Methodology for Predicting the Probability Distribution of the Amplitude of Solar Cycle 25

Iñigo Arregui[1,2]

© Springer ••••

Abstract A number of precursor-type methods for solar-cycle prediction are based on the use of regression models and confidence-level estimates. A drawback to these methods is that they do not permit one to make probability statements, nor do they offer straightforward ways to propagate the uncertainty from the observations to the quantities of interest. We suggest a method to calculate the probability for the maximum amplitude of Solar Cycle 25 using Bayesian inference. We illustrate this approach with the predictions made by one particular phenomenological model that relates the time interval between termination events of preceding cycles with the amplitude of the next cycle. Our results show well-constrained posterior-predictive distributions for the maximum sunspot-number. The impact of uncertainty on sunspot-number and time interval between terminators is quantified. A comparison between past maximum sunspot-number values and posterior-predictive distributions computed using the method enables us to quantify the quality of the inference and of the prediction.

Keywords: Solar cycle - Sunspots, Statistics

1. Introduction

Solar-cycle prediction has been a matter of notable interest over decades and has led to a vast literature (see, e.g., Hathaway, 2015; Petrovay, 2020; Nandy, 2021). Among the different techniques employed, precursor methods aim at predicting the amplitude of a given solar cycle based on some measure of solar activity or magnetism of the preceding cycle at a given moment of time.

✉ I. Arregui
iarregui@iac.es

[1] Instituto de Astrofísica de Canarias Via Lactea S/N, E-38205 La Laguna, Tenerife, Spain

[2] Departamento de Astrofísica, Universidad de La Laguna E-38206 La Laguna, Tenerife, Spain

For example, an empirical relationship between the time interval between termination events and the amplitude of the upcoming solar cycle was recently proposed by McIntosh et al. (2020). Termination events delimit epochs of toroidal-magnetic-activity-band interaction and mark the limit between 11-year sunspot cycles at the end of 22-year magnetic-activity cycles. According to the proposed relationship, widely separated terminators would correspond to low-amplitude sunspot cycles. Conversely, more narrowly separated terminators would lead to large-amplitude sunspot cycles. Given a prediction of a terminator event in 2020 by Leamon et al. (2020) and their phenomenological model linking it to the sunspot cycle, McIntosh et al. (2020) predict that Solar Cycle 25 will have a large magnitude, in sharp contrast to the prediction from the NOAA/NASA Solar Cycle 25 Prediction Panel (SC25PP)[1].

A drawback to the statistical method employed by McIntosh et al. (2020) (and similar studies) is that no statement can be made about the probability of the quantity of interest: the sunspot-number. For instance, the statement by McIntosh et al. (2020) "We predict with 95 % confidence that the Cycle 25 amplitude will fall between 153 and 305 spots" cannot be interpreted as "The probability that the Cycle 25 amplitude will fall between 153 and 305 spots is 95 %". The prediction by McIntosh et al. (2020) gives a range constructed from confidence intervals from a linear regression of data for the time interval between termination events and the sunspot-number. A confidence interval is the frequentist method to measure the sampling uncertainty of regression coefficients. An interval is said to have a 95 % confidence when, as a result of a sampling procedure, in 95 % of all the samples drawn the confidence interval will cover the true value of the parameters. Once the confidence interval is calculated, it is inaccurate to state that the probability for the sunspot-number to lie in the interval is 95 %. One can only say that, in the long run, 95 % of the confidence intervals that were generated during the sampling procedure will include the true value (see, e.g., Neyman, 1937; D'Agostini, 2003; von Toussaint, 2011; Morey et al., 2016, for further clarification).

A frequentist analysis does not permit us to make probability statements because interpreting confidence intervals in terms of probability statements is not straightforward. Because a confidence interval informs us about the long-run frequency of confidence intervals that contain the true parameter, no statement can be made about the uncertainty of the inferred sunspot-number depending on the uncertainty of observed quantities. We propose a method for computing the probability distribution of the maximum amplitude of Solar Cycle 25 using Bayesian inference. This method can be applied to individual prediction models and should therefore not be confused with estimating uncertainties from ensembles of predictions based on different models. Here, we illustrate this method using predictions made by the phenomenological model by McIntosh et al. (2020).

[1]The official NOAA press release with the Panel's forecast can be found at www.weather.gov/news/190504-sun-activity-in-solar-cycle

2. Methodology

In the Bayesian framework (see, e.g., Jaynes, 2003; Lindley, 2014), the probability of an event E is conditioned by another event F in a way dictated by Bayes's theorem to give the probability of event E given event F: $p(E|F)$. This scheme can be applied to scientific inference about an unknown parameter vector $\boldsymbol{\theta}$, conditional on the assumption of a model $\mathcal{M}$ and on observed data $\mathcal{D}$, to calculate the posterior probability of the parameters $p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})$ as

$$p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M}) = \frac{p(\mathcal{D}|\boldsymbol{\theta},\mathcal{M})\,p(\boldsymbol{\theta}|\mathcal{M})}{p(\mathcal{D}|\mathcal{M})}, \tag{1}$$

with $p(\mathcal{D}|\boldsymbol{\theta},\mathcal{M})$ being the likelihood function and $p(\boldsymbol{\theta}|\mathcal{M})$ the prior distribution over the parameter space. The denominator $[p(\mathcal{D}|\mathcal{M})]$ is the so-called marginal likelihood, evidence, or prior-predictive distribution of the data

$$p(\mathcal{D}|\mathcal{M}) = \int_{\boldsymbol{\theta}} p(\mathcal{D}|\boldsymbol{\theta},\mathcal{M})\,p(\boldsymbol{\theta}|\mathcal{M})\,\mathrm{d}\boldsymbol{\theta}. \tag{2}$$

The evidence is a measure of the quality of the model. It quantifies how well the data $\mathcal{D}$ are predicted by the model $\mathcal{M}$.

Once the inference of the model parameters $\boldsymbol{\theta}$ has been performed, a distribution of possible unobserved data $\widetilde{\mathcal{D}}$ conditional on the observed data $\mathcal{D}$ and the assumed model $\mathcal{M}$ is given by the so-called posterior-predictive distribution

$$p(\widetilde{\mathcal{D}}|\mathcal{D},\mathcal{M}) = \int_{\boldsymbol{\theta}} p(\widetilde{\mathcal{D}}|\boldsymbol{\theta},\mathcal{M})\,p(\boldsymbol{\theta}|\mathcal{D},\mathcal{M})\,\mathrm{d}\boldsymbol{\theta}. \tag{3}$$

The first factor in the integrand is the likelihood of the new unobserved data as a function of the parameter vector. The second factor is the posterior inferred from the old observed data.

3. Analysis and Results

Following McIntosh et al. (2020), let us assume a linear relationship between the maximum sunspot-number SSN and the time interval between termination events ΔT, $\mathrm{SSN} = \mathcal{M}(\Delta T|\alpha,\beta) = \alpha\Delta T + \beta$, and let us infer the posterior density for the slope α and the intercept β of the model $[p(\alpha,\beta|\mathcal{D},\mathcal{M})]$ with $\mathcal{D} = d_i = \{\mathrm{SSN}_i, \Delta T_{i-1}\}_{i=2}^{24}$ being the (past) observed data in Table 1 of McIntosh et al. (2020). The maximum sunspot-number amplitudes correspond to monthly-averages (see also, Table 1 of Pesnell, 2018).

In the Bayesian framework, the data are considered fixed and the parameters are unknowns with a given probability distribution. The likelihood of a given data realisation $[d_i]$ as a function of the parameter vector $[\boldsymbol{\theta}{=}(\alpha,\,\beta)]$ is the so-called likelihood function. A particular choice for the special case of a straight line model when there are independent errors in both data coordinates is (see, e.g., Gregory, 2005)

$$p(\mathcal{D}|\mathcal{M},\alpha,\beta) \;=\; (2\pi)^{-N/2}\left(\prod_{i=1}^{N}(\sigma^2_{\mathrm{SSN}_i}+\alpha^2\sigma^2_{\Delta T_{i-1}})^{-1/2}\right) \times\; \exp\left\{\sum_{i=1}^{N}\frac{-[d_i-\mathcal{M}(\Delta T_{i-1}|\alpha,\beta)]^2}{2(\sigma^2_{\mathrm{SSN}_i}+\alpha^2\sigma^2_{\Delta T_{i-1}})}\right\} \tag{4}$$

with each σ_{SSN} and $\sigma_{\Delta T}$ expressing the uncertainty on sunspot-number and time interval between termination events, respectively. As shown by Mathieu et al. (2019), the sunspot number tends to follow a stretched Poisson law. This is important in the context of studies based on the distribution of data. The likelihood function in Equation 4 does not represent the likelihood of different data realisations. Its purpose is first, to measure the discrepancy between model predictions and observed data as a function of the model parameters, taking as reference the uncertainty of the data. Second, to accordingly assign different levels of likelihood to alternative parameter combinations.

We adopt uniform priors for the two unknown parameters over broad enough ranges to accommodate our posteriors: $\mathcal{U}(\alpha,-60,-10)$ and $\mathcal{U}(\beta,400,600)$. The joint posterior $p(\alpha,\beta|\mathcal{D},\mathcal{M})$ is proportional to the likelihood function.

The marginal posterior for the slope α is

$$p(\alpha|\mathcal{D},\mathcal{M}) = \int \mathrm{d}\beta\, p(\alpha,\beta|\mathcal{D},\mathcal{M}) \propto p(\alpha|\mathcal{M}) \int \mathrm{d}\beta\, p(\beta|\mathcal{M}) p(\mathcal{D}|\mathcal{M},\alpha,\beta), \tag{5}$$

and likewise for the intercept β

$$p(\beta|\mathcal{D},\mathcal{M}) = \int \mathrm{d}\alpha\, p(\alpha,\beta|\mathcal{D},\mathcal{M}) \propto p(\beta|\mathcal{M}) \int \mathrm{d}\alpha\, p(\alpha|\mathcal{M}) p(\mathcal{D}|\mathcal{M},\alpha,\beta). \tag{6}$$

The two marginal posteriors for the slope and intercept are displayed in the top and middle panels of Figure 1, respectively. For this particular example, we assumed fixed values for the uncertainties on observed sunspot-number and terminator intervals, given by $\sigma_{\mathrm{SSN}} = 10$ and $\sigma_{\Delta T} = 0.1$, respectively. The value $\sigma_{\mathrm{SSN}} = 10$ corresponds roughly with an estimated error of about 1 sunspot group (Clette et al., 2014; Muñoz-Jaramillo and Vaquero, 2019). The posteriors show well-constrained probability densities, from which the following summaries are calculated: $\alpha = -31^{+3}_{-2}$ and $\beta = 532^{+17}_{-16}$, with the estimates corresponding to the median and the uncertainties given at the 68 % credible interval. This inference result encodes all the information that can be obtained from past data on sunspot-numbers and time intervals between termination events, the assumed model, and the adopted uncertainty values.

Following the definition in Equation 3, the posterior-predictive distribution for the future unobserved amplitude of Solar Cycle 25 [$\widetilde{\mathcal{D}} = \mathrm{SSN}_{25}$] based on the time interval for the preceding termination event [ΔT_{24}] can then be computed from the inferred posterior and by considering a Gaussian likelihood for the new data as a function of the parameter vector

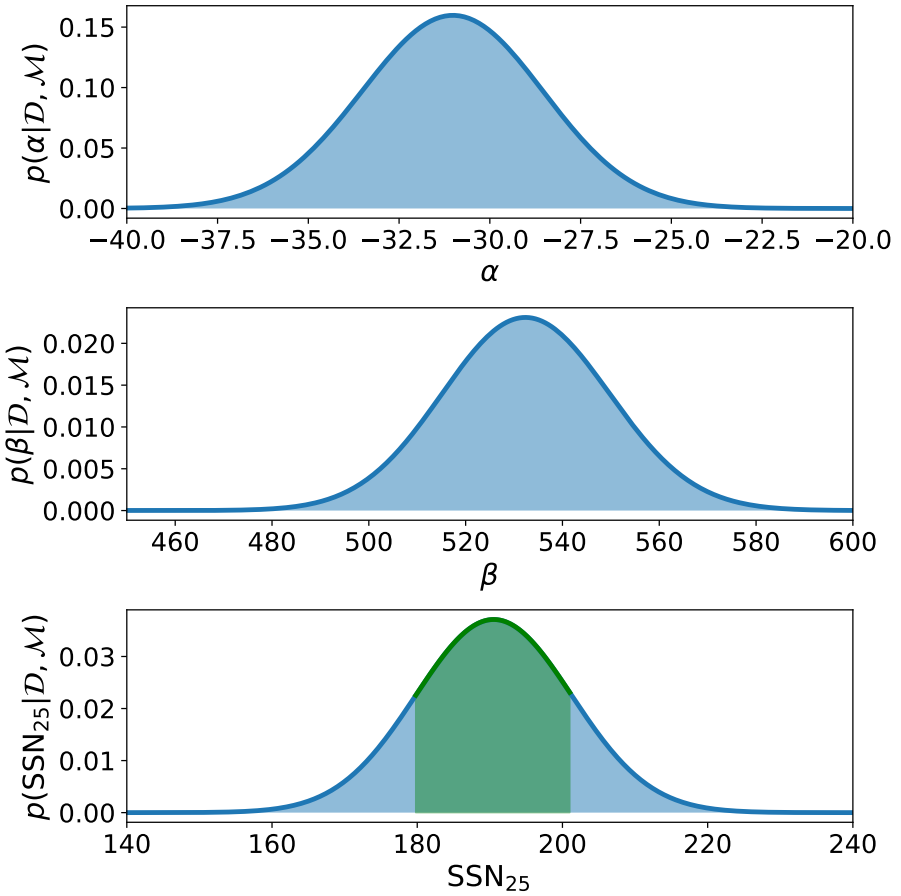

Figure 1. *Top and middle*: Marginal posterior densities for the slope α and the intercept β of the linear model $\mathcal{M}$ that account for the past sunspot-number data $\mathcal{D}$ computed with Equations 5 and 6. The calculations assume $\sigma_{\mathrm{SSN}_i} = 10$ and $\sigma_{\Delta T_{i-1}} = 0.1\,\forall i$. *Bottom*: Posterior-predictive distribution for the sunspot-number during Solar Cycle 25, based on the posterior from past data and the likelihood of new data, Equation 7. The *shaded green area* covers 68 % of the mass centred around the median: $\mathrm{SSN}_{25} = 190.6^{+10.6}_{-10.8}$. The calculation assumes $\Delta T_{24} = 10.72$.

$$p(\mathrm{SSN}_{25}|\mathcal{M},\alpha,\beta) = \frac{1}{\sqrt{2\pi}}\left((\sigma^2_{\mathrm{SSN}_{25}} + \alpha^2\sigma^2_{\Delta T_{24}})^{-1/2}\right) \times \exp\left\{\frac{-[\mathrm{SSN}_{25} - \mathcal{M}(\Delta T_{24}|\alpha,\beta)]^2}{2(\sigma^2_{\mathrm{SSN}_{25}} + \alpha^2\sigma^2_{\Delta T_{24}})}\right\}, \quad (7)$$

with $\sigma_{\mathrm{SSN}_{25}}$ and $\sigma_{\Delta T_{24}}$ expressing the uncertainty that we are willing to consider for the future sunspot-number and the last time interval between termination events, respectively.

Leamon et al. (2020) predicted an early date of 2020.37 for the termination event corresponding to Solar Cycle 24, leading to $\Delta T_{24} = 9.29$, which failed to happen (Leamon et al., 2021). Let us assume that the termination event occurs in October 2021. This leads to $\Delta T_{24} = 10.72$. The obtained posterior-predictive distribution for the maximum amplitude of Cycle 25 is displayed in the bottom panel of Figure 1 and shows a well-constrained posterior density. The rounded summary of the posterior-predictive distribution is $\mathrm{SSN}_{25} = 191^{+11}_{-11}$, with the estimate corresponding to the median and the uncertainties given at the 68 % credible interval. Employing the Leamon et al. (2020) prediction with $\Delta T_{24} = 9.29$ results in $\mathrm{SSN}_{25} = 236^{+11}_{-11}$, which is in good agreement with the estimate by McIntosh et al. (2020) of 233, the slight difference in the median being due to the inclusion of uncertainty in ΔT.

The main advantage of having the posterior-predictive distribution is that it is now perfectly possible and straightforward to make probability statements. For instance, according to the result displayed in Figure 1, the probability that the Solar Cycle 25 amplitude will fall between ≈ 180 and 201 spots is 68 %, the area under the green curve covering that percentage of the full probability mass.

Another advantage of this method is that it propagates uncertainty from the observations to the inferred quantities of interest. The posterior-predictive distribution in Figure 1 depends on the uncertainty of the data for the sunspot-number of past cycles and on the determination of the time interval between termination events. To produce the result, both quantities were fixed to certain values. Figure 2 shows how varying these two uncertainties impacts the resulting posterior-predictive distribution. The uncertainty on the sunspot-number affects the dispersion of the distribution, but location parameters such as the median remain unaffected (Figure 2 top panel). On the other hand, the uncertainty on the time interval between termination events has an effect not only on the dispersion, but also on the location parameters of the distribution (Figure 2 middle panel). Increasing $\sigma_{\Delta T}$ initially shifts the distribution to the right. Further increasing the value of $\sigma_{\Delta T}$ shifts the distribution back to the left while at the same time the dispersion of the distribution increases. This behaviour is further described and analysed in the Appendix.

Of course, these are just academic examples that may not reflect our current state of knowledge about the uncertainty on sunspot-number and time interval between terminators. The uncertainty on sunspot-number depends on the time and the observer. A better informed choice can be made by adopting the result by Dudok de Wit, Lefèvre, and Clette (2016) that for solar cycles after 1981 the error can be approximated by $\sigma_{\mathrm{SSN}} \approx 1.7\sqrt{\mathrm{SSN}+1}$. We adopt this expression noting that it seems to apply to daily values, for which a large fraction of the variability of the data comes from day-to-day scatter. Concerning the time interval between termination events, the times of the terminators are expressed as a decimal year by McIntosh et al. (2020), but as discussed by Leamon et al. (2021) the uncertainty in their determination is still unclear. The bottom panel in Figure 2 shows a comparison between the posterior-predictive distribution computed with a fixed error on sunspot-number, $\sigma_{\mathrm{SSN}} = 10$, and the one computed making use of the approximate formula by Dudok de Wit, Lefèvre, and Clette (2016). In the later case, the uncertainty in the obtained result is larger and the median

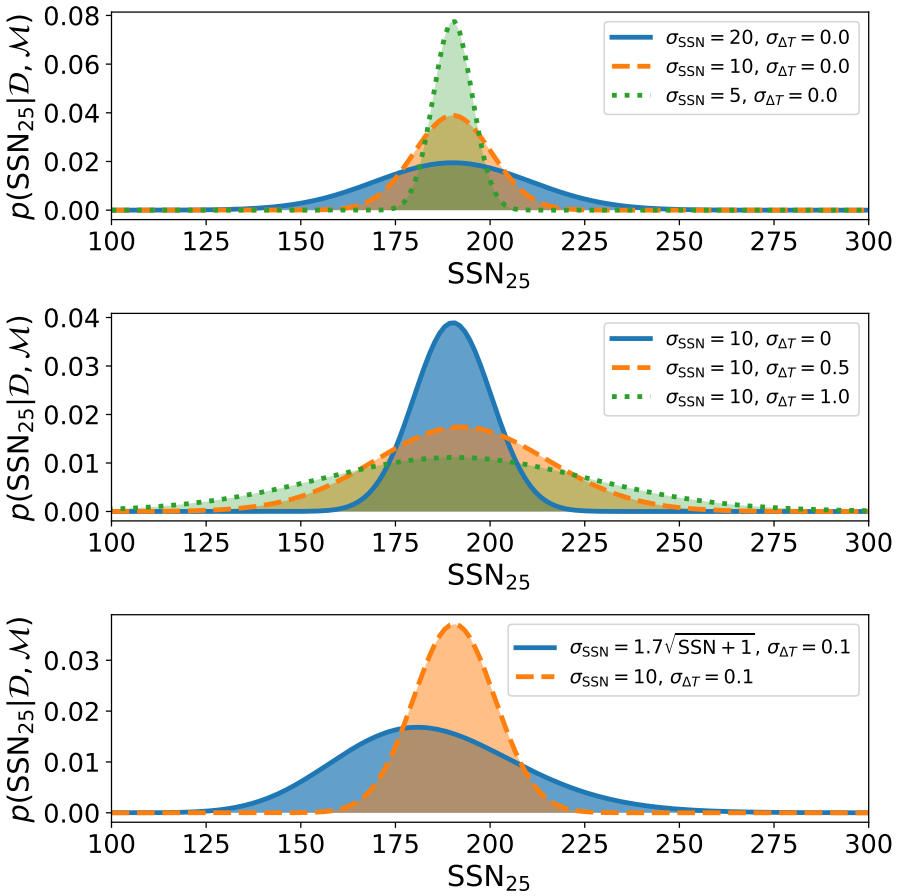

Figure 2. *Top*: Influence of uncertainty on sunspot-number [σ_{SSN}] on the posterior-predictive distribution for SSN_{25} for the case with no uncertainty on the time interval between termination events: $\sigma_{\Delta T} = 0$. *Middle*: Influence of uncertainty on the time interval between termination events [$\sigma_{\Delta T}$] on the posterior-predictive distribution for SSN_{25} for the case with $\sigma_{\mathrm{SSN}} = 10$. *Bottom*: A comparison between posterior-predictive distributions for SSN_{25} computed with a fixed error on sunspot-number and with the approximate formula by Dudok de Wit, Lefèvre, and Clette (2016).

of the distribution is affected too. The summaries are $\mathrm{SSN}_{25} = 191^{+11}_{-11}$ for the case with constant error on sunspot-number and $\mathrm{SSN}_{25} = 184^{+25}_{-22}$ for the solution that employs the formula from Dudok de Wit, Lefèvre, and Clette (2016), with the estimate corresponding to the median and the uncertainties given at the 68 % credible interval. The consequences of the relative dominance of the uncertainties on sunspot-number and time interval between termination events on the posterior-predictive distribution are further discussed in the Appendix.

Because the method provides us with a distribution of probability among different possible sunspot-number values, it makes it possible to assess in a straightforward and quantitative manner the goodness of the predictive capa-

bilities of the precursor ΔT and/or of the model $\mathcal{M}$, by comparing the actually occurring maximum sunspot-number with the predicted posterior distribution.

While Solar Cycle 25 is underway, we can do the exercise with past solar cycles and use the method to compute the posterior-predictive distribution for them, assuming that the inferred model $\mathcal{M}(\Delta T|\alpha,\beta)$ is valid and employing the time intervals for past termination events in Table 1 of McIntosh et al. (2020). The results are displayed in Figure 3. We can appreciate that the 95 % credible intervals of the computed posterior-predictive distributions cover the observed maximum sunspot-number in all except three cases (SC16, SC19, and SC21). The observed maximum sunspot-number values are within the 68 % credible interval of the prediction in 12 cases (SC2, SC3, SC4, SC8, SC12, SC15, SC17, SC18, SC20, SC22, SC23, and SC24). The median of the prediction is fairly accurate in six cases (SC2, SC4, SC8, SC15, SC17, and SC24). Of course, the optimism of these results must be moderated by the realisation that the predicted distributions are rather wide. The results in Figure 3 were computed with the $\sigma_{\rm SSN} = 1.7\sqrt{\rm SSN+1}$ approximate formula. In computations with $\sigma_{\rm SSN} = 10$ (not shown here) the 95 % credible intervals of the computed posterior-predictive distributions cover the observed maximum sunspot-number in 13 out of the 23 cases.

The Space Weather Prediction Center (NOAA) prediction is continuously updated at www.swpc.noaa.gov/products/predicted-sunspot-number-and-radio-flux. At the time of writing, the predicted minimum and a maximum peak sunspot-number are 104.6 and 124.6, respectively. This is in agreement with Carrasco and Vaquero (2021) who conclude that the maximum amplitude for SC25 could be small – moderate.

The SWPC/NOAA prediction interval, also shown in Figure 3, falls below our posterior-predictive distribution for Solar Cycle 25, with median 184^{+25}_{-22}. On the other hand, our estimate is in good agreement with the climatological forecast of 180 ± 60 by Pesnell (2018), which considers that the maximum amplitude of Solar Cycle 25 will be the average of all observed maxima.

Discrepancies between predictions based on different preferred solar activity parameters and relationships are not unusual. A meaningful comparison between our results and those from other precursor methods would need to first establish solid relationships between their respective indicators to shed light into the reasons for discrepancy. Also, the SWPC/NOAA panel prediction is the result of considerations about aspects such as when the minimum occurs; the evolution of several indicators, such as the sunspot number or the F10.7 flux; the presence or absence of high-latitude sunspots with the polarity of the new cycle; the dipolar nature of the coronal field; the flattening of the heliospheric current sheet, and other activity measures. Another points of discrepancy are the different perceptions about what the cycle maximum is, and the differences in the statistical modelling of the data.

For these reasons, the results presented in this study offer a method rather than a definite prediction. The accuracy of the results can only be as good as the model employed, on which they are conditional. High accuracy seems difficult to achieve with the use of a single precursor and a simple linear relationship between the indicators of interest. The offered advantage lies in a procedure that allows

a direct and meaningful comparison between observations and predictions that can be employed as both data and models are further refined.

4. Conclusion

We have presented a method to compute predictions for the maximum amplitude of Solar Cycle 25 based on Bayesian inference and adopting as an example application a precursor method recently proposed by McIntosh et al. (2020). The quantity of interest is the posterior-predictive distribution of the maximum sunspot-number. This quantity is a probability distribution and results from the combination of a posterior probability distribution of straight-line parameters inferred from past data for sunspot-number and time interval between terminators and a likelihood function for the unobserved future data on sunspot-number as a function of the model parameters. The resulting distributions are well-constrained and their properties depend on the uncertainties on sunspot-number and time interval between termination events. In our current state of knowledge about the sunspot-number, the distributions are rather broad. Observations and models should be improved to further constrain the probability distributions

The method has several useful qualities: First, it quantifies the probability of the quantity of interest, the maximum sunspot-number, so making statements about the possible values of this quantity is now straightforward. Second, the inference considers the propagation of uncertainty from the relevant observed quantities to the inferred parameters. Third, it serves to assess the quality of the predictor and/or the model, since the actually occurring outcome can be compared to the predicted probability distribution. The method is applicable to other predictive methods and alternative models as well. In that case, the Bayesian framework would also enable to compare their relative performance by assessing which one better explains the observed sunspot record.

Acknowledgements I thank the reviewer for constructive and insightful comments and suggestions. I am indebted to Scott W. Mcintosh for correspondence that motivated this study. I am grateful to Andrés Asensio Ramos, José Manuel Vaquero, and Víctor Manuel Sánchez Carrasco for helpful comments and suggestions. Calculations and figures were implemented with numpy (Harris et al., 2020) and matplotlib (Hunter, 2007).

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Funding This research was supported by project PGC2018-102108-B-I00 from Ministerio de Ciencia, Innovación y Universidades and FEDER funds.

Conflict of interest The author declares that there is no conflict of interest.

Appendix

Uncertainty Analysis

The uncertainty on the computed posterior-predictive distribution depends on the uncertainty on the sunspot-number [σ_{SSN}] and on the time interval between

termination events $[\sigma_{\Delta T}]$. These two quantities determine the shape and dispersion of the likelihood functions in Equations 4 and 7. Their possible values hence have an impact on the inferred posterior for α and β and the subsequent posterior-predictive distribution $p(\mathrm{SSN}_{25}|\mathcal{D},\mathcal{M})$. The uncertainty on the sunspot-number is subject to intense continuous investigation (see, e.g., Clette et al., 2014; Dudok de Wit, Lefèvre, and Clette, 2016; Muñoz-Jaramillo and Vaquero, 2019; Mathieu et al., 2019). The uncertainty on the time interval between terminator events needs to be quantified. Their relative dominance is important to quantify the uncertainty of our predictions.

Let us consider the approximate formula $\sigma_{\mathrm{SSN}} = 1.7\sqrt{\mathrm{SSN}+1}$ from Dudok de Wit, Lefèvre, and Clette (2016) for the uncertainty on sunspot-number data. The top panel of Figure 4 shows how the marginal posterior for the slope of the inferred model depends on different values for the uncertainty on the time interval between termination events $[\sigma_{\Delta T}]$. Increasing $\sigma_{\Delta T}$, the distribution initially moves towards the left and then towards the right. The distribution also becomes non-symmetric and has a larger dispersion. A similar effect is observed in the marginal posterior for the intercept $[\beta]$, not shown here. The middle panel of Figure 4 shows the corresponding solutions for the posterior-predictive distribution for Solar Cycle 25. Now the distribution first slightly shifts towards the right and then towards the left, as already mentioned for the case with $\sigma_{\mathrm{SSN}} = 10$ (see Figure 2). The uncertainty of the solution also increases with increasing $\sigma_{\Delta T}$.

This behaviour is closely linked to the relative dominance of the two sources of uncertainty in the inference and prediction processes. This is shown in the bottom panel of Figure 4. When the uncertainty on the time interval between termination events is negligible or relatively small, the full uncertainty of the inference $[\sigma^2]$ is determined by the uncertainty on the sunspot-number. As we increase the uncertainty on ΔT, the total uncertainty increases, thus the posteriors widen, and at the same time slightly shift. At a given moment, for $\sigma_{\Delta T} \approx 0.8$, both uncertainties contribute equally to the total uncertainty. From there on, the uncertainty on ΔT dominates and the posteriors move in the opposite direction while they become more asymmetric. For an error in the estimate of the time interval between events below ≈ 0.5 years, the uncertainty of the problem will be predominantly determined by the uncertainty on the sunspot-number.

References

Carrasco, V.M.S., Vaquero, J.M.: 2021, Solar Cycle 25 is Currently Very Similar to Solar Cycle 24. *Res. Notes Am. Astron. Soc.* **5**, 181. DOI. ADS.

Clette, F., Svalgaard, L., Vaquero, J.M., Cliver, E.W.: 2014, Revisiting the Sunspot Number. A 400-Year Perspective on the Solar Cycle. *Space Sci. Rev.* **186**, 35. DOI. ADS.

D'Agostini, G.: 2003, *Bayesian Reasoning in Data Analysis*, World Scientific, Singapore. DOI.

Dudok de Wit, T., Lefèvre, L., Clette, F.: 2016, Uncertainties in the Sunspot Numbers: Estimation and Implications. *Solar Phys.* **291**, 2709. DOI. ADS.

Gregory, P.C.: 2005, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support*, Cambridge University Press, Cambridge UK. ISBN 9780521841504. DOI. ADS.

Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk,

M.H., Brett, M., Haldane, A., del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E.: 2020, Array programming with NumPy. *Nature* **585**, 357. DOI. ADS.

Hathaway, D.H.: 2015, The Solar Cycle. *Liv. Rev. Solar Phys.* **12**, 4. DOI. ADS.

Hunter, J.D.: 2007, Matplotlib: A 2D Graphics Environment. *Comp. Sci. Eng.* **9**, 90. DOI. ADS.

Jaynes, E.T.: 2003, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge UK. ISBN 9780511790423. DOI.

Leamon, R.J., McIntosh, S.W., Chapman, S.C., Watkins, N.W.: 2020, Timing Terminators: Forecasting Sunspot Cycle 25 Onset. *Solar Phys.* **295**, 36. DOI. ADS.

Leamon, R.J., McIntosh, S.W., Chapman, S.C., Watkins, N.W.: 2021, Response to "Limitations in the Hilbert Transform Approach to Locating Solar Cycle Terminators" by R. Booth. *Solar Phys.* **296**, 151. DOI. ADS.

Lindley, D.V.: 2014, *Understanding Uncertainty*, John Wiley & Sons, Inc., Hoboken, NJ. ISBN 9781118650158. DOI.

Mathieu, S., von Sachs, R., Ritter, C., Delouille, V., Lefèvre, L.: 2019, Uncertainty Quantification in Sunspot Counts. *Astrophys. J.* **886**, 7. DOI. ADS.

McIntosh, S.W., Chapman, S., Leamon, R.J., Egeland, R., Watkins, N.W.: 2020, Overlapping Magnetic Activity Cycles and the Sunspot Number: Forecasting Sunspot Cycle 25 Amplitude. *Solar Phys.* **295**, 163. DOI. ADS.

Morey, R.D., Hoekstra, R., Rouder, J.N., Lee, M.D., Wagenmakers, E.-J.: 2016, The fallacy of placing confidence in confidence intervals. *Psychonomic Bull. Rev.* **23**, 103. DOI.

Muñoz-Jaramillo, A., Vaquero, J.M.: 2019, Visualization of the challenges and limitations of the long-term sunspot number record. *Nat. Astron.* **3**, 205. DOI. ADS.

Nandy, D.: 2021, Progress in Solar Cycle Predictions: Sunspot Cycles 24-25 in Perspective. *Solar Phys.* **296**, 54. DOI. ADS.

Neyman, J.: 1937, Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. *Phil. Trans. Roy. Soc. London Ser. A* **236**, 333. DOI. ADS.

Pesnell, W.D.: 2018, Effects of Version 2 of the International Sunspot Number on Naïve Predictions of Solar Cycle 25. *Space Weather* **16**, 1997. DOI.

Petrovay, K.: 2020, Solar cycle prediction. *Liv. Rev. Solar Phys.* **17**, 2. DOI. ADS.

von Toussaint, U.: 2011, Bayesian inference in physics. *Rev. Mod. Phys.* **83**, 943. DOI.

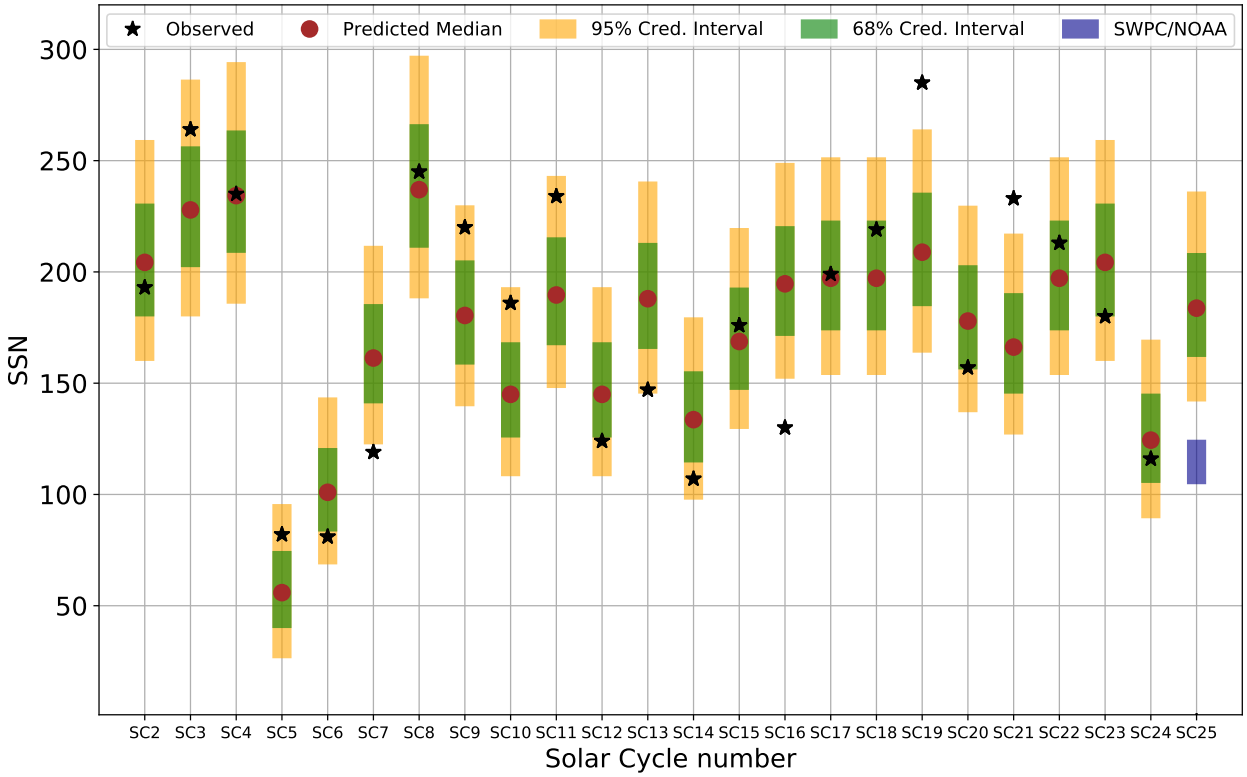

Figure 3.: Comparison between observed sunspot-number values for past Solar Cycles 2 to 24 from Table 1 of McIntosh et al. (2020) and posterior-predictive distributions computed using Equation 7 with data on Table 1 of McIntosh et al. (2020). In all calculations, $\sigma_{\rm SSN} = 1.7\sqrt{\rm SSN + 1}$ and $\sigma_{\Delta \rm T} = 0.1$. For Solar Cycle 25, $\Delta T = 10.72$ is employed. Also shown is the interval between the minimum and maximum peak sunspot-number values currently predicted by the SWPC/NOAA. Colours indicate different credible intervals of the obtained posterior distributions.

SOLA: manuscriptR2.tex; 30 January 2022; 17:59; p. 12

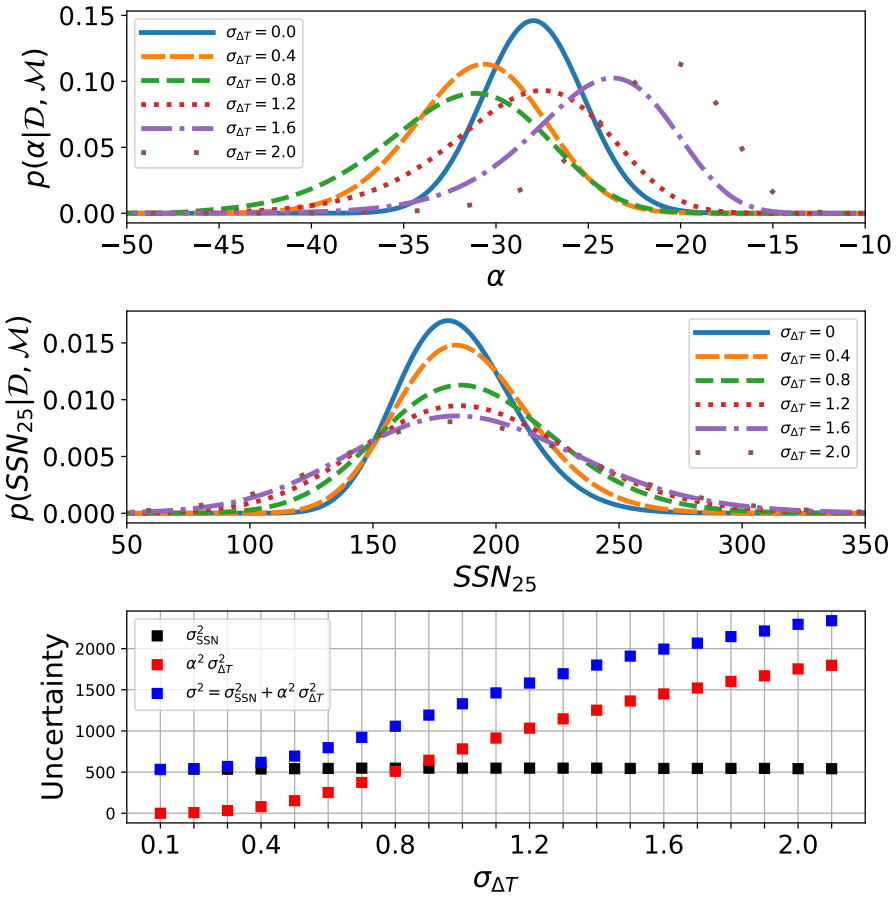

Figure 4. *Top*: Marginal posterior density for the slope α for inferences with $\sigma_{\rm SSN} = 1.7\sqrt{\rm SSN + 1}$ and different values of the uncertainty on the time interval between termination events $\sigma_{\Delta T}$. *Middle*: Corresponding posterior-predictive distributions for Solar Cycle 25. *Bottom*: Variation of the total uncertainty $[\sigma^2]$ and of each contribution $[\sigma^2_{\rm SSN}$ and $\alpha^2\sigma^2_{\Delta T}]$ in the likelihood functions, Equations 4 and 7, used in the inference and subsequent prediction.