The persistence of solar activity indicators and the descent of the Sun into Maunder Minimum conditions

M. Lockwood,\textsuperscript{1,2} M. J. Owens,\textsuperscript{1} L. Barnard,\textsuperscript{1} C. J. Davis,\textsuperscript{1,2} and F. Steinhilber\textsuperscript{3}

Received 30 September 2011; revised 25 October 2011; accepted 26 October 2011; published 30 November 2011.

[1] The recent low and prolonged minimum of the solar cycle, along with the slow growth in activity of the new cycle, has led to suggestions that the Sun is entering a Grand Solar Minimum (GSMi), potentially as deep as the Maunder Minimum (MM). This raises questions about the persistence and predictability of solar activity. We study the autocorrelation functions and predictability $R^2(t)$ of solar indices, particularly group sunspot number $R_G$ and heliospheric modulation potential $\Phi$ for which we have data during the descent into the MM. For $R_G$ and $\Phi$, $R^2(t) > 0.5$ for times into the future of $t \approx 4$ and $\approx 3$ solar cycles, respectively: sufficient to allow prediction of a GSMi onset. The lower predictability of sunspot number $R_Z$ is discussed. The current declines in peak and mean $R_G$ are the largest since the onset of the MM and exceed those around 1800 which failed to initiate a GSMi. Citation: Lockwood M., M. J. Owens, L. Barnard, C. J. Davis, and F. Steinhilber (2011), The persistence of solar activity indicators and the descent of the Sun into Maunder Minimum conditions, Geophys. Res. Lett. 38, L22105, doi:10.1029/2011GL049811.

1. The Recent Solar Minimum

[2] The minimum in solar activity between solar cycles 23 and 24 (SC23 and SC24) was unprecedentedly low and long-lived for the space age [e.g., Lockwood, 2010; Russell et al., 2010]. For example, the open solar flux and the near–Earth IMF fell to values not seen before in the space age [Smith and Balogh, 2008], indeed such low open flux values had not existed since about 1920 [Lockwood et al., 2009]. This minimum is part of a decline in average solar activity, as quantified by a variety of parameters, which has been present since about 1985 [Lockwood and Fröhlich, 2007]. Abreu et al. [2008] studied the durations of Grand Solar Maxima (GSMa) in solar activity during the past 9300 years, using the composite variation of the heliospheric cosmic-ray modulation potential $\Phi$ compiled by Steinhilber et al. [2008]. The GSMa were defined to be when 25-year means of $\Phi$ exceeded a fixed threshold of 616 MV. Abreu et al. [2008] deduced that recent decades formed a GSMa which was uniquely long-lived and so is due to end soon. This was supported by extrapolations of recent trends in heliospheric parameters by Lockwood et al. [2009]. Lockwood [2010] composites the evolution of the $\Phi$ reconstruction by Steinhilber et al. around the ends to the previous 24 GSMa in that dataset and so made an analogue forecast of how $\Phi$ is likely to evolve in future, including the probability that the Sun enters a Grand Solar Minimum (GSMi) similar to the Maunder Minimum (MM, circa 1645–1715). Barnard et al. [2011] have extended this forecast to other parameters, including sunspot number, using empirical relationships with $\Phi$.

[3] Owens et al. [2011b] have recently studied the evolution to date of cycle SC24 in sunspot number $R_Z$, Heliospheric Current Sheet (HCS) tilt and mean sunspot latitude. By assuming SC24 will continue to follow the average solar cycle behaviour, they predict a peak $R_Z$ of 65 $\pm$ 10 about the middle/end of 2012. This is consistent with the prior prediction by Svalgaard et al. [2005] from the solar polar fields but is significantly lower than NOAA’s most recent expert-panel prediction of peak $R_Z = 90$ around the middle of 2013. (http://www.swpc.noaa.gov/SolarCycle/SC24/index.html) [see also Pesnell, 2008]. Owens et al. [2011b] show that this evolution of $R_Z$ during SC24 is in the lowest 5 percentile of the potential future variations in $R_Z$ predicted by Barnard et al. [2011] which makes it consistent with a return to MM conditions within about 40 years. However, the relationship between $\Phi$ and $R_Z$ used by Barnard et al. [2011] was based on 25-year means and the fractional deviation of annual values of $R_Z$ from its 25 year means was then evaluated as a function of solar cycle phase. Both these steps introduce uncertainties which mean that the predictions of Barnard et al. [2011] for $R_Z$ do not have the same level of certainty as those for $\Phi$, which limits the extent to which individual cycles in $R_Z$ can be used to predict longer-term changes. We here study the variability and predictability of sunspot number $R_Z$, group sunspot number $R_G$, $\Phi$, and modelled open solar flux $F_S$, in order to evaluate the extent to which recent data can be used as an indicator of a solar decline towards a GSMi.

2. Variations Over the Past 400 Years

[4] Solar activity is often quantified using the Zurich (also called the Wolf or International) sunspot number defined as $R_Z = k (10G + N)$, where $G$ is the number of sunspot groups, $N$ is the total number of individual sunspots and $k$ is a calibration factor which allows for differences between observational techniques, sites and instruments (see reviews by Clette et al. [2007] and Vaquero [2007]). The data sequence extends back to 1700 but, like all such records, is increasingly unreliable at earlier times. Hoyt and Schatten [1998] developed the group sunspot number, defined by $R_G = (12.08/n) \sum_{i=1}^{n} k_G G_i$, where $n$ is the number of independent observers, $G_i$ is the number of sunspot groups
observer and have much higher persistence (broader ACF) (green). All to the corresponding minima in before the MM (SC-11 to Figure 1c deduced from in-situ magnetic R records: the has considerably lower. This may, in R R R is the calibration factor [2006] reconstruction to quantify the pro- F whereas R values before the MM has recently been 10 F from (blue), R for 1610 to the present. Annual means are shown during the MM but are observer. The factor 12.08 ensures that and R is evaluated reconstruction by M. J. Owens and data. Note that the conclusions of this paper do not can be although we note that in the 9300-year Long term variations in solar activity indicators. [2008], F F R (annual means in blue). (a) Group sunspot number, R G (annual mean in red). (b) Zurich sunspot number R Z (annual means in blue). (c) Signed open solar flux F S (annual means from in-situ data and geomagnetic activity data in mauve, the 10-year means of modelled values by Owens and Lockwood (submitted manuscript, 2011) in green). (d) Heliospheric modulation potential Φ (black histogram shows solar cycle means from the composite of Steinhilber et al. [2008], Φ S, cyan and orange show 10-year running means from 10 Be and 14 C, respectively Φ 10Be and Φ 14C).

recorded by the i th observer and k Gi is the calibration factor for that i th observer. The factor 12.08 ensures that R Z and R G are very similar for modern data. Like R Z, R G is evaluated daily but we here use annual averages. Because early observations of the number of individual spots N are relatively rare and often unreliable, R G has been considered more reliable than R Z in early data (see review by Usoskin [2008]). In addition, R G extends almost continuously back to 1610 and so includes the MM. However, an important correction to R G values before the MM has recently been made by Vaquero et al. [2011] using additional observations and this is incorporated in the present paper.

[5] Figure 1a shows the time series of group sunspot number R G for 1610 to the present. Annual means are shown in red and solar cycle means by the black histogram. Solar cycles cannot be detected in R G during the MM but are present in the abundance of the 10 Be cosmogenic isotope during this interval [Beer et al., 1998]. To define cycles within the MM, we here take the times of peak 10 Be abundance from the Dye-3 ice core: when cycles in R G can be defined (i.e. outside the MM), these peaks can be “wiggle-matched” to the corresponding minima in R G to within a dating uncertainty of ±2 years. Using the standard numbering of solar cycles, the grey and white vertical bands in Figure 1 define, respectively, the even- and odd-numbered solar cycles deduced this way. The cycle numbers are then extended back in time (i.e. to negative values) to the start of the R G data. Note that the conclusions of this paper do not depend on the number of cycles defined during the MM; however, this procedure does allow us to assign identification numbers to the cycles in R G before the MM (SC-11 to SC-9) which will be used in this paper. The MM itself covers SC-8 to SC-5, the first signs of a recovery were in SC-4 and this recovery continued during SC-3 and SC-2. Figure 1b shows the corresponding plot for R Z and Figure 1c the signed open solar flux F S deduced from in-situ magnetic field data and geomagnetic activity [Lockwood and Owens, 2011]. The green line in 1(c) shows 10-year running means of the recent model F S reconstruction by M. J. Owens and M. Lockwood (Cyclic loss of open solar flux since 1868: The link to heliospheric, current sheet tilt and implications for the Maunder Minimum, submitted to Journal of Geophysical Research, 2011), based on open flux continuity [Vieira and Solanki, 2010] using R G to quantify the production rate and a loss rate that varies with the cyclic HCS tilt [Owens et al., 2011a]. This reconstruction allows for a base-level coronal mass ejection rate during the solar minima (including the MM). Figure 1d shows the heliospheric cosmic ray modulation potential Φ. The black histogram shows solar cycle means derived from interpolated annual values of the composite (Φ S) generated by Steinhilber et al. [2008] from 10 Be cosmogenic isotope abundances and modern neutron monitor, combined using numerical modeling of the effects of galactic cosmic ray bombardment of Earth’s atmosphere. This composite was based on three independent Φ records: the Vonmoos et al. [2006] reconstruction is used prior to 1645 and is derived from the 10 Be abundance in the Greenland GRIP core; the McCracken et al. [2004] reconstruction is used for 1645–1951 and is derived from the 10 Be abundance in the South Pole core; for after 1951 neutron monitor data are used [Usoskin et al., 2005]. The three records have different temporal resolution and, in order to obtain a homogeneous record, filters were used which generate data that are 25-year means. These are here linearly interpolated to give annual values that are used to compute solar cycle means. The cyan and orange lines show 10-year running means of Φ from 10 Be and 14 C (Φ 10Be and Φ 14C from, respectively, Usoskin et al. [2003] and Muscheler et al. [2007]). We note that although there are clear similarities between the long term variations of different parameters shown in Figure 1, there are also significant differences. Some of these are due to differences in what the parameters are actually a measure of, others may be due to different and varying measurement uncertainties. While the former could cause differences in the inherent predictabilities, that latter would influence the apparent predictabilities, as evaluated from past data.

3. Cycle-to-Cycle Autocorrelation Functions

[6] To study how much the various solar activity indices vary from one cycle to the next, we here take autocorrelation functions (ACFs) of data. In Figure 2a ACFs of annual 25-year running mean values are taken and the lag expressed in units of an average solar cycle length (11.1 yrs). Figure 2a compares the ACFs of R Z (blue), R G (red), Φ S (black) and the modelled open solar flux, F S (green). All show a peak at around 9 solar cycles (the “Gleissberg” period) – although we note that in the 9300-year Φ S data series, this peak is broader and extended to longer periods, possibly by other century-scale variations that have been less evident in recent centuries. It is noticeable that R G and modelled F S have much higher persistence (broader ACF) than Φ S whereas R Z has considerably lower. This may, in part, be due to the shorter data series available (specifically
The predictability, since L22105 (in units of $F_N$ to be $R$ and $R$ is now reduced but still present, and must be for each solar cycle. There is very $\approx$ and $R$ Hong and Billing corrections by Hoyt and Schatten (in blue), $t$ uses $R$ to quantify the source around and in the MM. The red line incorporates $F$ is a little broader and $[e.g., R$, 1998; and for sunspot for $R$ with perfect accuracy, but is zero if no information on data sequence. a small amount, but it has $R$ in Figure 2 and will be, in part, 0.3) as many as 8 cycles into the are similar to that $R$, 2007; E. Cliver, private communication, 2011] that $R$, 1700, 1610 and 7300BC. (b) ACFs for 10-year running means of data after 1700: red, blue and green are for $R_G$, $R_Z$ and $F_S$ and cyan and orange are for the heliospheric modulation potential from $^{10}$Be and $^{14}$C isotope abundances, $\Phi_{10Be}$ and $\Phi_{14C}$, respectively. The dashed red and blue lines in Figure 2b are for $R_G$ and $R_Z$ (respectively) for after 1885 only.

the persistent near zero values in grand minima such as the MM are not included in the $R_Z$ data series); hence in Figure 2b the analysis is repeated using only data from after 1700 so that direct comparisons between the parameters can be made. In addition, the ACFs in 2(b) are taken from 10-year running means of the data. The difference between $R_G$ and $R_Z$ is now reduced but still present, and must be associated with the variability in the observed number of spots, $N$. The ACFs for $\Phi_{10Be}$ and $\Phi_{14C}$ are similar to that for $R_Z$. The modelled $F_S$ uses $R_G$ to quantify the source term and this accounts for the similarity in their ACFs. Although many authors have considered early $R_G$ to be more reliable than $R_Z$ [e.g., Hoyt and Schatten, 1998; Hathaway et al., 2002; Clette et al., 2007; Vaquero, 2007; Usoskin, 2008], it has been suggested [Svalgaard, 2010; E. Cliver, private communication, 2011] that $R_G$ values are systematically too low before 1885. To test for any effect of this on our analysis we have added the ACFs of $R_Z$ and $R_G$ for after 1885 only as a blue and red dashed lines, respectively. For the range of lags over which we can do this for the shorter data interval, the ACF of $R_G$ is a little broader than for the longer data series. However, the ACF of $R_Z$ for after 1885 is much broader than for after 1700, such that it is similar to that for $R_G$. In other words, the pre-1885 data has narrowed the ACF for $R_G$ a small amount, but it has narrowed that for $R_Z$ a great deal and is the major cause of the differences between $R_Z$ and $R_G$. This shows that random measurement errors in the number of spots $N$ and hence $R_Z$ are much higher before 1885. It does not directly tell us about long-term systematic errors.

[7] In Figure 3 (right), the analysis is repeated using the maximum (in red), the cycle mean (in black) and the minimum (in blue) of $R_G$ for each solar cycle. There is very little persistence in the solar minimum values (the ACF falling to zero within 2 solar cycles), but the ACF for the peak and mean values are both similar to those for the corresponding 25-year means in Figure 2a.

[8] Figure 3 (left) analyses what the ACFs in Figure 2b mean for the predictability of the various long-term solar activity indicators in decadal means. We use the procedure developed by Hong and Billing [1999] to quantify forecast predictability $R^2(t)$ (to avoid confusion we here adopt their nomenclature despite the use of $R_Z$ and $R_G$ for sunspot number and group sunspot number) using the ACFs of 10-year running mean data. $R^2(t)$ is derived from the Yule-Walker equations of the autoregressive (AR) model of Walker: it is unity if the parameter can be predicted at a time $t$ with prefect accuracy, but is zero if no information on that parameter is derivable for that time. Note the relatively low predictability of sunspot number $R_G$ (in blue), $R^2(t)$ falling to zero after just 1.5 solar cycles. This is consistent with the narrow ACF for $R_G$ in Figure 2 and will be, in part, due to the inherent unpredictability of $N$, the number of individual spots. However, the dashed lines in Figure 2b indicate a large contribution is due to erroneous values of $N$ early in the $R_Z$ data sequence. $R_G$ and modelled $F_S$ are the most predictable, and cosmogenic isotopes somewhat less so, consistent with the variety of factors which influence the propagation of GCRs through the heliosphere and measurement uncertainties. The plot shows that for $R_G$, $R^2(t)$ still exceeds 0.5 after 4 solar cycles and there is still some predictability ($R^2(t) \approx 0.3$) as many as 8 cycles into the future. For the cosmogenic isotope data $R^2(t)$ exceeds 0.5 for up to 3 solar cycles.

4. The Onset of the Maunder Minimum

[9] Figure 4 shows detail of the variations of $R_G$, $\Phi$ and modelled $F_S$ around and in the MM. The red line incorporates the $R_G$ corrections by Vaquero et al. [2011] (the dotted red line showing the previous best estimates): it can be seen
that these corrections greatly reduce the peak in SC-9 and mean that the decline in peak $R_G$ from $>100$ to $≈0$ at the start of the MM occurs smoothly in just 3 solar cycles. This decline is slightly faster than in the $R_G$ ACFs in Figures 2 and 3 and faster than the average for the end of the 24 GSMs in the past 9300 years inferred by Barnard et al. [2011]. The heliospheric potential estimates $\Phi_S$, $\Phi_{10Be}$ and $\Phi_{14C}$ (in black, cyan and orange, respectively) decline before and during the period of near-zero sunspot number and reach a minima near the end of the GSMi. This late minimum is a strong feature of the South Pole $^{10}$Be record [McCracken et al., 2004], and other records show slightly different behaviour. These differences may be due to site-dependent climatic influences on the $^{10}$Be deposition caused by associated total solar irradiance changes and/or volcanic activity [Field et al., 2009]. Although the temporal behaviour is not exactly the same in the different records, all show the expected increase in radionuclide production during the Maunder minimum due to lower solar activity, which dominates over climate-induced changes in the transport and deposition [Heikkilä et al., 2008; Field et al., 2009]. Some other reconstructions based on the inferred production rate of the $^{14}$C cosmogenic isotope do not show this slow decline persisting to the end of the Maunder minimum [Solanki et al., 2004; Muscheler et al., 2007] whereas others are more similar to the $^{10}$Be records [Muscheler et al., 2007], as shown here in Figure 4 (see review by Usoskin [2008, Figure 12]). The modelled $F_S$ (in green) decays slightly faster than the $\Phi$ estimates, on a timescale similar to $R_G$.

[10] Figure 5 analyses the changes in the $R_G$ values, thereby comparing the recent decline with the descent towards the MM. The plot shows the changes in the maximum (red), mean (black) and minimum (blue) $R_G$ from one cycle to the next, $\Delta R_G$. To reduce the cycle-to-cycle noise and reveal the underlying trends, a 2-point running mean has first been applied to all 3 data sequences. The mean, maximum and minimum behave in similar ways. The changes during the descent into the MM are seen as negative values between SC-11 and SC-8 and the MM itself shows up as zero values until SC-5. The recovery from the MM gives positive $\Delta R_G$ between SC-5 and SC-1. The solar minimum data include the value from the most recent minimum (2009). The open red and black points and dashed lines show the values using the SC24 peak (65 $±$ 10) and mean (32 $±$ 5), respectively, predicted by Owens et al. [2011b]. It is interesting to compare the recent values with the two previous periods of consistently-negative $\Delta R_G$. The first of these is the decline into the MM discussed above and the second is around 1800 (SC2-SC6). The latter did not result in a GSMi, but gave the less-deep Dalton Minimum (DM). The recent $\Delta R_G$ for solar minimum was as negative as immediately before the MM but not as negative as before the DM. The $\Delta R_G$ for cycle means was similar to that before the MM and slightly larger than that prior to the DM. Including the SC24 prediction, the solar-maximum $\Delta R_G$ is larger in amplitude than for before the DM but not quite as large as for the MM. It seems clear that the recent $\Delta R_G$ reveal the onset of a minimum – however, there is no clear indicator of the depth of that minimum (i.e., it is not clear if it will be as deep and as long-lasting as the MM, and hence a GSMi, and could be more like the DM).

[11] The compositing study of past variations of $\Phi$ by Lockwood [2010] found that the chance of $\Phi$ falling below Maunder minimum values is 8% for within the next 40 years, rising to 43% for within the next 100 years. The MM data shown in Figure 4 confirms that a descent in peak group sunspot number as rapid as predicted by Barnard et al. [2011] is certainly possible and has occurred in the past. The open solar flux has, by each solar minimum, migrated to the polar photosphere and is thought to act as the seed field for the solar dynamo at the tachocline [Charbonneau, 2005].

Figure 4. Detailed view of variations around the Maunder minimum. The grey and white vertical bands are as in Figure 1. The red line is the group sunspot number, $R_G$ (the red dotted line is without the correction by Vaquero et al. [2011]). The green line are annual modelled values of signed open solar flux $F_S$ (times 2 to allow use of the same scale as for $R_G$). The black line shows solar cycle means of the composite heliospheric modulation parameter $\Phi_S$ and the cyan and orange are decadal means from $^{10}$Be and $^{14}$C, $\Phi_{10Be}$ and $\Phi_{14C}$ (heliospheric modulation parameter estimates are all divided by 5 to allow use of the same scale as for $R_G$).

Figure 5. Variations of the changes in group sunspot number from one cycle to the next ($\Delta R_G$): (red) the maximum values, (blue) the minimum values, and (black) the solar-cycle averages. All data have been passed through a 2-point running mean before the $\Delta R_G$ values taken. The vertical grey and white bands are as in Figures 1 and 4. The open red and black points with error bars and the dashed red and black lines use the predicted peak and mean value, respectively, for SC24 by Owens et al. [2011b].
This being the case, the decay in $F_S$ heralds a continuing slowing-down of the solar dynamo. The entry into the MM shows that the weak solar cycles were inadequate to prevent a fall into a GSMi. It seems that the HCS remained sufficiently tilted [Owens et al., 2011a; Owens and Lockwood, submitted manuscript, 2011] and/or other open flux loss mechanisms were sufficient to ensure that the decay in open flux continued. The study presented here shows that $R_C$ and $\Phi$ is predictable ($\langle \tilde{R}_C(t) > 0.5 \rangle$ for at least 4 and 3 cycles (respectively) into the future and thus, because the amended group sunspot numbers of Vaquero et al. [2011] show that the decay into MM conditions took less than 3 cycles, it should be possible to predict the onset of GSMi conditions.

[12] Acknowledgments. L.B. was supported by a PhD studentship from the UK Natural Environment Research Council, M.L., M.A.H. and C.J.D. are funded by the Science and Technology Research Council and FS by NCCR climate (Swiss climate research).

[13] The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References


