Heliospheric modulation of galactic cosmic rays during grand solar minima: Past and future variations

M. J. Owens, I. Usoskin, and M. Lockwood

Received 16 July 2012; revised 4 September 2012; accepted 4 September 2012; published 5 October 2012.

[1] Galactic cosmic ray flux at Earth is modulated by the heliospheric magnetic field. Heliospheric modulation potential, \( \Phi \), during grand solar minima is investigated using an open solar flux (OSF) model with OSF source based on sunspot number, \( R \), and OSF loss on heliospheric current sheet inclination. Changing dominance between source and loss means \( \Phi \) varies in- (anti-) phase with \( R \) during strong (weak) cycles, in agreement with \( \Phi \) estimates from ice core records of \(^{10}\)Be concentration, which are in-phase during most of the last 300 years, but anti-phase during the Maunder Minimum. Model results suggest “flat” OSF cycles, such as solar cycle 20 result from OSF source and loss terms temporarily balancing throughout the cycle. Thus even if solar activity continues to decline steadily, the long-term drop in OSF through SC21 to SC23 may plateau during SC24, though reemerge in SC25 with the inverted phase relation. Citation: Owens, M. J., I. Usoskin, and M. Lockwood (2012), Heliospheric modulation of galactic cosmic rays during grand solar minima: Past and future variations, Geophys. Res. Lett., 39, L19102, doi:10.1029/2012GL053151.

1. Introduction

[2] During the space-age, both photospheric magnetic flux and the heliospheric magnetic flux (HMF) have varied approximately in phase with the sunspot cycle. On longer timescales the HMF has been evaluated using geomagnetic activity as a proxy indicator [Lockwood et al., 2009; Lockwood and Owens, 2011]. The HMF modulates the galactic cosmic ray (GCR) flux incident on the terrestrial atmosphere, allowing HMF properties to be inferred from GCR records. Ground-based neutron monitors detect secondary particles produced by GCRs in the atmosphere, allowing the modulation potential, \( \Phi \), to be reproduced over the last half century [Usoskin et al., 2005]. Ice core records of \(^{10}\)Be [Beer, 2000; Usoskin et al., 2003; Usoskin, 2008], a product of the GCR-atmosphere interaction, can also be used to quantify the \( \Phi \), and hence the HMF, on timescales from 1-year to thousands of years. Another cosmogenic isotope, radiocarbon \(^{14}\)C, takes part in the carbon cycle and is stored, for example, in ancient tree trunks. It is very useful for long-term solar activity reconstructions [Solanki et al., 2004], but cannot resolve individual solar cycles [e.g., Bard et al., 1997]. To infer \( \Phi \) from \(^{10}\)Be records, it is necessary to account for the geomagnetic field strength and beryllium deposition rates. The geomagnetic variation is reasonably well characterised for the last few millennia [Donnadini et al., 2010] and can be subtracted [e.g., Masarik and Beer, 2009; Kovaltsov and Usoskin, 2010], whereas spatial variability in deposition rate can be largely mitigated by examining ice cores from well-separated geographic locations. As stronger HMF results in fewer GCRs reaching Earth, the \(^{10}\)Be concentration is generally found to be in anti-phase with the sunspot number variation [Beer, 2000].

[3] During the Maunder Minimum, a grand solar minimum approximately lasting from 1645 to 1700, the Sun was well observed by professional astronomers but displayed remarkably few sunspots [e.g., Hoyt and Schatten, 1998; Vaquero, 2007]. There is evidence that the solar cycle continued throughout this period, as \(^{10}\)Be concentration shows an approximately 11-year periodicity in the Dye3 ice core throughout [Beer et al., 1998]. However, the Maunder Minimum \(^{10}\)Be variation is in phase with the expected sunspot cycle, contrary to expectations and the subsequent behaviour [Usoskin et al., 2001]. Thus there is debate whether the Maunder Minimum \(^{10}\)Be cycling is a solar modulation effect, or the result of a large-scale change in deposition by precipitation resulting from, e.g., the North Atlantic Oscillation [Heikkilä et al., 2009].

[4] The last 5 or 6 solar cycles, which includes the space-age, have displayed higher average sunspot numbers than the rest of the record, suggesting grand solar maximum (GSM) conditions [Solanki et al., 2004], in agreement with geomagnetic [Lockwood et al., 2009; Lockwood and Owens, 2011] and GCR reconstructions of the HMF [McCracken, 2007; Steinhilber et al., 2010]. However, over the last 2 or 3 solar cycles, the solar magnetic field has declined, suggesting the current GSM is ending [Abreu et al., 2008; Lockwood et al., 2009, 2012]. From the GCR record, around 10% of previous GSM exits have resulted in Maunder Minimum-like conditions within 50 years [Steinhilber et al., 2010; Lockwood, 2010; Barnard et al., 2011].

[5] In this study, we use a continuity model to investigate HMF evolution through a declining solar magnetic field and show that cycling \(^{10}\)Be observations during the Maunder Minimum are consistent with a continued solar cycle, despite the phase shift. We then use the model to explore how cycles 24 and 25 could develop given different scenarios for the sunspot cycle.

2. Modelling Heliospheric Modulation Potential

[6] Open solar flux (OSF), the total unsigned magnetic flux threading a heliocentric sphere at the solar wind formation height, can be estimated from both extrapolations of the observed photospheric magnetic field [Wang and Sheeley,
From top to bottom: (first row) imposed sunspot number (R) and (second row) HCS tilt, (third row) modelled unsigned open solar flux (OSF) and (fourth row) heliospheric modulation potential (Φ). Alternate grey/white panels show alternate cycles. Red (blue) vertical lines show times of R (Φ) maxima. R and Φ are in phase during high R, but in anti-phase during low R.

1995] and from in situ measurements of the HMF [e.g., Owens et al., 2008a; Lockwood and Owens, 2009]. Solanki et al. [2000] modelled the solar cycle variation in OSF as a source term, S, assumed to follow the sunspot number (R), and a loss term, L, which allows the OSF to decay with given time constants. Owens and Lockwood [2012] used the observed OSF and observed R to show the fractional OSF loss rate (\(\chi\)) was essentially cyclic over the last century and closely followed the variation in the heliospheric current sheet (HCS) tilt angle. Loss of OSF at regions of high HCS inclination is in agreement with observations of coronal inflows and collapsing loops [Sheeley and Wang, 2001]. A tilted HCS allows differential rotation to force together OSF with opposite polarity, resulting in OSF loss through reconnection [e.g., Owens et al., 2011a].

To examine the effect of a declining solar magnetic field, Figure 1 (first row) shows a simulated R variation using the observed mean variation over cycles 12–23 [Owens et al., 2011b], with amplitude linearly ramped down and up. Times of maximum sunspot number, \(T_R\), are shown here and in Figure 1 (fourth row) as vertical red lines, with shaded areas showing R within 80% of the cycle maximum. Figure 1 (second row) shows the mean variation in \(\chi\) over cycles 12–23 [Owens and Lockwood, 2012], linearly scaled up by a factor 1.74 to match the HCS tilt angle variation. The HCS tilt variation is more asymmetric than the R variation, with a sharper rise, earlier peak, and a more extended decline. The HCS variation is assumed identical each cycle, which holds to first order [Owens and Lockwood, 2012], but the details of the HCS variation can be important for the resulting OSF, as discussed in Section 4.

The R and HCS tilt variations are used as the basis of OSF source and loss terms, in the same manner as by Owens and Lockwood [2012], namely using \(S = a(R + R_0)\) and \(L = \chi OSF\), where \(a = 1 \times 10^{-12} \text{ Wb} \cdot \text{Cr}^{-1}\) (\(C\) = Carrington Rotation) and \(R_0 = 10\). The \(R_0\) term gives OSF production even at times of \(R = 0\), as suggested by coronal mass ejection rates at \(R = 0\) during the recent sunspot minimum [Owens et al., 2008b]. The resulting model OSF and Φ variations are shown in Figures 1 (third row) and (fourth row), respectively. Φ is computed from the OSF and HCS tilt [Alanko-Huotari et al., 2007]. Times of maximum Φ, \(T_\Phi\), are shown here and in Figure 1 (first row) as vertical blue lines, with shaded regions showing times when Φ is within 80% of the peak cycle value.

For cycles with peak R, \(R_{MAX} > 75\) (cycles 1–4 and 14–18 of the plot), the Φ variation is in approximate phase with R. However, as \(R_{MAX}\) drops to around 60, the OSF variation flattens, while Φ peaks later in the cycle. (The exact threshold at which this occurs will depend on the OSF parameters and the form of the source and loss terms.) Stepping down further, \(R_{MAX} \approx 40\) gives OSF in anti-phase with R, which in turn gives a flat Φ variation. For \(R_{MAX} < 30\), grand minimum-like conditions, the Φ variation reemerges in anti-phase with R. This switch is due to the solar cycle variation in OSF primarily resulting from a cycling source term at times of high R, but the cycling loss term at times of low R.

During grand minimum-like conditions, the amplitude of the Φ variation is greatly reduced, despite the relatively high amplitude of the OSF variation. This is because the Alanko-Huotari et al. [2007] form for Φ results in the HCS tilt and OSF combining during times of high R, but canceling during times of low R. Consequently, the solar modulation signal should be more difficult to detect in \(^{10}\)Be records during grand minima. Conversely, lower Φ will result in higher GCR fluxes at Earth, giving elevated \(^{10}\)Be production and hence a weaker solar modulation signal may still be detectable. Furthermore, the Alanko-Huotari et al. [2007] relation is based on space-age observations, so the HCS tilt may not cancel the OSF variation to the same extent during grand solar minima.

The phase between R and Φ can be expressed as \(\Delta T = T_R - T_\Phi\). Black symbols in Figure 2a show |\(\Delta T\)| as a function of \(R_{MAX}\). Circles are times of peak Φ, lines show times when Φ is within 80% of the cyclepeak. |\(\Delta T\)|, rather than \(\Delta T\), is used as Φ peaks near the start/end of the cycle, creating a 1-cycle ambiguity. The dark (light) grey panel shows times when Φ and R are in anti-phase (“transition phase” where Φ is essentially flat and \(T_\Phi\) is difficult to define). For cycles with \(R_{MAX} > 50\), the model Φ and R are in phase to within 1 to 1.5 years.

3. Comparison With \(^{10}\)Be Observations

Model results are now compared with \(^{10}\)Be observations for the period 1610 to 1980, the start of the group sunspot record to the last reliable ice-core observations. Minima in \(^{10}\)Be are assumed to coincide with maxima in Φ (see also Figure 3) [Usoskin et al., 2001]. A 1-year lag has been added to the \(^{10}\)Be-derived \(T_\Phi\) to allow for deposition time [Beer, 2000; Heikkinä et al., 2009]. Uncertainties in atmospheric transport and ice-core dating mean there remains an approximately 1-year uncertainty in the \(^{10}\)Be reconstructions of \(T_\Phi\) [Beer, 2000]. Here we use two independent \(^{10}\)Be records, the Dye3 [Beer et al., 1990] and NGRIP [Berggren et al., 2009] ice cores, more than 1000 km apart. We note that \(^{14}\)C measured in annual tree rings [Stuiver and Quay, 1980] cannot be used in this study as it is unable to resolve individual solar cycles or their exact phase, due to the attenuation and non-
linear delay of the signal caused by the global carbon cycle [Bard et al., 1997; Usoskin and Kromer, 2005]. $R_{\text{MAX}}$ and $T_R$ are determined from the group sunspot record [Hoyt and Schatten, 1998]. During the Maunder Minimum, $T_R$ is estimated by the mean of different reconstructions [Usoskin et al., 2001, Table 1, and references therein]. Sparse sunspot data means there is likely an additional 1-year uncertainty in $T_R$ at these times.

[13] Blue and red lines in Figure 2 show mean $|\Delta T|$ as a function of mean $R_{\text{MAX}}$ using $^{10}$Be records from NGRIP and Dye3 ice core estimates of $T_\Phi$, respectively. Bins are set to contain a minimum of 5 data points. There is a change in $\Phi$-$R_{\text{MAX}}$ phase relation from approximate anti-phase at $R_{\text{MAX}} < 20$, to approximately in-phase at $R_{\text{MAX}} > 40$, in agreement with the model prediction outlined in Section 2. We note that the NGRIP data set was not employed by Usoskin et al. [2001], but it also shows the inverted phase relation during the Maunder Minimum. These results suggest the $^{10}$Be cycling during the Maunder Minimum was a solar modulation effect.

4. Cycles 24 and 25

[14] The model is now used to investigate OSF over the remainder of cycle 24 and cycle 25, given possible scenarios for the sunspot number variation. The black line in Figure 3 (first row) shows the observed sunspot number over the space age. The green line for solar cycle 24 (SC24) shows a forward extrapolation which follows the average variation over past cycles [Owens et al., 2011b]. The red and blue curves for SC24 show one standard deviation above and below the average variation, respectively. For SC25, the green line shows the assumption that SC24 will be repeated exactly, while red (blue) shows an increase (decrease) in $R_{\text{MAX}}$ of 40% compared to SC24. Figure 3 (second row) shows the HCS inclination index [Owens et al., 2011a], factored up by 1200 so that it has a maximum value of 90, allowing comparison with HCS tilt angle. The green line shows the cycle average variation [Owens and Lockwood, 2012].

[15] The black line in Figure 3 (third row) shows the observed OSF variation using 1-day averages of the OMNI [King and Papitashvili, 2005] radial magnetic field. The green/red/blue lines show the OSF model results using the three R scenarios from Figure 3 (first row) and the HCS tilt from Figure 3 (second row). The model is initiated at the minimum between SC23 and SC24. The initial rise in OSF observed in SC24 is well matched. All three curves then show a reasonably small OSF variation over SC24. The green and blue curves for SC24 are quite flat, not dissimilar to SC20 and the $R = 60$ case in Figure 1. Interestingly, even the lowest R estimate does not give a large drop in OSF between cycle minima SC23/SC24 and SC24/SC25. However, in the continued decline scenario for SC25, the OSF does then go on to drop below the SC23 minimum, suggesting that even for a steady fall in R through SC21 to SC25, SC24 could form a temporary plateau in OSF. Such a stalling of the OSF decline was not considered in the solar cycle projections of Barnard et al. [2011].

[16] The model OSF is sensitive not only to the R variation, but also to solar cycle length and the shape of HCS inclination variation. The black line in Figure 2b shows the mean...
fractional OSF loss over cycles 12–23 [Owens et al., 2011b], converted to an equivalent HCS tilt. White and yellow lines show the HCS inclination for SC22 and SC23, respectively, converted to HCS tilt angle [Owens et al., 2011a]. The curves have been normalised, so the integrated fractional OSF loss over each cycle is equal, but the solar cycle phase at which OSF loss occurs is also important. Assuming the mean R variation for SC24, Figure 2c shows the change in OSF between the ends of SC23 and SC24 as a function of solar cycle length and HCS tilt. In all cases longer cycles result in lower OSF, though the magnitude of this effect depends on the form of the HCS variation. The ordering of the yellow, black and white lines show that an elevated HCS during the declining phase of the solar cycle is key to reducing OSF over a solar cycle. Indeed, the large drop in OSF from the end of SC22 to the end of SC23 can be largely attributed to the high HCS inclination through the protracted declining phase of SC23.

Returning to Figure 3 (third row), the red line shows the increasing sunspot number scenario for SC25. The increase in sunspot number between SC24 and SC25 is similar to that in SC20 and SC21: After a flat OSF in SC20 and SC24, both SC21 and SC25 show a slight OSF increase early in the cycle, before a more pronounced increase late in the cycle.

Finally, the black line in Figure 3 (fourth row) shows the neutron-monitor derived heliospheric modulation potential [Usoskin et al., 2005] in black. The red, green and blue lines show \(^{10}\)Be estimated from the model OSF and cyclic HCS tilt angle. The white and yellow lines show \(^{10}\)Be concentration from the Dye3 and NGRIP ice cores, scaled to fit on the same axis, which as expected are in anti-phase with \(\Phi\). The blue, green and red lines are in reasonable agreement with the predictions of Barnard et al. [2011] for probabilities of \(\Phi\) exceeding the given value of \(P[>\Phi]=15\%\), 50\% and 85\%. Given that the red and blue lines are for \(1\sigma\) deviation from the average prediction of \(R\), the continuity modelling presented here is in general agreement with the trends that Barnard et al. [2011] derived from a super-posed epoch analysis of the ends of grand solar maxima.

5. Discussion and Conclusions

A continuity model of open solar flux (OSF) was used to investigate the heliospheric modulation potential, \(\Phi\), during varying solar magnetic field conditions, in particular during grand solar minima. Sunspot number, \(R\), is used as a proxy for the OSF production rate. During strong cycles (i.e., high peak \(R\)), both \(\Phi\) and \(R\) vary closely in phase. However, for weakening cycles (i.e., decreasing peak \(R\)), OSF and \(\Phi\) first go through a transition phase, wherein neither parameter shows much of solar cycle variation, before emerging with a solar cycle variation that is in anti-phase with \(R\). This is because during times of low \(R\), the OSF loss term, which closely follows the heliospheric current sheet (HCS) inclination, dominates over the solar cycle variation in the OSF source.

The modelled behaviour of OSF is in good agreement with the variation of \(^{10}\)Be obtained from ice cores. It suggests that \(^{10}\)Be cycling during the Maunder Minimum was a solar modulation, rather than a local climate effect, as proposed by Usoskin et al. [2001] in the absence of an appropriate model of the heliospheric variation. As the present OSF model provides a natural explanation for the phase change, there is no need to invoke a speculative meteorological influence dominating \(^{10}\)Be data around the Maunder Minimum. Moreover, it would be difficult to explain the similar \(^{10}\)Be variability in two sites separated by >1000 km by the action of an unspecified local effect.

In the model, the change in phase between the heliospheric modulation potential and sunspot maxima is a direct result of assuming the HCS continues to cycle in the same manner, regardless of the strength of the sunspot cycle. Thus we suggest that the solar cycle polarity reversals where unaffected by the relative absence of sunspots during the Maunder Minimum. In turn, this argues that the solar dynamo continues to cycle with an approximately 11-year period throughout grand solar minima, but that the necessary photospheric flux emergence takes place in flux tubes with field too weak, or of too small a diameter, to form sunspots [Spruit, 1977]. The presence of small flux tubes with no dark sunspots may imply that although UV emission from the chromosphere would be lower during grand minima (such as the Maunder Minimum) than during grand maxima (such as recent decades), the visible and infrared emissions from the photosphere would actually be greater than in recent years. During the recent long and extended solar minimum SC23/SC24 observations by the SORCE satellite [Harder et al., 2009] show that the decline in UV was larger than expected but that IR and visible emissions increased. There has been much debate about the veracity of these results but independent ozone measurements do lend some support to them (see discussion by Lockwood [2012]).

Applying the OSF model to future variations, we note a continued decline in solar activity over cycles 24 and 25 may not produce a steady decline in OSF. The model predicts that even a steady decline in sunspot number may result in a plateau in OSF during SC24, before a continued decline in SC25 with the inverted phase relation between \(R\) and OSF.

Acknowledgments. We thank T. Hoeksema of Stanford University for WSO magnetograms. This work was facilitated by the ISSI workshop 233, “Long-term reconstructions of solar and solar wind parameters” organised by L. Svalgaard, E. Clilver, J. Beer and M. Lockwood.

The Editor thanks two anonymous reviewers for assisting in the evaluation of this paper.

References


Beer, J., et al. (1990), Use of \(b=10\) in polar ice to trace the 11-year cycle of solar activity, Nature, 347, 164–166, doi:10.1038/347164a0.


