The heliospheric Hale cycle over the last 300 years and its implications for a "lost" late 18th century solar cycle

J. Space Weather Space Clim., **5**, A30 (2015)
DOI: 10.1051/swsc/2015032

SWSC

RESEARCH ARTICLE OPEN ACCESS

The heliospheric Hale cycle over the last 300 years and its implications for a "lost" late 18th century solar cycle

Mathew J. Owens[1,*], Ken G. McCracken[2], Mike Lockwood[1], and Luke Barnard[1]

[1] Space and Atmospheric Electricity Group, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK
[*]Corresponding author: m.j.owens@reading.ac.uk
[2] 100 Mt. Jellore Lane, Woodlands, NSW, 2575, Australia

Received 21 May 2015 / Accepted 20 August 2015

ABSTRACT

A Hale cycle, one complete magnetic cycle of the Sun, spans two complete Schwabe cycles (also referred to as sunspot and, more generally, solar cycles). The approximately 22-year Hale cycle is seen in magnetic polarities of both sunspots and polar fields, as well as in the intensity of galactic cosmic rays reaching Earth, with odd- and even-numbered solar cycles displaying qualitatively different waveforms. Correct numbering of solar cycles also underpins empirical cycle-to-cycle relations which are used as first-order tests of stellar dynamo models. There has been much debate about whether the unusually long solar cycle 4 (SC4), spanning 1784–1799, was actually two shorter solar cycles combined as a result of poor data coverage in the original Wolf sunspot number record. Indeed, the group sunspot number does show a small increase around 1794–1799 and there is evidence of an increase in the mean latitude of sunspots at this time, suggesting the existence of a cycle "4b". In this study, we use cosmogenic radionuclide data and associated reconstructions of the heliospheric magnetic field (HMF) to show that the Hale cycle has persisted over the last 300 years and that data prior to 1800 are more consistent with cycle 4 being a single long cycle (the "no SC4b" scenario). We also investigate the effect of cycle 4b on the HMF using an open solar flux (OSF) continuity model, in which the OSF source term is related to sunspot number and the OSF loss term is determined by the heliospheric current sheet tilt, assumed to be a simple function of solar cycle phase. The results are surprising; Without SC4b, the HMF shows two distinct peaks in the 1784–1799 interval, while the addition of SC4b removes the secondary peak, as the OSF loss term acts in opposition to the later rise in sunspot number. The timing and magnitude of the main SC4 HMF peak is also significantly changed by the addition of SC4b. These results are compared with the cosmogenic isotope reconstructions of HMF and historical aurora records. These data marginally favour the existence of SC4b (the "SC4b" scenario), though the result is less certain than that based on the persistence of the Hale cycle. Thus while the current uncertainties in the observations preclude any definitive conclusions, the data favour the "no SC4b" scenario. Future improvements to cosmogenic isotope reconstructions of the HMF, through either improved modelling or additional ice cores from well-separated geographic locations, may enable questions of the existence of SC4b and the phase of Hale cycle prior to the Maunder minimum to be settled conclusively.

Key words. Galactic cosmic rays – Space climate – Open solar flux – Solar cycle – Sunspot number

1. Introduction

Solar magnetic activity varies over a huge range of time scales, from seconds to millennia. The most pronounced quasi-periodic variation is the approximately 11-year Schwabe cycle, first identified in sunspot records, but since observed in all associated solar magnetic activity indices of sufficient length and time resolution. Telescopic sunspot observations form the longest near-continuous record of direct solar measurement. As such, they provide useful long-term proxies for solar magnetism, solar irradiance and heliospheric conditions, and hence underpin a wide range of solar, solar-terrestrial and climate studies. (The evidence for effects of solar activity cycles on global and regional climates has recently been reviewed by, e.g., Lockwood 2012).

There are two widely used sunspot records for and around the late 18th century, the original Wolf sunspot number, R_W (Wolf 1861), and the more extensive group sunspot number, R_G (Hoyt & Schatten 1998). While there are a number of ongoing efforts to better quantify uncertainties and even recalibrate for known discontinuities in both records (e.g., Clette et al. 2014; Lockwood et al. 2014b and references therein), both sunspot records generally agree with each other (and independent proxies, such as auroral occurrence and cosmogenic isotope abundance) about the continued existence of the approximately 11-year Schwabe cycle back to at least 1750. During the end of the 18th century, however, the R_W record displays an unusually long cycle, solar cycle 4, spanning 1784–1799. It has been suggested (e.g., Usoskin et al. 2001) that this may be the result of a dearth of observations, which led to Wolf interpolating across two shorter solar cycles and hence losing a solar cycle. Indeed, the group sunspot record does show a small increase around 1794–1799. An extra solar cycle, here referred to as solar cycle 4b (SC4b), would obviously change the numbering of cycles prior to this time. Correct and consistent numbering of Schwabe cycles also underpins empirical cycle-to-cycle sunspot variations (Gnevyshev & Ohl 1948), which are often used as zeroth-order tests of solar dynamo models (Charbonneau et al. 2007). Direct comparison between group sunspot numbers and cosmogenic

isotope abundance is consistent (or at least, not inconsistent) with the existence of SC4b (Usoskin et al. 2001, 2002; Karoff et al. 2014). It should be noted, however, that the relation between GCR intensity at Earth and sunspot number is not a direct one: GCRs are modulated by the total open solar flux (OSF) and structure within heliospheric magnetic field (HMF) (Usoskin 2013), which is related to sunspot number in a non-linear manner.

The latitude of sunspots, a good indication of the phase of the solar cycle independent of cycle amplitude (Waldmeier 1975; Owens et al. 2011b), also appears to support the existence of a cycle 4b (Usoskin et al. 2009). These data, however, are sparse, leading other researchers (e.g., Zolotova & Ponyavin 2011) to suggest the unusual length of cycle 4 and the jump in sunspot latitude in the declining phase was the result of a small burst of sunspot activity in the northern solar hemisphere and not an additional solar cycle.

Hale & Nicholson (1925) noted that the magnetic polarity of leading sunspots reverses every Schwabe cycle, reflecting the dominant solar magnetic pole polarity, and meaning that a complete solar magnetic cycle is approximately 22 years in length (Babcock 1961). Neutron monitor data show that this "Hale cycle" is present in the flux of galactic cosmic rays (GCRs) arriving in near-Earth space: following the reversal of the polar field at sunspot maximum, the GCR intensity increases slowly to the sunspot minimum value in odd-numbered cycles, but rapidly in even-numbered cycles. That is, the GCR intensity immediately prior to solar minimum at the end of odd-numbered solar cycles displays "peaked" temporal variations, while it is "domed" prior to even-numbered cycles (Potgieter & Moraal 1985). This behaviour is primarily attributed to differing GCR drift patterns (Jokipii et al. 1977), though 22-year variations in the heliospheric magnetic field may also have an effect (Cliver & Ling 2001; Rouillard & Lockwood 2004; Thomas et al. 2013). While the NM record started in 1951, the cosmogenic radionuclides (primarily ^{10}Be sequestered in polar ice) provide a record of the GCR intensity near Earth (Beer et al. 2012) that allows us to test the phase of the Hale cycle, and the cycle numbering prior to solar cycle 4.

This study uses the GCR intensity as recorded by cosmogenic isotope measurements to investigate the phase of the Hale cycle before and after the putative cycle 4b. We then estimate the near-Earth heliospheric magnetic field (HMF) variation, with and without the existence of a solar cycle 4b, using the continuity model of OSF of Owens & Lockwood (2012) which assumes OSF source varies with sunspot number and OSF loss is determined by the phase of the solar cycle through heliospheric current sheet variations, as has been predicted theoretically (Owens et al. 2011a). Thus the addition of solar cycle 4b has an effect on the HMF through the OSF loss rate, even though the sunspot number is unchanged.

2. Heliospheric Hale cycle variations

There are a number of direct and indirect proxies for solar magnetic variability before the space age. Geomagnetic observations can be used to accurately reconstruct annual means of heliospheric magnetic field (HMF) intensity and solar wind speed back to 1845 (e.g., Lockwood et al. 2014a), but data before that date are more scarce and not a homogeneous extension of later data.

Figure 1 shows the available solar, GCR, heliospheric and geomagnetic variations at annual resolution over the last 400 years. Panel (a) shows R_G*, the corrected group sunspot number (Lockwood et al. 2014b and references therein), which is based on the International sunspot number and the Hoyt & Schatten (1998) group number, with various adjustments to both records (most notibly for these purposes, Leussu et al. 2013). The Dalton minimum, spanning 1800–1825, immediately follows the longest cycle in the R_G* record, solar cycle 4. This is shown in more detail in the top panel of Figure 4. Panel (b) shows the number of nights per year on which aurora were recorded below 61-degree geomagnetic latitude, N_A, using aurora sighting catalogues for New England (Silverman), Rest of USA (Silverman), Europe and USA (Fritz), Europe and USA (Lovering), global (Krivsky), Sweden (Rubenson), Finland (Nevanlinna), and a previously unpublished one for Great Britain and Ireland (Lockwood et al.). (See Lockwood & Barnard 2015 for full description of all sources.) While the changing numbers of observers at different geomagnetic latitudes makes quantitative assessment of long-term trends in N_A extremely difficult, we note that solar cycle variations are clearly visible, as are distinct longer-term features such as the Dalton minimum. We also note that subsets of the data at different longitudes (specifically from the USA, British Isles/Ireland and Sweden/Finland) all give variations in N_A that are very similar indeed to the overall variation shown here and what differences there are are all explicable in terms of the effects of the changes in the geomagnetic field on the range of magnetic latitudes covered by each region. It thus seems probable that the secular variation in low-latitude aurorae is driven by a slightly different set of solar wind parameters than geomagnetic activity (see discussion by Lockwood 2013). Panel (c) shows the annual ^{14}C production rate estimated by Roth & Joos (2013). While these data are unable to resolve individual solar cycles, longer-term trends resulting from solar activity are present. Panel (d) shows annual values of the average concentration of ^{10}Be in two ice cores, Dye 3 and North GRIP both from Greenland (McCracken & Beer 2015), relative to the 1944–1987 mean computed by Steinhilber et al. (2012). The black line shows annual values, which are quite "noisy", due to 20% standard deviation of the annual ^{10}Be data. Typically, a 1,4,6,4,1-binomial filter is applied to these annual data to account for the spread in possible ^{10}Be production and deposition times, producing the smoothed value shown in red. Panel (e) shows near-Earth heliospheric magnetic field intensity, B, estimated from ^{10}Be concentration, referred to as B(^{10}Be). There is general qualitative agreement between R_G*, N_A and B*(^{10}Be) and B(OSF), discussed further below, both in terms of solar cycles and the long-term trends.

These data are first used to determine the likely phase of the Hale cycle before 1800 and hence whether an additional solar cycle, 4b, is necessary to explain the observations. The left-hand panels of Figure 2 show super-posed epoch plots of solar polarity cycles as a function of solar cycle phase, ϕ. Data are binned into 11 equally spaced ϕ bins, giving approximately annual resolution. Polarity cycles run from the maximum of cycle N ($\phi = -\pi$) to the maximum of cycle $N + 1$ ($\phi = \pi$), when the dominant solar polarity reversals occur. Solar minimum is at $\phi = 0$. Polarity reversal times are approximated as occurring at $\phi = 125°$, i.e., around sunspot maximum (see Figure 1 of Thomas et al. 2013). This approximation is a source of noise in our analysis, as will be discussed later. Polarity cycles for which N is odd are

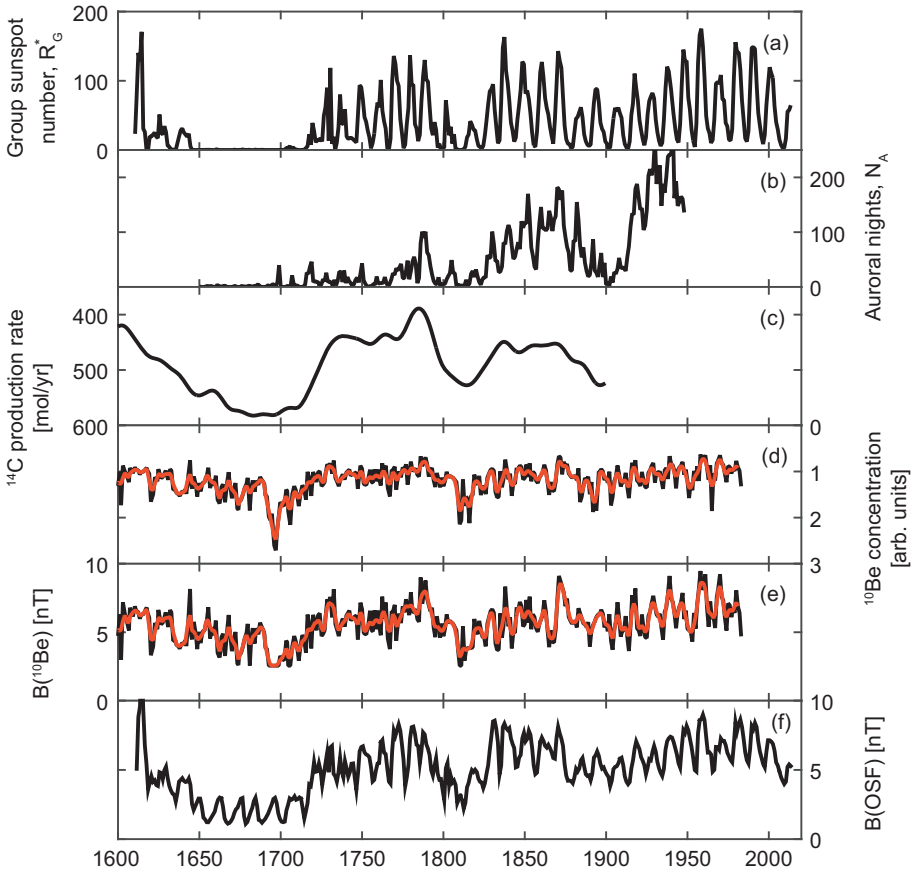

Fig. 1. Annual values of solar, GCR and heliospheric parameters from 1600 to present. (a) Corrected group sunspot number, R_G*; (b) Number of auroral nights below 61° geomagnetic latitude; (c) ^{14}C production rate; (d) Average concentration of ^{10}Be at annual resolution (black) and smoothed using a running 1,4,6,4,1 filter window (red); (e) near-Earth heliospheric magnetic field intensity, B, computed from ^{10}Be concentration, in the same format; (f) B estimated from the open solar flux continuity model using R_G*. Note that the y-axis direction is inverted for panels (c) and (d), as cosmogenic isotope production falls with rising solar activity.

associated with outward polarity magnetic flux at the northern solar pole and are denoted $qA < 0$. These are shown in green. $qA > 0$ polarities, for which N is even, are associated with inward polarity in the northern solar pole and are shown in mauve. All data have been normalised to the maximum and minimum value in each polarity cycle, in order to remove the effect of secular trends. We also remove the effect of solar cycle length by considering variations as a function of ϕ, rather than time. Figure 2a shows R_G* for the period 1800–2014, when the solar cycle numbering is known. Around solar minimum, i.e., $\phi = 0$, the variations are very similar. Around solar maximum, there is a slight divergence, due to even-numbered solar cycles generally exhibiting a higher peak value than odd-numbered solar cycles, often regarded as a simplified Gnevyshev-Ohl rule (e.g., Cliver et al. 1996). Figure 2b shows neutron monitor counts over the period 1960–2014 from the McMurdo station (e.g., Krüger et al. 2008), which is at high geomagnetic latitude and therefore displays strong solar modulation. For the neutron monitor data, the "peaked" (green) and "domed" (mauve) variations are clearly visible. They primarily result from differences during the declining phase of the solar cycle (i.e., phase < 0). Figure 2c shows ^{10}Be concentration over 1800–1983. While these data are noisier than the neutron monitor data, the use of nearly two centuries of data means the same Hale cycle trends are visible. When considering shorter spans of ^{10}Be data, however, the noise introduced by the uncertainty in the polarity reversal times overwhelms the weak signal of the Hale cycle. Thus we also determine the phase of the Hale cycle using superposed epoch analysis of Schwabe cycles (i.e., from solar minimum of cycle N to solar minimum of cycle $N + 1$), as the timings of solar minima can be reliably determined from sunspot latitude (Owens et al. 2011b), when data are available, or minima in annual R_G*, when they are not. The results are shown in the right-hand panels of Figure 2 for odd (red) and even (blue) numbered solar cycles. Figure 2d shows sunspot number over the period 1800-present. The differences between sunspot number during odd- and even-numbered cycles are reduced to a slightly more rapid decline from solar maximum during odd cycles. As the amplitudes and lengths of cycles are normalised, this is consistent with a more rapid rise during odd cycles. Figure 2e shows the normalised neutron monitor counts. Here it is clear that the "domed" waveform of $qA < 0$ polarity cycles is due primarily to the more rapid recovery of GCR flux in the declining phase of even-numbered solar cycles, while the slower recovery in odd-numbered cycles results in the "peaked" waveform. Figure 2f shows the same analysis for ^{10}Be concentration. There is better agreement with the neutron monitor variations than for the polarity cycle analysis.

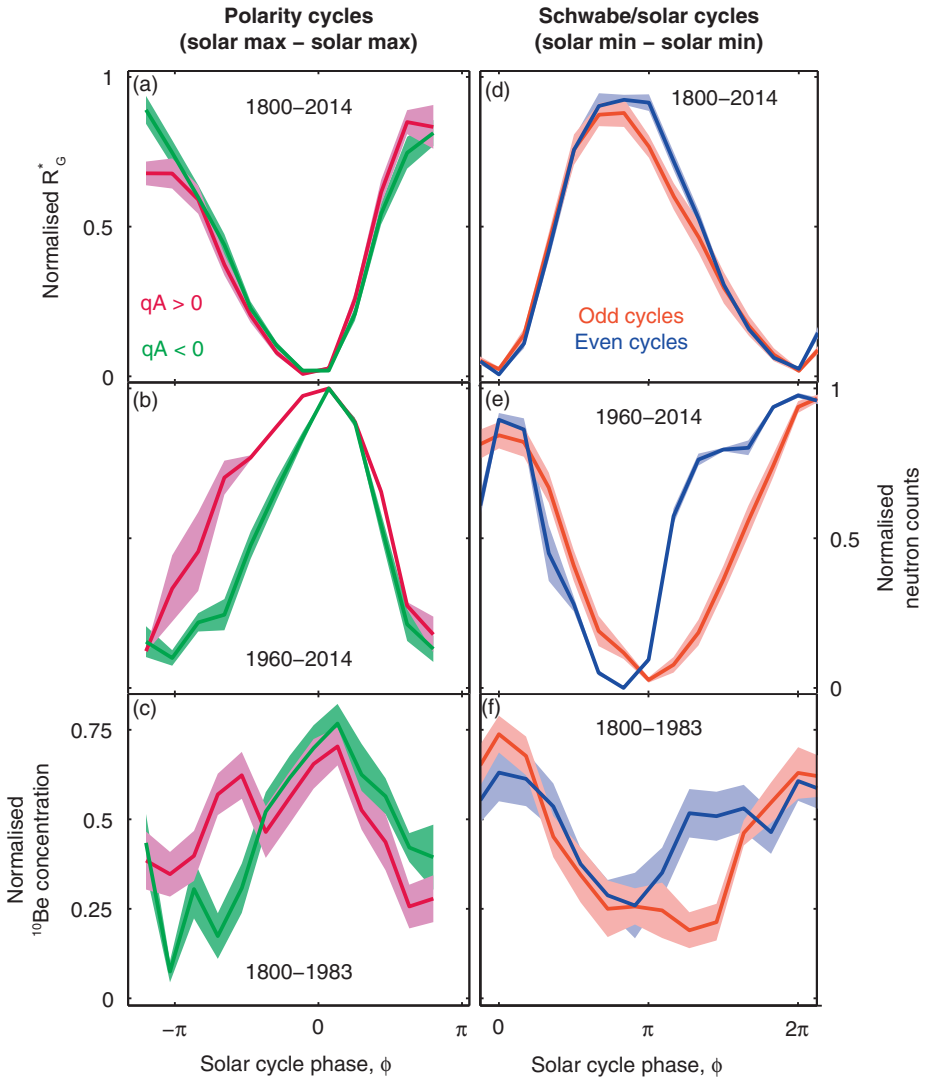

Fig. 2. Super-posed epoch plots of sunspot and GCR measurements over solar polarity cycles (left) and solar cycles (right). Shaded areas show plus-and-minus one standard error in the mean. Data have been normalised to maximum/minimum values within each cycle. Polarity cycles in which the northern solar pole contains outward (inward) polarity magnetic flux, denoted $qA < 0$ ($qA > 0$), are shown in green (mauve). Odd- (even-) numbered solar cycles are shown in red (blue). (a) and (d) show sunspot number, (b) and (e) show neutron monitor counts from McMurdo, while (c) and (f) show ^{10}Be concentration.

Panels (b) and (c) of Figure 3 show the same analysis for annual ^{10}Be concentration over the period 1712–1800. Prior to 1712 is the Maunder minimum. While the solar cycle likely continued throughout this interval (e.g., Owens et al. 2012), the absence of sunspots means that correctly numbering solar cycles is problematic (e.g., Riley et al. 2015; Usoskin et al. 2015). Figure 3b shows analysis using the original solar cycle numbering (i.e., without the addition of SC4b), while Figure 3c shows the cycles renumbered to include SC4b. Both plots show less ordered variations than for the longer 1800–1983 interval (panel a). It is fairly clear, however, that the "no SC4b" scenario provides a between agreement with the post-1800 interval, particularly in the "fast rise" in ^{10}Be concentration for odd-numbered cycles following solar maximum (see also McCracken 2001). Panels (d) and (e) show the whole 1712–1983 interval. These plots share a lot of common data (1800–1983), but here again it is clear that "no SC4b" scenario provides a better reproduction of the "peaked" and "domed" variations seen in neutron monitor data.

The next section considers the implications of solar cycle 4b on the heliospheric magnetic field and compares the two scenarios with the observed $B(^{10}\mathrm{Be})$ and auroral occurrence variations.

3. Reconstructing the heliospheric magnetic field

Variations in the total (unsigned) open solar flux (OSF) can be approximated by a continuity equation (Solanki et al. 2000). Here we use a recent model which quantifies the emergence of open flux from sunspot number, based on an analysis of the occurrence rate and magnetic flux content of coronal mass ejections as a function of sunspot number over recent solar cycles (Owens & Crooker 2006; Owens & Lockwood 2012). The OSF fractional loss rate is varied over the solar cycle with the current sheet tilt, as predicted theoretically by Owens et al. (2011a). Start times of each solar cycle are taken from sunspot numbers, as described in the previous section. The one free

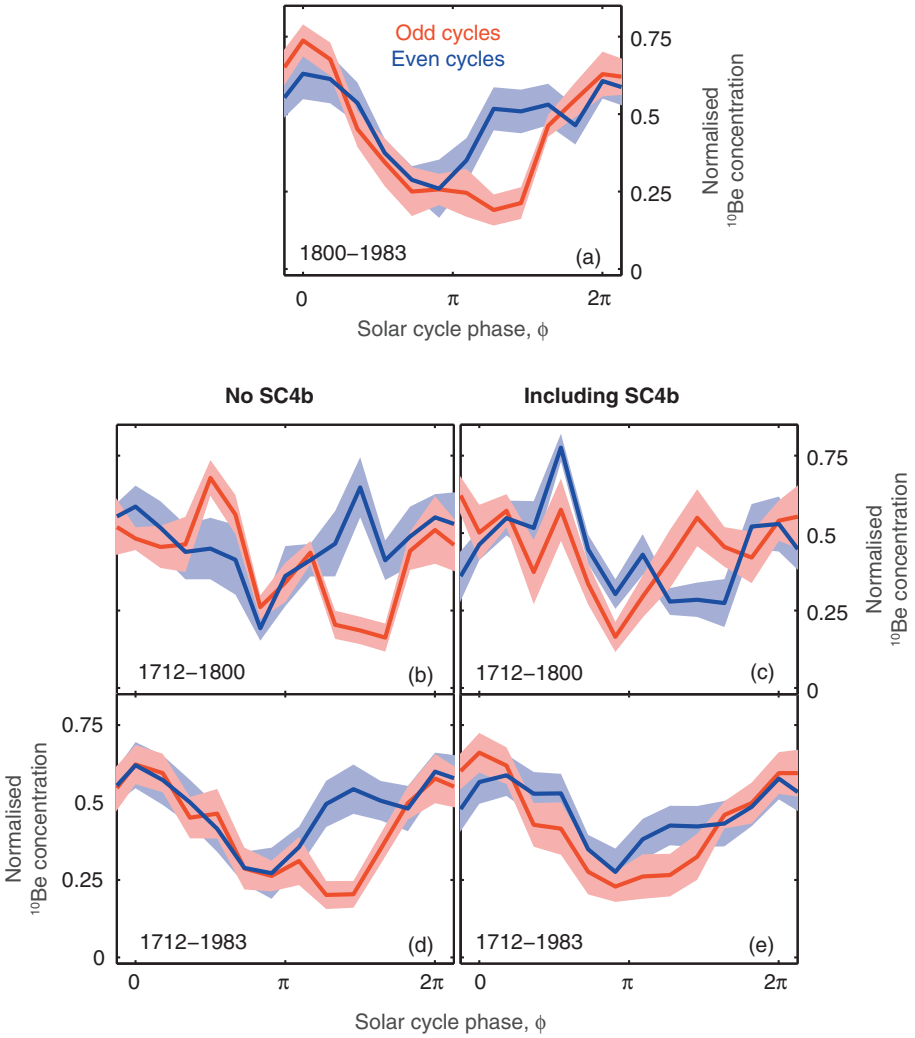

Fig. 3. Super-posed epoch plots of ^{10}Be concentration over solar cycles. Panel (a) is simply Figure 2f, repeated for reference. Panels (b) and (d) show solar cycles numbered without the addition of SC4b, while panels (c) and (f) renumber the cycles prior to 1800 to include SC4b. Panels (b) and (c) only show data from the period 1712 to 1800, while panels (d) and (e) cover 1712–2014 (i.e., include the data shown in panel a).

parameter needed to solve the continuity equation, and so model the open solar flux variation, is obtained by fitting to the open flux reconstruction derived from geomagnetic data for 1845–2012 by Lockwood et al. (2014a). This approach has shown to be successful in reconstructing the observed near-Earth heliospheric magnetic field intensity, B, over both the space age and the extended geomagnetic reconstruction period (Owens & Forsyth 2013; Lockwood & Owens 2014; Usoskin et al. 2015).

Panel (a) of Figure 4 shows the R_G*, the corrected group sunspot number (Lockwood et al. 2014b). This record is used without any further modification to investigate the OSF variation both with and without the presence of SC4b. Panel (b) shows the fractional OSF loss rate, χ, which is purely a function of solar cycle phase (Owens & Lockwood 2012). The red and black lines show χ with and without SC4b, respectively. The start date of SC4b is taken from the sunspot butterfly diagram to be 1793.5, proposed by Usoskin et al. (2009). The pink-shaded area shows the effect of varying this date by ± 1 year.

These two time series are used to constrain the OSF continuity model. The computed OSF is converted to near-Earth magnetic field intensity, B(OSF), shown in panel (c), using Equation 12 of Lockwood et al. (2014a). Consider first the results without cycle 4b (black): R_G*, which acts as the OSF source term, increases rapidly through the rise phase of cycle 4, 1785–1789. But the near-simultaneous increase in OSF loss, χ, means the resulting B(OSF) peaks slightly earlier (1787) than the sunspot maximum (1789). It reaches a peak magnitude of 7.7 nT, just lower than the previous two solar maxima in SC2 and SC3. B(OSF) then declines steadily until 1793, when the small rise in R_G* causes a similar rise in B(OSF). In the B(OSF) time series, this looks rather like a "mini solar cycle" spanning 1793–1798, but it is important to note that this variation does not arise from an actual solar cycle (i.e., a solar polarity reversal and associated HCS rotation), just a "blip" in sunspot number.

The addition of SC4b (red) causes the OSF loss rate, χ, to rise faster at the start of cycle 4. Thus the predicted B rise is slower than the scenario without SC4b. The addition of

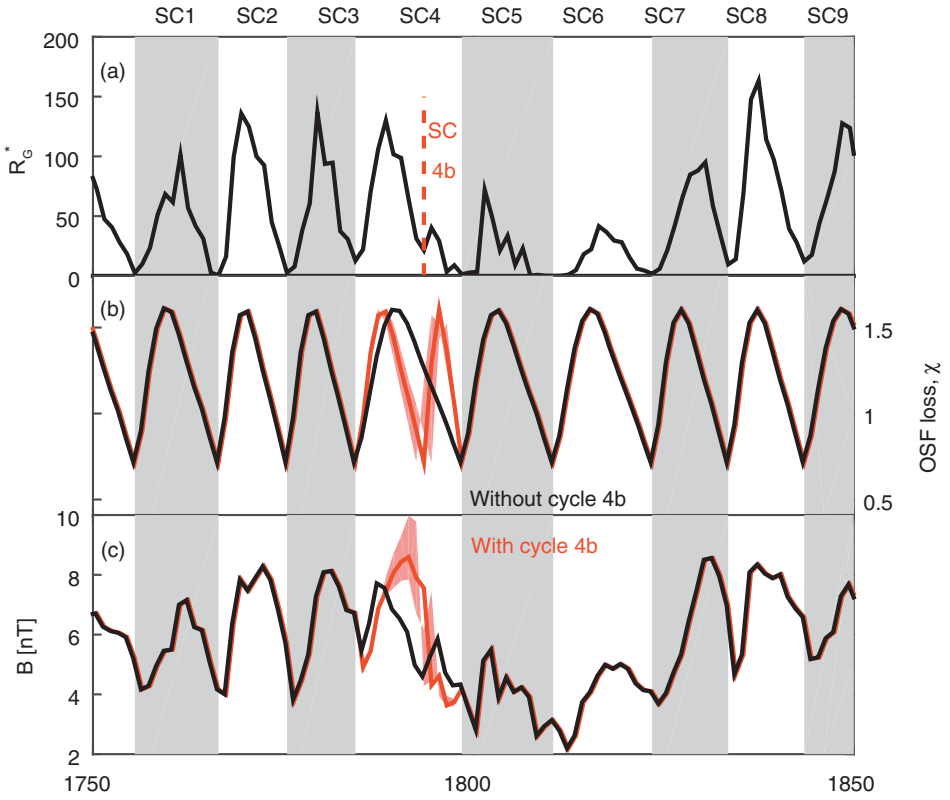

Fig. 4. (a) Detail of annual corrected group sunspot number, R_G*, for the interval 1750–1850. Alternating white/grey panels show solar cycles (without the putative extra cycle 4b). The proposed start of solar cycle 4b is shown by the red dashed line. (b) Fractional open solar flux loss rate, χ, which is assumed to be purely a function of solar cycle phase. Red (black) lines show χ with (without) the addition of solar cycle 4b. Pink-shaded areas show the range of values obtained by varying the SC4b start date by ± 1 year. (c) Near-Earth magnetic field magnitude, *B*, from the modelled OSF. In the same format as panel (b).

SC4b also results in χ falling rapidly just as R_G* peaks in 1789. Consequently, *B*(OSF) continues to rise even as R_G* falls, peaking late in SC4 (1792) with a value of 8.6 nT, higher than previous two solar maxima. In 1793, the increase in χ associated with the onset of SC4b causes a precipitous drop in *B*(OSF) and completely negates the small rise in R_G* at that time. Thus, contrary to expectations, the addition of cycle 4b effectively removes the "mini solar cycle" variation in B(OSF) from 1793 to 1798, while increasing the magnitude, but delaying the timing, of the solar cycle 4 peak in *B*(OSF). The next step is to compare these two scenarios with the available observations from the period.

4. Comparison with ^{10}Be, ^{14}C and auroral records

The blue line in Figure 5a shows annual number of auroral nights, N_A, smoothed by a 1,4,6,4,1-binomial filter to give N_A*. There are no statistically meaningful differences in linear correlation coefficients between N_A* and *B**(OSF) with/without cycle 4b present (see also Table 1). As discussed earlier, long-term trends in historical aurora data are subject to large baseline drift, making quantitative cycle-to-cycle comparisons problematic. The timing of individual aurora events, however, is precise, thus trends within an individual solar cycle are likely to be the result of solar variations. We note that the timing of the SC4 N_A* peak is in agreement with the "no SC4b" scenario. Conversely, there is no secondary increase in N_A* at the end of SC4, with the continuous decline into the Dalton minimum in better agreement with the SC4b scenario. Similarly, the amplitude of the N_A* peak in SC4 relative to the preceding cycles is also more consistent with the SC4b scenario.

The blue line in Figure 5b shows *B* estimated from ^{10}Be observations, *B*(^{10}Be) (McCracken & Beer 2015) passed through the same 1,4,6,4,1-binomial filter to compute the smoothed *B**. The standard deviation of the filtered data is estimated to be 0.6 nT and the uncertainty in ^{10}Be dating can be up to several years. Over the 1750–1850 interval, there is general agreement between *B**(OSF) (black and red) and *B**(^{10}Be) (blue) in terms of the number of solar cycles and the long-term trends, particularly the existence of the "Dalton minimum" through ~1800–1825, i.e., SC5 and SC6. It should be noted, however, that the details of individual solar cycles in *B**(OSF) and *B**(^{10}Be) do not perfectly agree, either in terms of magnitude or waveform (see also Table 1). Therefore, care must be taken not to "over interpret" the results for SC4 and here we make primarily qualitative comparisons, summarised in Table 2. We also note that, as for N_A, there is no statistically meaningful difference in the linear correlation coefficients between *B*(^{10}Be) and *B*(OSF) with/without the addition of SC4b.

The main *B**(^{10}Be) peak in SC4 occurs relatively early (1789), which is in better agreement with *B**(OSF) for the "no SC4b" scenario. The amplitude of the SC4 *B**(^{10}Be) peak, however, is higher than either of the previous three cycles, in agreement with *B**(OSF) when SC4b is included.

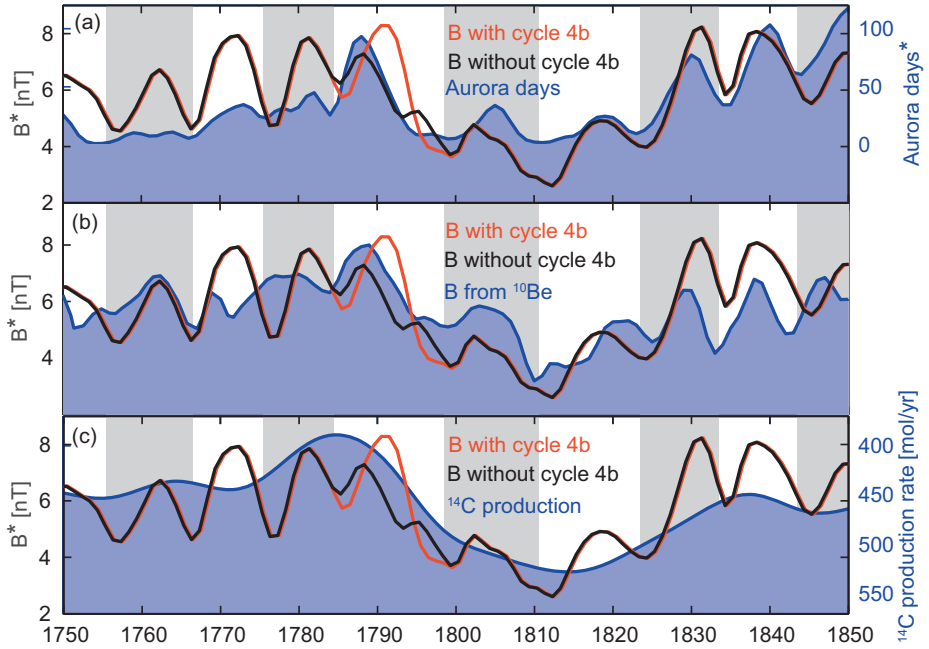

Fig. 5. Annual variations from 1750 to 1850, smoothed using a 1,4,6,4,1-binomial filter. Red (black) lines show output of the OSF continuity model with (without) the addition of solar cycle 4b. Blue lines and the shaded regions show (a) annual auroral nights, N_A (b) $B^*(^{10}Be)$ and (c) ^{14}C production rate, with y-axis inverted, to follow the B(OSF) variations.

Table 1. Linear (Pearson) correlation coefficients over the period 1750–1850. As $N = 100$, all correlations have p-values < 0.00001.

	B^*(OSF) no SC4b	B^*(OSF) with SC4b
$B^*(^{10}Be)$	0.600	0.635
N_A^*	0.683	0.698

Looking towards the end of SC4, there is some suggestion of a small rise in $B^*(^{10}Be)$ prior to the onset of SC5, as is found in the "no cycle 4b" B^*(OSF) time series. These findings are in broad agreement with the photospheric flux-transport modelling of Jiang et al. (2011), though they find the addition of SC4b to have a much larger and lasting effect on OSF.

Finally, Figure 5c compares the two B^*(OSF) reconstructions with the ^{14}C production rate (Roth & Joos 2013). While individual solar cycles cannot be resolved, the general long-term trends are present. The decrease in ^{14}C production (note inverted y-axis) around 1780–1790 is somewhat suggestive of SC4 (~1790) being larger than SC2 (~1770), providing support for the SC4b scenario.

These qualitative results are summarised in Table 2.

5. Summary and discussion

The approximately 22-year Hale cycle means that odd- and even-numbered Schwabe (solar/sunspot) cycles are associated with different patterns in the galactic cosmic ray intensity. We have used recent ^{10}Be concentration measurements (McCracken & Beer 2015) to demonstrate that this pattern can be observed in average values back to the end of the Maunder minimum around 1712. The phase of the Hale cycle in ^{10}Be data prior to 1800 is more consistent with the current numbering of solar cycles, without the addition of a cycle "4b". However, the uncertainties in the ^{10}Be data are presently too high to draw strong conclusions. We also note that Usoskin et al. (2001) used the sunspot record to conclude SC4b was required to restore the broken Gnevyshev-Ohl rule before 1800.

Annual ^{10}Be data can be extended back to the 14th century, meaning the phase of the Hale cycle should be observable prior to the start of the Maunder minimum. If the solar cycle is assumed to have continued through the Maunder minimum (Beer et al. 1998; Owens et al. 2012), this method could be used to provide the first reliable estimate of the number of solar cycles during the Maunder minimum.

We have also investigated the effect of SC4b on the heliospheric magnetic field, B, using a continuity model of open

Table 2. Properties of the modelled open solar flux variations with/without the addition of SC4b compared with observed ^{10}Be, ^{14}C and auroral variations.

	B^*(OSF) no SC4b	B^*(OSF) with SC4b	$B^*(^{10}Be)$	N_A^*	^{14}C production
Time of primary SC4 peak	1788	1790–1792	1789	1788	–
Amplitude of SC4 peak relative to SC2 and SC3	Smaller	Larger	Larger	Larger	Larger
Secondary peak in 1794–1799?	Yes	No	Yes	No	–

solar flux (OSF), in which the OSF source terms vary with group sunspot number and the OSF loss term varies with heliospheric current sheet (HCS) inclination. Even without any alteration to the sunspot record, the addition of a cycle "4b" in 1794–1799 makes considerable difference to the reconstructed *B*. This is due to solar cycle 4b (SC4b) requiring an extra reversal of the polar magnetic fields, which necessitates an additional cycle in the orientation of the heliospheric current sheet and thus changing the rate at which open solar magnetic flux is lost (Owens et al. 2011a). We find that without the addition of SC4b, the period 1784–1799 contains two *B* peaks, which are a direct result of the two peaks in group sunspot number and hence OSF source. Somewhat surprisingly, the addition of SC4b means the period 1784–1799 contains only a single *B* peak: the second *B* peak is removed due to the increase in group sunspot number being completely offset by the additional increase in OSF loss. The timing and magnitude of the main SC4 *B* peak is also modified. Without SC4b, the main SC4 peak in *B* occurs early in the solar cycle and is smaller in magnitude than the previous two solar cycles. The addition of SC4b delays the peak in *B* until later in SC4 and elevates its magnitude above the previous two solar cycles.

Available observations at Earth against which these two model scenarios can be compared are limited to auroral nights and cosmogenic isotope data. Despite the significant differences in model *B* with/without the addition of SC4b, the inherent uncertainties in the observations mean it is not possible to conclusively differentiate between the two possible scenarios. We do note, however, that the addition of SC4b provides marginally better agreement with the qualitative features in the observed time series. While there is unlikely to be much scope for significant improvement of the 18th century auroral record, *B* reconstructions from cosmogenic isotope data may in the future be able to determine the existence (or otherwise) of SC4b. This could be enabled through both more sophisticated modelling of the heliospheric modulation of galactic cosmic rays and the subsequent deposition of ^{10}Be, and additional ice cores to improve the global nature of the ^{10}Be measurement.

Of the two analytical methods used herein, that based on detection of the phase of the Hale cycle is based on substantially more ^{10}Be data than the curve matching in the case of the OSF models. The phase of the Hale cycle between 1712 and the present argues against the hypothesis that solar cycle 4 is, in reality, a double cycle.

Acknowledgements. We thank the Bartol Research Institute of the University of Delaware for the neutron monitor data from McMurdo, which is supported by NSF Grant No. ATM-0527878. MO, ML and LB are part-funded by Science and Technology Facilities Council (STFC) Grant No. ST/M000885/1. MO acknowledges support from the Leverhulme Trust through a Philip Leverhulme Prize. KGMcC acknowledges the consistent support he has received since 2005 from the International Space Science Institute, Bern, Switzerland. The editor thanks Ilya Usoskin and an anonymous referee for their assistance in evaluating this paper.

References

Babcock, H.W.. The topology of the Sun's magnetic field and the 22-year cycle. *Astrophys. J.*, **133**, 527–587, 1961.

Beer, J., K.G. McCracken, and R. von Steiger. *Cosmogenic Radionuclides: Theory and applications in the terrestrial and space environments*, Berlin, Springer, 2012.

Beer, J., S. Tobias, and N. Weiss. An active sun throughout the maunder minimum. *Sol. Phys.*, **181**, 237–249, 1998.

Charbonneau, P., G. Beaubien, and C. St-Jean. Fluctuations in Babcock-Leighton dynamos. II. Revisiting the Gnevyshev-Ohl rule. *Astrophys. J.*, **658**, 657, 2007.

Clette, F., L. Svalgaard, J.M. Vaquero, and E.W. Cliver. Revisiting the sunspot number. *Space Sci. Rev.*, **186**, 35–103, 2014.

Cliver, E.W., V. Boriakoff, and K.H. Bounar. The 22-year cycle of geomagnetic and solar wind activity. *J. Geophys. Res. [Space Phys.] (1978–2012)*, **101**, 27091–27109, 1996.

Cliver, E.W., and A.G. Ling. 22 Year patterns in the relationship of sunspot number and tilt angle to cosmic-ray intensity. *Astrophys. J. Lett.*, **551**, L189–L192, 2001, DOI: 10.1086/320022.

Gnevyshev, M., and A. Ohl. On the 22-year cycle of solar activity. *Astron. Zh.*, **25**, 18, 1948.

Hale, G.E., and S.B. Nicholson. The law of sun-spot polarity. *Astrophys. J.*, **62**, 270–300, 1925.

Hoyt, D.V., and K.H. Schatten. Group sunspot numbers: A new solar activity reconstruction. *Sol. Phys.*, **181**, 491–512, 1998.

Jiang, J., R.H. Cameron, D. Schmitt, and M. Schüssler. The solar magnetic field since 1700-II. Physical reconstruction of total, polar and open flux. *A&A*, **528**, A83, 2011.

Jokipii, J.R., E.H. Levy, and W.B. Hubbard. Effects of particle drift on cosmic-ray transport. I – General properties, application to solar modulation. *Astrophys. J.*, **213**, 861–868, 1977, DOI: 10.1086/155218.

Karoff, C., F. Inceoglu, M.F. Knudsen, J. Olsen, and A. Fogtmann-Schulz. The lost sunspot cycle: New support from ^{10}Be measurements. *A&A*, **575**, A77, 2014, DOI: 10.1051/0004-6361/201424927.

Krüger, H., H. Moraal, J. Bieber, J. Clem, P. Evenson, K. Pyle, M. Duldig, and J. Humble. A calibration neutron monitor: Energy response and instrumental temperature sensitivity. *J. Geophys. Res.*, **113**, A08101, 2008, DOI: 10.1029/2008JA013229.

Leussu, R., I.G. Usoskin, R. Arlt, and K. Mursula. Inconsistency of the Wolf sunspot number series around 1848. *A&A*, **559**, A28, 2013, DOI: 10.1051/0004-6361/201322373.

Lockwood, M. Solar influence on global and regional climates. *Surv. Geophys.*, **33**, 503–534, 2012, DOI: 10.1007/s10712-012-9181-3.

Lockwood, M. Reconstruction and prediction of variations in the open solar magnetic flux and interplanetary conditions. *Living Rev. Sol. Phys.*, **10**, 4, 2013, DOI: 10.12942/lrsp-2013-4.

Lockwood, M., and L. Barnard. An arch seen in the UK: a new catalogue of auroral observations made in the British Isles and Ireland, *Astron. Geophys.*, **56**, 4–25, 2015.

Lockwood, M., H. Nevanlinna, L. Barnard, M.J. Owens, R.G. Harrison, A.P. Rouillard, and C.J. Scott. Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr – Part 4: Near-Earth solar wind speed, IMF, and open solar flux. *Ann. Geophys.*, **32**, 383–399, 2014a, DOI: 10.5194/angeo-32-383-2014.

Lockwood, M., and M.J. Owens. Centennial variations in sunspot number, open solar flux and streamer belt width: 3. Modeling. *J. Geophys. Res.*, **119**, 5193–5209, 2014, DOI: 10.1002/2014JA019973.

Lockwood, M., M.J. Owens, and L. Barnard. Centennial variations in sunspot number, open solar flux, and streamer belt width: 1. Correction of the sunspot number record since 1874. *J. Geophys. Res.*, **119**, 5172–5182, 2014b, DOI: 10.1002/2014JA019970.

McCracken, K. Variations in the production of ^{10}Be due to the 11 year modulation of the cosmic radiation, and variations in the vector geomagnetic dipole. *Proc. Int. Conf. Cosmic Rays*, 4129–4132, 2001.

McCracken, K.G., and J. Beer. The Annual Cosmic-radiation Intensities 1391–2014; the annual Heliospheric Magnetic Field Strengths 1391–1983; and identification of solar cosmic ray events in the cosmogenic record 1800–1983. *Sol. Phys.*, 2015, in press.

Owens, M.J., and N.U. Crooker. Coronal mass ejections and magnetic flux buildup in the heliosphere. *J. Geophys. Res.*, **111**, A10104, 2006, DOI: 10.1029/2006JA011641.

Owens, M.J., N.U. Crooker, and M. Lockwood. How is open solar magnetic flux lost over the solar cycle? *J. Geophys. Res.*, **116**, A04111, 2011a, DOI: 10.1029/2010JA016039.

Owens, M.J., and R.J. Forsyth. The heliospheric magnetic field. *Living Rev. Sol. Phys.*, **10**, 5, 2013, DOI: 10.12942/lrsp-2013-5.

Owens, M.J., and M. Lockwood. Cyclic loss of open solar flux since 1868: the link to heliospheric current sheet tilt and implications for the Maunder Minimum. *J. Geophys. Res.*, **117**, A04102, 2012, DOI: 10.1029/2011JA017193.

Owens, M.J., M. Lockwood, C.J. Davis, and L. Barnard. Solar cycle 24: implications for energetic particles and long-term space climate change. *Geophys. Res. Lett.*, **38**, L19106, 2011b, DOI: 10.1029/2011GL049328.

Owens, M.J., I. Usoskin, and M. Lockwood. Heliospheric modulation of galactic cosmic rays during grand solar minima: past and future variations. *Geophys. Res. Lett.*, **39**, L19102, 2012, DOI: 10.1029/2012GL053151.

Potgieter, M., and H. Moraal. A drift model for the modulation of galactic cosmic rays. *Astrophys. J.*, **294**, 425–440, 1985.

Riley, P., R. Lionella, J.A. Linker, et al. Inferring the structure of the solar corona and inner heliosphere during the maunder minimum using global thermodynamic magnetohydrodynamic simulations. *Astrophys. J.*, **802**, 105, 2015, DOI: 10.1088/0004-637X/802/2/105.

Roth, R., and F. Joos. A reconstruction of radiocarbon production and total solar irradiance from the Holocene ^{14}C and CO_2 records: implications of data and model uncertainties. *Climate of the Past*, **9**, 1879–1909, 2013.

Rouillard, A., and M. Lockwood. Oscillations in the open solar magnetic flux with a period of 1.68 years: imprint on galactic cosmic rays and implications for heliospheric shielding. *Ann. Geophys.*, **22**, 4381–4395, 2004, DOI: 1432-0576/ag/2004-22-4381.

Solanki, S.K., M. Schüssler, and M. Fligge. Evolution of the Sun's large-scale magnetic field since the Maunder minimum. *Nature*, **408**, 445–447, 2000, DOI: 10.1038/35044027.

Steinhilber, F., J.A. Abreu, J. Beer, et al. 9,400 years of cosmic radiation and solar activity from ice cores and tree rings. *Proceedings of the National Academy of Science*, **109**, 5967–5971, 2012, DOI: 10.1073/pnas.1118965109.

Thomas, S.R., M.J. Owens, and M. Lockwood. The 22-year hale cycle in cosmic ray flux – evidence for direct heliospheric modulation. *Sol. Phys.*, **289**, 407–421, 2013, DOI: 10.1007/s11207-013-0341-5.

Usoskin, I., R. Arlt, E. Asvestari, et al. The Maunder minimum (1645–1715) was indeed a Grand minimum: A reassessment of multiple datasets. *A&A*, 2015, in press, DOI: 10.1051/0004-6361/201526652.

Usoskin, I.G. A history of solar activity over millennia. *Living Rev. Sol. Phys.*, **10**, 2013, DOI: 10.12942/lrsp-2013-1.

Usoskin, I.G., K. Mursula, R. Arlt, and G.A. Kovaltsov. A solar cycle lost in 1793–1800: Early sunspot observations resolve the old mystery. *Astrophys. J. Lett.*, **700**, L154, 2009.

Usoskin, I.G., K. Mursula, and G.A. Kovaltsov. Was one sunspot cycle lost in late XVIII century? *A&A*, **370**, L31–L34, 2001.

Usoskin, I.G., K. Mursula, and G.A. Kovaltsov. Lost sunspot cycle in the beginning of Dalton minimum: New evidence and consequences. *Geophys. Res. Lett.*, **29**, 2183, 2002.

Waldmeier, M. The beginning of a new cycle of solar activity. *Nature*, **253**, 419, 1975.

Wolf, R.A. Abstracts of his latest results, *Mon. Not. R. Astron. Soc.*, **21**, 77, 1861.

Zolotova, N., and D. Ponyavin. Enigma of the solar cycle 4 still not resolved. *Astrophys. J.*, **736**, 115, 2011.

Cite this article as: Owens M, McCracken K, Lockwood M & Barnard L. The heliospheric Hale cycle over the last 300 years and its implications for a "lost" late 18th century solar cycle. *J. Space Weather Space Clim.*, **5**, A30, 2015, DOI: 10.1051/swsc/2015032.