Centennial variations in sunspot number, open solar flux and streamer belt width: 3. Modelling

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Space Physics

RESEARCH ARTICLE
10.1002/2014JA019973

This article is a companion to *Lockwood et al.* [2014] doi:10.1002/2014JA019970 and *Lockwood et al.* [2014] doi:10.1002/2014JA019972.

Key Points:
- Streamer belt and coronal hole open solar flux modeled
- Streamer belt width is higher for lower open solar flux
- Earth remained within streamer belt for duration of the Maunder minimum

Supporting Information:
- Readme
- JGR_Lockwood_OSF3_supplementarydata

Correspondence to:
M. Lockwood,
m.lockwood@reading.ac.uk

Citation:
Lockwood, M., and M. J. Owens (2014), Centennial variations in sunspot number, open solar flux and streamer belt width: 3. Modeling, *J. Geophys. Res. Space Physics*, *119*, doi:10.1002/2014JA019973.

Received 12 MAR 2014
Accepted 10 JUN 2014
Accepted article online 12 JUN 2014

Centennial variations in sunspot number, open solar flux and streamer belt width: 3. Modeling

M. Lockwood[1] and M. J. Owens[1]

[1]Department of Meteorology, University of Reading, UK

Abstract From the variation of near-Earth interplanetary conditions, reconstructed for the mid-19th century to the present day using historic geomagnetic activity observations, Lockwood and Owens (2014) have suggested that Earth remains within a broadened streamer belt during solar cycles when the Open Solar Flux (OSF) is low. From this they propose that the Earth was immersed in almost constant slow solar wind during the Maunder minimum (c. 1650–1710). In this paper, we extend continuity modeling of the OSF to predict the streamer belt width using both group sunspot numbers and corrected international sunspot numbers to quantify the emergence rate of new OSF. The results support the idea that the solar wind at Earth was persistently slow during the Maunder minimum because the streamer belt was broad.

1. Introduction

1.1. Reconstructions of Open Solar Flux from Historic Geomagnetic Activity Data

The (signed) open solar flux (OSF) F_S was first reconstructed by *Lockwood et al.* [1999] from geomagnetic data for 1868 onward using the aa activity index and its 27 day recurrence index. *Svalgaard and Cliver* [2007] noted that both the near-Earth interplanetary magnetic field (IMF), B, and solar wind speed, V_{SW}, could both be reconstructed using pairs of indices that have different dependencies on these near-Earth interplanetary parameters. This was exploited by *Rouillard et al.* [2007], who used the separate B and V_{SW} estimates with Parker spiral theory (and a statistical allowance for excess flux effect, discussed below in section 1.4) to reconstruct the open solar flux variation. The development of these reconstructions has been reviewed by *Lockwood* [2013].

A recent analysis has been presented in a series of four papers by *Lockwood et al.* [2013a, 2013b, 2014c, 2014d] which used four pairings of magnetic indices to reconstruct B, V_{SW}, and F_S from 1845 to the present day with full analysis of the uncertainties in each case. These uncertainties were computed at the 2σ level using a Monte-Carlo fitting technique and allow for regression fit errors, the IMF orientation factor (relevant because geomagnetic activity responds to the southward component of the IMF), in situ satellite measurement errors, geomagnetic instrument intercalibration errors, geomagnetic observatory calibration drift, and site changes. The resulting variations are shown in Figure 1 (from *Lockwood et al.* [2014d]). The grey areas mark the ±2σ uncertainty band in each case, and these uncertainties are largest for the earliest data (because of geomagnetic observation uncertainties) but are remarkably small for all years. Note that the annual means of the in situ data (the blue dots) sometimes lie outside the reconstruction uncertainty band because of the effect of data gaps in the satellite measurements: these are not a factor in the reconstructions which are based on correlations between annual means with piecewise removal of geomagnetic data at times corresponding to satellite data gaps [*Finch and Lockwood*, 2007]. *Lockwood et al.* [2014d] show that the reconstruction of F_S using the method of *Lockwood et al.* [1999] also lies within the uncertainty band in Figure 1f.

Thus, we can reconstruct these three parameters with a high degree of certainty back to 1845, just 13 years after the establishment by Gauss of the first magnetic observatory in Göttingen. It is important to make a clear distinction between OSF which, like sunspot number, is an indicator of the global state of the Sun, and B and V_{SW} which are samples of the heliosphere at a single point in the ecliptic plane.

1.2. Definition and Measurement of Open Solar Flux

It is important to work to a clear, consistent, and measurable definition of "open solar flux." As discussed by *Lockwood et al.* [2009c], any schematic of heliospheric field lines that is of finite size allows a false

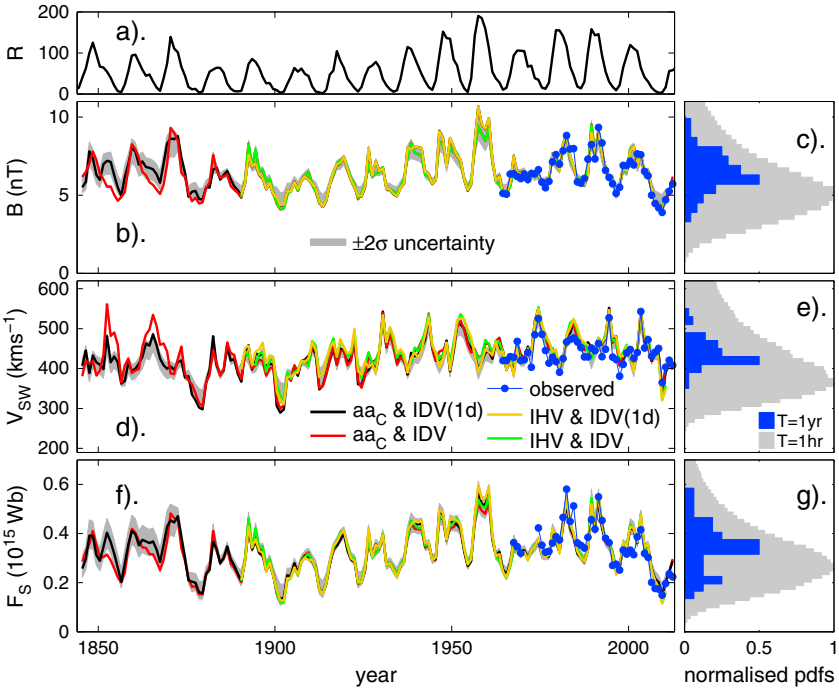

Figure 1. Reconstructions of annual means of solar and near-Earth heliospheric parameters: (a) the international sunspot number, R_C, with the correction defined in Paper 1 [*Lockwood et al.*, 2014a]; (b) the near-Earth IMF, *B*; (d) the near-Earth solar wind speed, V_{SW}; and (f) the signed open solar flux, F_S. The reconstructions in Figures 1b, 1d, and 1f are based on four pairings of geomagnetic activity data: (black line) aa$_C$ and *IDV*(1d); (red line) aa$_C$ and *IDV*(1d); (orange line) *IHV* and *IDV*(1d); and (green line) *IHV* and *IDV*. The gray shows the extent of the 2σ uncertainty band. Blue dots joined by the blue line show annual means of available in situ data from near-Earth interplanetary spacecraft. In Figures 1b, 1d, and 1f, the agreement between the four reconstructions is so close that some lines cannot be seen as they lie beneath others. Figures 1c, 1e, and 1g show the probability distribution functions (pdfs) from the in situ spacecraft data on *B*, V_{SW}, and F_S, respectively, for 1966–2012 (inclusive) for hourly means (in gray) and annual means (in blue). Note that the pdf scale has been halved for annual data for clarity. [From *Lockwood et al.*, 2014d].

distinction between "open" solar flux as being field lines that extend beyond the edge of the schematic and "closed" field lines that form closed loops within the schematic. Such schematics often lead to the idea that open flux that only has one "foot point" still attached to the Sun instead of two: however, this is a false definition because, unless one invokes magnetic monopoles, all field lines that are still connected to the Sun have two such foot points, however distended the loop between them. Hence, a more precise definition of open flux is that the loop between the two foot points threads a certain surface in the heliosphere. This could either be a real but movable (and non-spherical) boundary (such as the heliopause, the heliospheric termination shock, or the Alfvén surface where the solar wind flow becomes super-Alfvénic) or it could be a nominal boundary, such as a fixed coronal source surface. A less precise definition is that the loop between the two foot points is so distended that one cannot tell where it maps to. In theory, unidirectional or bidirectional heat flux or "strahl" electron distribution functions could distinguish between open and closed flux by this latter definition; however, in practice electron scattering by heliospheric structure into other populations such as "halo" makes even this imprecise distinction ambiguous [*Larson et al.*, 1997; *Fitzenreiter et al.*, 1998; *Owens et al.*, 2008a]. However, these definitions are all academic when it comes to quantifying the total open solar flux, because there is no way of reliably mapping all the field lines (to see if they thread a certain heliospheric surface) or of remotely sensing the electron flows along them (to see if they are unidirectional or bidirectional) as a way of empirically quantifying the total open flux.

The only definition of OSF that gives a measurable quantity is that it is the magnetic flux threading the top of the solar corona. Usually, the "coronal source surface" is employed, taken to be a heliocentric sphere of radius $r = 2.5\ R_\odot$, where $R_\odot$ is a mean solar radius. OSF defined this way can be measured in two ways: (1) using "Potential Field Source Surface" (PFSS) modeling of the solar corona based on photospheric

magnetograms and several assumptions [see review by *Mackay and Yeates*, 2012]; or (2) from in situ measurements or estimates of the radial field in the heliosphere (for example near Earth) [see review by *Lockwood*, 2013]. However, in the second case, there are some important caveats and corrections caused by effects taking place between the source surface and the observation point which are discussed in section 1.4.

Because it renders it quantifiable, we here adopt the definition that the (signed) OSF is the total flux of one polarity that threads a fixed coronal source surface at $r = 2.5\ R_\odot$ and consider the implications of that definition. So-called "interchange reconnections" [*Crooker et al*., 2002] take place between open and closed field line loops. Using this definition of OSF, a closed loop does not thread the source surface. Hence, interchange reconnections make no difference to the OSF and simply move the photospheric foot point of the open field line. For this definition, all flux that threads the source surface is "open." The effects of any reconnection involving these open flux loops (however far they extend into the heliosphere or beyond and wherever their photospheric foot points lie) depends only on where the reconnection takes place relative to the Alfvén surface [e.g., *Schwadron et al*., 2010]. Reconnections taking place inside this surface will refigure the two open flux loops such that one of them can move under the influence of the MHD magnetic curvature force against the solar wind flow back toward the Sun through the coronal source surface, halving the OSF that was involved in the reconnection. Reconnections taking place beyond the Alfvén surface make no difference to the OSF because both the reconfigured loops continue to move away from the Sun.

Thus, for this definition of OSF, interchange interactions are not directly relevant to OSF continuity equation: for loops that do not thread the coronal source surface they make no direct difference to the OSF (although they can influence how the OSF subsequently evolves), and for loops that do thread the source surface all reconnections are the same (and cause loss of open flux provided the reconnection is inside the Alfvén surface). Note that the "Fisk circulations" of OSF that interchange reconnections below the Alfvén surface enable [*Fisk et al*., 1999; *Fisk and Schwadron*, 2001]) take place simultaneously with the cycles of OSF growth and loss (as the balance between OSF emergence and disconnection changes).

1.3. Modeling of the Open Solar Flux Variation

The concept used to model the OSF variation was first introduced by *Solanki et al*. [2000]. These authors were not concerned with the spatial distribution of open flux on the Sun but applied a global continuity equation to all open flux. They used sunspot number as a basis of their quantification of the emergence rate of open flux and an assumed linear loss rate with a time constant that was fitted to match the OSF reconstruction of *Lockwood et al*. [1999]. The open flux continuity equation has subsequently been used by many authors with refinements to the production and loss rate formulations used [*Lockwood*, 2003; *Owens and Crooker*, 2006, 2007; *Vieira and Solanki*, 2010; *Schwadron et al*., 2010; *Owens and Lockwood*, 2012; *Goelzer et al*., 2013].

The main difference in the modeling by *Lockwood* [2003] was only that instead of running the continuity equation forward in time from an assumed open flux of zero at the end of the Maunder minimum (as has been done by *Solanki et al*.), it was run backward in time starting from the observed open flux at the minimum between solar cycles 22 and 23: this yielded an estimate of the total open flux at the end of the Maunder minimum that was roughly a quarter of the average of space-age values. *Owens and Crooker* [2006] introduced the use of coronal mass ejection (CME) occurrence frequency and flux content into the computation of the emergence rate, which has subsequently been used by *Schwadron et al*. [2010], *Owens and Lockwood* [2012], and *Goelzer et al*. [2013] and is also employed here. The other differences between the models relate to the computation of the loss rate. *Vieira and Solanki* [2010] divided open flux into a short-lived and a long-lived component, each with its characteristic loss timescale. *Schwadron et al*. [2010] subdivided the open solar flux into a CME-associated part and the remainder and derived continuity equations for both parts including the effect of interchange reconnections on both. As discussed in section 1.2, for the measurable definition of OSF used here, this does not change the overall continuity equation so essentially these authors have refined both the production and loss terms. However, this has been achieved at the expense of an increase in the number of free variables. One criticism of the original *Solanki et al*. [2000] modeling that was often raised was that, in quantifying the emergence and loss rates, there are several parameters that are either free or poorly defined to the extent that fits to observations or reconstructions are not significant. Subdividing the production and loss terms, as done by *Vieira and Solanki* and *Schwadron et al*. gives more degrees of freedom, thereby making the fits achieved less statistically significant.

For this reason, the modeling by *Owens and Lockwood* [2012] was importantly different. These authors argued that new OSF injection is predominantly in the form of CMEs and so used CME fluxes and occurrence rates as a function of sunspot number, as introduced by *Owens and Crooker* [2006]. Using the OSF variation reconstructed from geomagnetic activity these authors showed that this production term requires a fractional OSF loss rate that depends only on the phase of the solar cycle and is approximately the same in each solar cycle (i.e., it is independent of the cycles' amplitude in sunspot numbers or in other solar activity indices). Furthermore, these authors showed that the solar cycle variation in this fractional OSF loss was very similar to the variation of the Heliospheric Current Sheet tilt—as had been predicted theoretically by *Owens et al*. [2011b] by considering how current sheet tilt facilitates OSF loss by disconnection. By making the assumption that the emergence rate in the long, low minimum between solar cycles 23 and 24 fell to its Maunder minimum value, *Owens and Lockwood* [2012] prescribed all inputs to the model and so had removed free fit parameters except one, which was the scaling factor between fractional OSF loss rate and the current sheet tilt.

1.4. The Relationship Between Open Solar Flux and the Near-Earth IMF

The continuity models described in section 1.2 predict the OSF and so to compare with observations and reconstructions of the near-Earth interplanetary magnetic field (IMF, B), the relationship between B and OSF must be known and understood. Part of this relationship is given by the theory of the Parker spiral in the heliospheric field [*Parker*, 1958]. However, there is a second factor that must be recognized and understood, namely the role of heliospheric field lines that do not follow the Parker spiral configuration, such as "folded" and "ortho-gardenhose" flux. This section deals with this second factor.

The Ulysses satellite showed that the *average* of the radial component of the heliospheric field (B_r) was independent of latitude. How the averaging is done is clearly important (for example if 27 day periods are used, "away" (from the Sun) and "toward" flux will largely cancel in the streamer belt giving a small residual value, whereas they would not cancel in a unipolar polar coronal hole. *Smith and Balogh* [1995] found latitudinal constancy by separately averaging over toward and away interplanetary sectors in the streamer belt, whereas *Lockwood et al*. [2004] found that it also applied to the average of the modulus of the radial field, $|B_r|$. *Smith* [2011] argues that the latter introduces an artefact and that only the sector averaging method should be used: however, as pointed out by *Lockwood and Owens* [2013], this does not answer the question of how sectors should be defined: for example, *Lockwood and Owens* showed that using the polarity of B_r observed at the spacecraft to define the sectors is mathematically identical to taking the modulus. Hence, the sector-average method depends entirely on the accurate identification of the sector boundaries (at the source surface, and not at the spacecraft), and there will be unknown errors associated with inaccurate definition of these boundaries.

The latitudinal constancy of $|B_r|$ found by *Lockwood et al*. [2004] means that averaging over a solar rotation allows us to quantify the signed flux threading a heliocentric sphere of radius r as $F(r) = 4\pi r^2 < |B_r| > /2$. Applying this to $r = 2.5\ R_\odot$ gives, by the definition we employ, the OSF, $F_S = F(r = 2.5\ R_\odot)$. *Owens et al*. [2008b] showed that $F(r)$ increases with r, and the rise in $F(r) - F_S$ with r is what *Lockwood et al*. [2009b, 2009c] define as the "excess flux." *Smith* [2011] argued that excess flux is an unphysical artefact caused by the use of the modulus: a misconception that has been repeated in subsequent papers, even though in reality they are actually developing methods to deal with the same physical effects [e.g., *Erdös and Balogh*, 2014].

There are three key points that must be understood. First, excess flux, as defined by *Lockwood et al*. [2009b, 2009c] (from the modulus of the radial field component), is not unphysical. It arises from any mechanism that causes some heliospheric field lines to bend back toward the Sun such that for one crossing of the coronal source surface a field line can thread a heliocentric sphere of radius r three (and potentially even five or more) times. Such deviations from Parker spiral orientation can be associated with ortho-gardenhose IMF orientations on large spatial scales or with Alfvén waves on smaller scales. Excess flux is also caused by "folded flux" associated with interchange reconnection in dipolar streamers and pseudostreamers, and this folded flux can remain in gardenhose orientation [*Owens et al*., 2013].

Second, the excess flux concept is in no way inconsistent with the use of the Ulysses result of latitudinal constancy of the radial component. This constancy is caused by the dominance of magnetic pressure in the very near-Sun heliosphere, which drives non-radial flow until the tangential pressure (and hence the

modulus of the radial field) is independent of latitude and particle pressure gradients at greater r are too small to perturb this coronal imprint [*Suess et al*., 1996]. The constancy of $|B_r|$ is established by $r \approx 5\ R_\odot$, i.e., very close to the coronal source surface and so the phenomena that cause excess flux act to perturb the radial constancy because they do not act equally at all latitudes. *Lockwood and Owens* [2009] have shown that if the excess flux is first removed and the latitudinal constancy then applied, the accuracy of the OSF value deduced using the Ulysses result is indeed increased.

The third key point is that although there are procedures that can be applied to modern satellite data to average out, reduce, or even eliminate excess flux, these are not possible for historical reconstructions from geomagnetic activity. One example of such a procedure is the sector averaging advocated by *Smith* [2011], which can be applied provided that electron heat flux measurements are available to define source sector boundaries with sufficient accuracy [*Lockwood and Owens*, 2013]. Similarly, the methods of *Erdös and Balogh* [2014] can be applied if simultaneous high time-resolution local plasma velocity measurements are available. However, these options are not possible when we are dealing with reconstructions of near-Earth interplanetary conditions: in these cases, we can infer B which, using a solar wind speed estimate and Parker spiral theory, we can relate to $|B_r|$—but as a consequence there is no option other than to quantify the excess flux and then subtract it. Hence, when comparing modeled OSF to reconstructions from geomagnetic activity, we must explicitly allow for the excess flux and use the formula (which is the definition of excess flux):

$$F_S = 2\pi {R_1}^2 < |B_{rE}| > -\Delta F = 2\pi {R_1}^2 < B\sin(\eta) > -\Delta F \quad (1)$$

where B_{rE} is the near-Earth radial field and η is the IMF gardenhose angle, and ΔF is the excess flux caused by folds in heliospheric fields on all spatial scales. Because Parker spiral theory shows that η depends on the solar wind speed, equation (1) shows that there is not a simple relationship between the near-Earth IMF B and the OSF, F_S: this relationship has been studied empirically by *Lockwood et al*. [2009a], *Lockwood and Owens* [2011], and *Lockwood et al*. [2014d]. The annual mean open solar fluxes shown in Figures 1f and 1g, both observed and reconstructed, have been corrected for excess flux using equation (1), where ΔF has been computed from the method of *Lockwood et al*. [2009c] for the observed data and using the statistical regression fit employed by *Lockwood et al*. [2014d] for the reconstructed OSF. (N.B. the hourly OSF values in Figure 1g are made using the approximate method of averaging B_r over one day centred on the hour in question [*Lockwood et al.,* 2009a]).

1.5. Streamer Belt Width

De Toma [2011] noted that in 2009 (during the long, low minimum between cycles 23 and 24), the near-Earth interplanetary satellites spent a larger fraction of the time in the slow solar wind than they had during previous solar minima in the space age. From eclipse observations of the solar corona, *Cliver and Ling* [2011] inferred that this was because the streamer belt was broader than during other recent solar minima and indeed showed that in images of the solar eclipse on 22 July 2009 the streamer belt appeared as wide as in images on 18 May 1901, a solar minimum in which geomagnetic reconstructions show the OSF fell to values very similar to those in 2009 [*Lockwood et al*., 1999, 2009c, 2014d].

Lockwood and Owens [2014] recently used the reconstructions of solar wind speed shown in Figure 1d to discuss the origins of the much lower average slow solar wind speeds derived for the solar minima around 1879 and 1901 (between cycles 11 and 12, and between cycles 13 and 14, respectively). They estimated the latitudinal width of the streamer belt and inferred that it was exceptionally wide during these minima and concluded that the low solar average solar wind speed occurred because the Earth remained consistently within the streamer belt during these minima. Figure 1d shows that the annual means of V_{SW} in these two minima are very close to the lowest values of the distribution of hourly means observed by near-Earth satellites shown by the grey histogram in Figure 1e. From this similarity, *Lockwood and Owens* [2014] suggest that the speed of the slow solar wind is unchanged but the average solar wind speed $< V_{SW} >$ at Earth is low because Earth resides almost continuously in the band of slow solar wind, rather than because of a change in the fundamental properties of the fast or slow solar wind. Evidence to support this idea comes from the recent solar minimum between cycles 23 and 24. That the streamer belt width was greater during minimum 23/24 has also been inferred from the Thompson-scattered light from the solar corona, as detected by coronagraphs, and from interplanetary scintillation observations of the solar wind [*Manoharan*, 2012]. *Lockwood and Owens* [2014] used the method of *Owens et al*. [2013, 2014] to define the streamer belt width by identifying dipole streamers and pseudostreamers using magnetograph data and PFSS modeling.

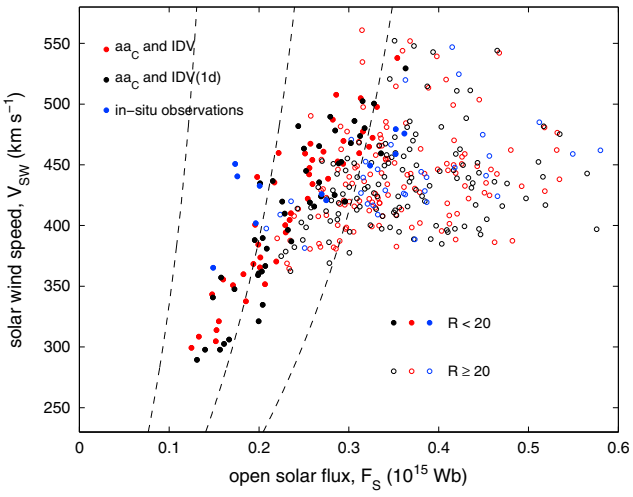

Figure 2. Scatterplot of annual means of near-Earth solar wind speed, V_{SW}, as a function of annual means of the signed open solar flux F_S for 1845–2012, inclusive. The black and red dots are reconstructed values using the geomagnetic index pairings of *IDV*(1d)-aa$_C$ and *IDV*-aa$_C$, respectively. The blue dots are the in situ observations for 1966–2012, inclusive. For all three data sets, years with corrected international number $R_C < 20$ are shown by filled circles, years for $R_C \geq 20$ by open circles. The dashed lines show the variation predicted by Parker spiral theory alone for three constant fields at the coronal source surface. The fact that the dashed lines have a steeper gradient than the best fit regression to the data shows that the variation cannot be attributed to the effect on F_S of the unwinding of the Parker spiral as solar wind speed, V_{SW} increases.

Both types of streamer act as sources of the slow solar wind [*Antiochos et al*., 2011; *Riley and Luhmann*, 2012; *Owens et al*., 2013, 2014], and *Lockwood and Owens* [2014] showed that this belt decreased in latitudinal width over all three declining phases of the observed solar cycles but did not become as thin during minimum 23/24 as it had at the previous two minima.

Cliver and Ling [2011] and *Lockwood and Owens* [2014] extrapolated their findings to the Maunder minimum and argued that the low open solar flux at this time would have caused a persistently broad streamer belt and that the Earth would have remained consistently within a band of slow solar wind as a result. In this paper, we extend the continuity equation modeling of the open solar flux to predict the streamer belt width. We make use of the correction to sunspot numbers quantified in Paper 1 [*Lockwood et al*., 2014a]. Paper 2 [*Lockwood et al*., 2014b] discussed the divergence of the (corrected) international sunspot number and the group sunspot number in the years before the start of the photographic observations by the Royal Greenwich Observatory in 1874. Because there is no way to know which of these two sunspot records best quantifies OSF emergence in these years, in the present paper, we use both the group sunspot number R_G and the (corrected and extended) international/Zürich/Wolf sunspot number composite R_C and discuss the implications for the modeling of the differences between the two.

2. Relationship Between Reconstructed Solar Wind Velocity and Open Solar Flux

In a series of four papers, *Lockwood et al*. [2013a, 2013b, 2014a, 2014b] have recently extended the aa index and the OSF, near-Earth IMF, and solar wind speed reconstructions back to 1845, as shown by Figure 1. The longest reconstructions are based on two pairings, the corrected aa index, (aa$_C$) with the *IDV* index and aa$_C$ with the *IDV*(1d) index. The results are very similar after 1872 and the tests presented in Papers 1 and 2 show that where they do differ before this date, it is the *IDV* index that is behaving differently. Thus, for 1845–1872 the reconstructions for aa$_C$ and *IDV*(1d) are the most reliable.

Figure 2 shows a scatterplot of both reconstructed and observed annual values of the near-Earth solar wind speed, V_{SW}, as a function of annual means of the signed OSF F_S for, respectively, 1845–2012 and 1966–2012, inclusive. For all data sets, solar minimum years (defined as with corrected sunspot number $R_C < 20$) are shown by filled circles, while all other years (with $R_C \geq 20$) by open circles. For much of the data ($F_S > 0.25 \times 10^{15}$ Wb) there is no coherent relationship. However, considering only $F_S < 0.25 \times 10^{15}$ Wb a clear relationship is seen with the annual means in V_{SW} roughly proportional to F_S. The lowest F_S are generally at sunspot minimum and the relationship is most clear if we restrict the data to solar minimum years ($R_C < 20$), as shown by the solid points.

The F_S values are computed from reconstructed annual means of B and V_{SW} and allow for both the excess flux and the effect of V_{SW} on the Parker spiral. The dashed lines show the variation expected for Parker spiral theory for three different values of the source field. (As V_{SW} increases, the spiral for a given coronal source field unwinds, such that the ratio $|B_r|/B$ increases and hence the F_S values increase for a given B). This

effect has been allowed for in the F_S values, and Figure 2 shows that the relationship between V_{SW} and F_S cannot be explained by an overestimation of this effect. Hence, average solar wind speeds at Earth during solar minimum are genuinely lower when the OSF is low. In the remainder of this paper we model the streamer belt width to analyze the concept of *Lockwood and Owens* [2014] that this is associated with increased width of the streamer belt at sunspot minimum, and to extend this understanding to the Maunder minimum.

3. Model of Open Solar Flux and Streamer Belt Width

3.1. The Basic Equations

Following *Solanki et al*. [2000], we model the signed OSF, F_S (here defined as the flux of one polarity threading the coronal source surface) using the global solar continuity equation

$$dF_S/dt = S - f \times F_S \tag{2}$$

where S is the source rate and f is the fractional OSF loss rate (so that the total loss rate is $f \times F_S$) which, following *Owens and Lockwood* [2012], we generalize to be a function of solar cycle phase, Φ, only. This quantifies the degree of heliospheric current sheet (HCS) warping and its effect on the OSF disconnection rate.

Owens and Crooker [2007] generated a more complex formulation of the OSF budget using OSF subdivisions based on different suprathermal electron populations, and this was employed by *Schwadron et al*. [2010] who added an explicit time constant for converting "closed" heliospheric loops to "open" coronal hole flux (rather than using the assumption of *Owens and Crooker* that the counterstreaming signature is lost within about 1 month of emergence) and allowed for separate disconnection and interchange reconnections. This provides a very elegant formulation of the OSF budget, but when it comes to fitting the observed and reconstructed OSF suffers from the larger number of free parameters, as discussed in section 1.3, and which are not essential for the working definition of OSF, as discussed in section 1.2. Instead of building up to a total OSF solution using the subdivision and equations of *Schwadron et al*. [2010], we here adopt the simpler formulations of production and loss rates of *Owens and Lockwood* [2012] with its single free fit parameter to compute the total OSF. However, we then investigate the suggestion of *Lockwood and Owens* [2014] by applying some of the principles discussed by *Schwadron et al*. [2010] to subdivide the OSF into two categories: streamer belt (SB) flux and coronal hole (CH) flux, the latter including both polar coronal holes and coronal holes at lower latitudes. Hence,

$$F_S = F_{SB} + F_{CH} \tag{3}$$

We assume that all newly emerged OSF is injected into the SB (i.e., the total flux emergence rate through the coronal source surface S is source term for SB flux). Some of this emerged flux is lost by disconnection at reconnection X-lines in the HCS which is embedded in the SB. The remainder that stays connected evolves into CH flux as the flux tube expands into the heliosphere and its foot points migrate poleward across the Sun under the influence of meridional flow and interchange reconnections. Given that flux tubes reach the termination shock within about a year of emergence, after which no information about the flux tube evolution can be received in the inner heliosphere, this evolution from SB to CH on timescales greater than a year is associated with flux transport in the photosphere and not evolution of the OSF flux loop in the heliosphere. Interchange reconnections move the foot points of the CH flux, but its loss requires reconnection between adjacent CH field lines of opposite polarity. From these considerations, the continuity equation for SB flux is:

$$dF_{SB}/dt = S - f_{SB} \times F_S - S_{CH} \tag{4}$$

where $f_{SB} \times F_S$ is the rate of disconnection of SB flux at the HCS and S_{CH} is the rate at which SB flux is transferred to coronal hole flux. The corresponding continuity equation for CH flux is

$$dF_{CH}/dt = S_{CH} - f_{CH} \times F_S \tag{5}$$

Note that f_{SB} and f_{CH} are the "fractional" loss rates, defined as the disconnection rate (for SB and CH flux, respectively) per unit total OSF. Adding equations (4) and (5), using equation (3) and comparing with equation (2) gives

$$f = f_{SB} + f_{CH} \tag{6}$$

Hence, the total fractional loss rate can be divided into part taking place in the streamer belt and part taking place in coronal holes [*Schrijver et al*., 2002]. *Owens and Lockwood* [2012] used the reconstructed open solar flux variation with equation (2) to show that the fractional loss rate depended only on solar cycle phase and that the cyclic part was highly correlated with HCS tilt, as expected theoretically for the disconnection of SB flux [*Owens et al*., 2011b].

In order to solve these equations we have to know the variation in the rate at which SB flux evolves into CH flux, giving the loss term S_{CH} in equation (4). Because all newly emerged OSF is envisaged as being injected into the SB, S_{CH} is the only source term in equation (5). The rate S_{CH} can be computed for a given temporal variation in S with a few simple assumptions, as discussed in section 3.3 below.

Once the total open solar flux F_S and streamer belt flux F_{SB} have been computed, the latitudinal half width of the streamer belt Λ_{SB} can be computed because of the latitudinal constancy of the radial component of the heliospheric field giving the equation used by *Lockwood and Owens* [2014]:

$$\Lambda_{SB} = \sin^{-1}\{1 - F_{CH}/F_S\} \tag{7}$$

3.2. Input Emergence Rate Variation

The only source of information on the long-term variation in OSF emergence available to us is the sunspot number record, and *Solanki et al*. [2000] demonstrated that it could be used in the OSF continuity equation. We here use both the international/Zürich/Wolf sunspot number composite and group sunspot number sequences, the former having been corrected to allow for calibration problems around the "Waldmeier" and "Wolf" discontinuities, as discussed in Papers 1 and 2. As noted in Paper 2, the corrected international sunspot number, R_C, and the group sunspot number, R_G, diverge as we go back in time—a divergence that is present after 1874 (the start of the RGO data) but is most pronounced before then but which fluctuates with time. There is no way to say if R_C or R_G better describes the OSF emergence rate before the mid-19th century, and so we here consider the implications of both. The R_G sequence extends back to 1610, before the start of the Maunder minimum, and we here extend R_C back to the same starting date using $1.3R_G$, as proposed in Paper 2.

From estimates of the CME occurrence rate over recent solar cycles and the magnetic flux emerging into the SB, *Owens and Lockwood* [2012] derived the flux emergence rate per year. We employ a slight variant of that equation (based on the best fit between CME rates and sunspot numbers [*Owens et al*., 2008c]) to predict the total OSF emergence rate:

$$S = a \times [F_{CME} \times (44.64 \times \tanh(0.00675R) + 8.196)] - b \tag{8}$$

where F_{CME} is the estimated average signed flux content injected by a CME, and R is the sunspot number estimate (we employ R_C and R_G). We here use $F_{CME} = 3.5 \times 10^{11}$ Wb [*Owens et al*., 2008c]. The terms a and b give a correction to allow for the fact that we do not know how much smaller the emergence rate in the Maunder minimum S_{MM} was, compared to the baselevel detected in the recent low solar minimum, S_o. The term in square brackets is the rate used by *Owens and Lockwood* [2012] (with slightly changed coefficients obtained by fitting to a slightly updated data set) which assumes that the recent minimum set a baselevel emergence rate that also existed during the Maunder minimum (i.e., they assumed $S_{MM} = S_o$). Hence, the use of $b = 0$ and $a = 1$ recovers this assumption. The largest possible value for $b = S_o - S_{MM}$ is 1.43×10^{14} Wb yr^{-1} (because any larger b would result in negative S_{MM}). To give an uncertainty analysis, we varied b over the full range up to this maximum value. The best fit of the modeled F_S variation to that derived from igeomagnetic observations was found to be for $b = 0$ (and $a = 0.95$). The uncertainty in this value and its effects will be discussed later. The best fit variation in the OSF source rate, S, from equation (6) using these best fit a and b values with R_C and R_G is shown by the black and mauve lines, respectively, in Figure 3.

3.3. The Transfer of Streamer Belt Flux to Coronal Holes

In order to compute the rate of flux transfer S_{CH} from the streamer belt to the coronal holes, we assume that the loss of SB flux by disconnection occurs during the first year after emergence (after that foot points

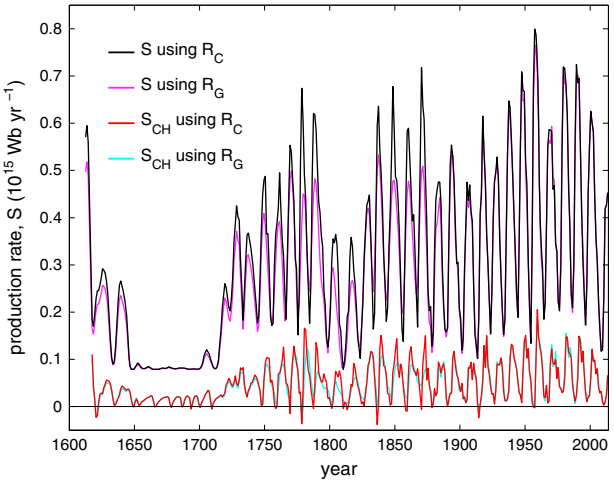

Figure 3. The upper two lines give the emergence rate of open solar flux *S*, based on the analysis of CME occurrence frequency and flux content as a function of sunspot number. The sunspot variations used are R_C and R_G for the black and mauve lines, respectively. The open flux production rate is given by equation (8) of the present paper for the best fit coefficients $b = 0$ and $a = 0.95$. The lower two lines giver the rate of flux transfer from the streamer belt to coronal holes, S_{CH}, for an example value of the fractional coronal hole loss rate, $f_{CG} = 0.22$: the red line is for R_C and the cyan line for R_G.

are deemed too have migrated poleward too far from HCS for such disconnection to happen). We also allow for the SB flux evolving into CH flux on a distribution of timescales. We here assume that of the emerged flux SB that survives the first-year disconnection, 10%, 50%, 30%, and 10% are transferred to CH flux after $y = 1$, 2, 3, and 4 years, respectively, following emergence. Hence, after 4 years all emerged SB flux has either been disconnected or evolved into CH, and the average time for this evolution is 2.4 years. This evolution profile was derived iteratively, but the results were largely insensitive to the precise profile as long as the mean value remained close to 2.4 years (the main effect of the distribution to be to add some smoothing to the modeled F_{CH} and F_{SB} sequences). Thus, S_{CH} can be computed for a given variation of the emergence rate *S*, provided that the loss of that emerged flux in the first year subsequent to emergence ($f_{SB} \times F_S$) is known:

$$S_{CH}(t) = \Sigma_{y=1}^{5} w(t-y)\{S(t-y) - f_{SB}(t-y) \times F_S(t-y)\} \quad (9)$$

From the above, we employ the temporal transfer profile of $w(t-1) = 0.1$, $w(t-2) = 0.5$, $w(t-3) = 0.3$, $w(t-4) = 0.1$, and $w(t-5) = 0$, giving the mean time before transfer of 2.4 years. The average was adjusted to 2.4 years to make the predicted F_{CH} in phase with the solar polar fields measured by the Wilcox Solar Observatory (WSO) magnetograph (see discussion below about Figure 8). We note that *Schwadron et al.* [2010] employed 2 years for a broadly equivalent timescale.

3.4. OSF Loss Rate Variation

Using the reconstructed OSF variation and a source rate variation similar to that shown in Figure 3, *Owens and Lockwood* [2012] noted that the fractional loss rate *f* required to model the data using equation (1) varied with solar cycle phase Φ in a regular manner for each cycle. Furthermore, the variation of *f* with Φ required was very similar to the variation in the current sheet tilt with Φ, consistent with the proposal by *Owens et al.* [2011b] that the loss of open flux varied with disconnection of flux of opposite polarity across the heliospheric current sheet, brought into contact by solar rotation such that the fractional loss rate is enhanced when the current sheet tilt is high.

The upper two panels of Figure 4 describe how the solar cycle phase is determined for after 1874, using the method of *Owens et al.* [2011a]. The upper panel shows the "butterfly diagram" from the RGO/SOON sunspot group data, by plotting the number of sunspot groups as a function of heliographic latitude, Λ_H, and time. This plot employs the SOON/RGO correction factor described in Paper 1. The second panel shows the mean of the absolute latitude of the groups $< |\Lambda_H| >$ as a function of time, revealing the regular variation over each cycle, independent of the amplitude of the cycle (in terms of spot area or sunspot number). The start of each new cycle is marked by the grey vertical line when the mean sunspot latitude increases most rapidly. Defining these as the start and end of each cycle ($\Phi = 0$ and $\Phi = 2\pi$), the $< |\Lambda_H| >$ value can be used to define Φ at any time from the average variation of $< |\Lambda_H| >$ with Φ for all cycles.

The bottom panel of Figure 5 shows the fractional loss rate *f* required to match the OSF variation for the source rate variation *S* prescribed by Figure 4. The blue and red dots are derived using equation (2) of *Owens*

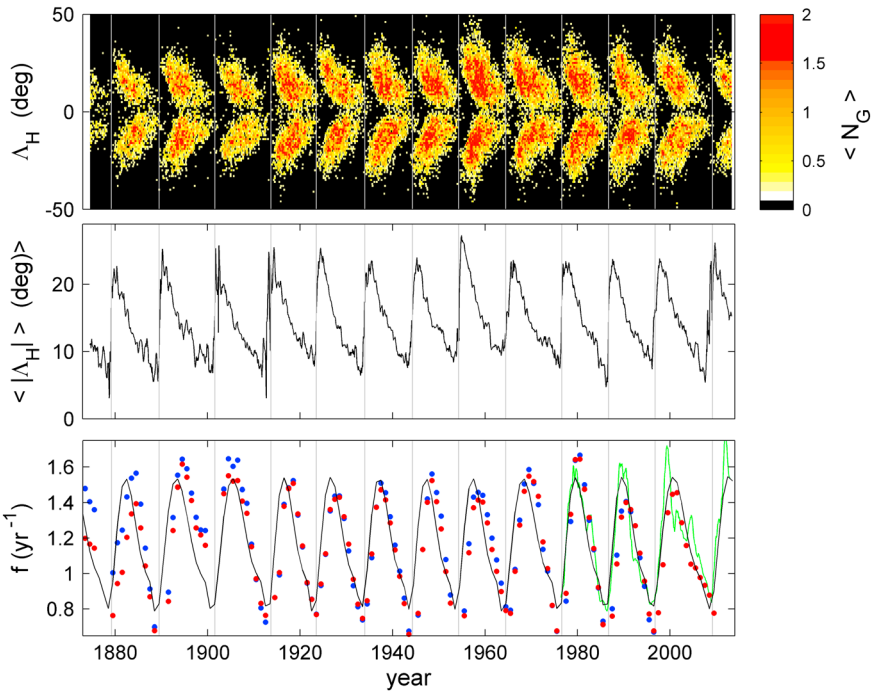

Figure 4. (Top) The "butterfly diagram" showing the number of sunspot groups per pixel, N_G, as a function of date and heliographic latitude, Λ_H (where a pixel is 27 days in time and 1° wide in Λ_H). (Middle) the mean of the modulus of the heliographic latitude of sunspot groups, $<|\Lambda_H|>$. Vertical grey lines are the start of each cycle ($\Phi = 0$), determined from $<|\Lambda_H|>$ and used to evaluate the phase of the solar cycle phase Φ as a function of time by the method of *Owens et al.* [2011a]. (Bottom) The fractional loss rate f of open solar flux required by the continuity model to yield the reconstructed OSF data using the source rate variations, S, given in Figure 4. The blue and red dots are the values of f from the geomagnetic OSF reconstruction data for the prescribed S using, respectively, the R_C and R_G sunspot records, and the black line is the fit as a function of Φ only made from the superposition of all cycles. The green line is the best linear-regression fit of 1 year means of the heliospheric current sheet tilt index of *Owens et al.* [2011b].

and Lockwood [2012] from the geomagnetic OSF reconstruction data: blue dots are the results if R_C is used to compute S and red dots if R_G is used. The OSF data have been passed through a three-point running mean to smooth out fluctuations. The black line is the fit to the average variation with solar cycle phase Φ, plotted as a function of time using the variation of Φ derived from $<|\Lambda_H|>$.

Also shown in the bottom panel of Figure 4 (in green) is the best fit linear regression of the Heliospheric Current Sheet index [*Owens et al.*, 2011b] derived from application of the PFSS method to the WSO magnetograph data. It can be seen that the variation of this index with Φ is very similar to that of the average fractional loss rate. This gives strong support to the idea that the cycling of the fractional loss rate, shown by the blue and red dots in Figure 4c, is caused by the disconnection of SB flux by reconnection in the HCS, as proposed by *Owens et al.* [2011b]. Thus, this cyclic variation is best explained as a feature of the loss of SB flux. However, there is a baselevel fractional loss rate f of about 0.8 year^{-1}, some of which could be baselevel SB disconnection, and some of which is disconnection of CH flux [*Schrijver et al.*, 2002].

Given that the cyclic behavior of the total fractional loss rate f is well explained by its streamer belt component f_{SB} (see equation (6)) there is no evidence that the coronal hole component f_{CH} is not constant. However, the value of f_{CH} is not known. We must have $f_{CH} > 0$ or, by equation (4) coronal hole flux will continuously grow and the streamer belt disappear. From Figure 5 we know $f_{CH} < 0.8$ year^{-1}, the minimum value of f seen every solar cycle. To determine the optimum f_{CH}, we vary it between these two limits and use the equations of section 3.1 to generate the time series for 1610–2012 of the OSF F_S, the SB flux F_{SB}, the CH flux F_{CH}, the latitudinal half width of the streamer belt Λ_{SB}, and the transfer rate between the streamer belt and coronal holes, S_{CH}. Note that the solar cycle phase Φ was determined for solar cycles 8–11 using the extension of the butterfly diagram by *Arlt et al.* [2013]. Before 1825 there is no continuous record of

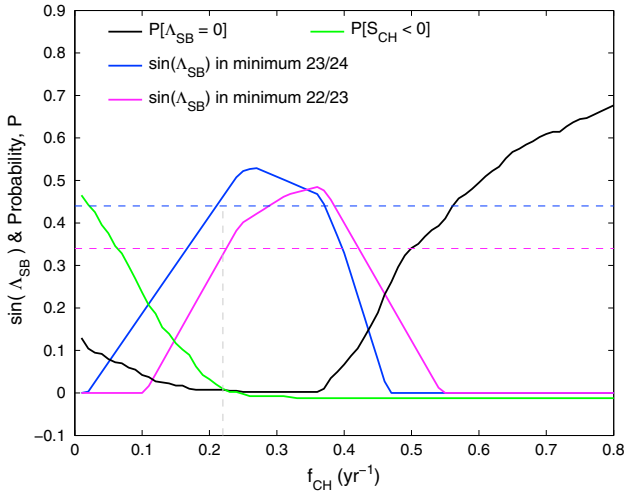

Figure 5. Determination of the fractional coronal hole loss rate f_{CH}. Parameters are computed for the full range of possible f_{CH} values between 0 and 0.8 year^{-1}. The black line shows the fraction of annual values of the modeled streamer belt width Λ_{SB} for 1620–2012 that are zero (i.e., the streamer belt has disappeared), the green line the fraction of the streamer belt to coronal hole transfer rate values, S_{CH}, that are negative. The modeled minimum values of sin(Λ_{SB}) for solar minima 22/23 and 23/24 are shown by the mauve and blue lines. The values determined from the PFSS mapping of dipolar streamers and pseudostreamers (set by the limit to the occurrence of dipolar and pseudostreamers) for these minima are shown by the horizontal dashed lines. The vertical gray dashed line is at $f_{CH} = 0.22$.

sunspot latitudes from which to determine the variation of Φ, and so the times of the minimum in R_G are used to define the times of $\Phi = 0$ and a linear variation of Φ with time assumed in each cycle. During the Maunder minimum the maxima in the cycles of 10Be isotope abundance are used [*Beer et al.*, 1998], as in the study by *Owens and Lockwood* [2012].

The parameter f_{CH} is the major free parameter in modeling the streamer belt width presented in this paper, and Figure 5 explains how the value of this parameter is set. Figure 5 shows the results for f_{CH} values in the allowed range between zero and 0.8 year^{-1}. Using both R_C and R_G, the initial results show some influence of the initial (1610) values of F_S and F_{SB} (and hence F_{CH}) from which the solution of equations (1), (3), and (4) are initiated. However, by 1620 all influence of these spin-up values was found to be lost. The black line shows the fraction of the 391 annual values after 1620 for which F_{SB} (and hence Λ_{SB}) fall to zero, $P[\Lambda_{SB} = 0]$, i.e., when the streamer belt disappears. As expected, if f_{CH} is too small, this happens quite frequently. It is also found to occur increasingly often if f_{CH} is too large because this causes f_{SB} to be too small and too much flux to be transferred to the coronal holes, i.e., S_{CH} is too large. Only for the limited range $0.18 < f_{CH} < 0.37$ year^{-1} does the streamer belt not regularly vanish. The best determination of f_{CH} comes from comparing the latitudinal half width of the streamer belt Λ_{SB} in recent minima with that obtained from the occurrence frequency of dipole streamers (DS) and pseudostreamers (PS) determined from PFSS maps of the solar corona from magnetograph data using the method of *Owens et al.* [2013, 2014]. This is discussed further later in the present paper. The dashed horizontal lines in Figure 5 are the minima of the streamer belt half-width Λ_{SB} for the solar minima between cycles 22 and 23 (mauve dashed line) and between cycles 23 and 24 (blue dashed line), as determined from where DS and PS streamer frequency falls to zero. The solid lines of the same colors show the model predictions for Λ_{SB} at these minima as a function of f_{CH}. In both cases, the values closest to the observed values occur for $f_{CH} = 0.22$ year^{-1}. Thus, we adopt $f_{CH} = 0.22$ year^{-1} as the value giving the best match to the SB width during recent solar minima.

The green line in Figure 5 shows another consequence of the f_{CH} value adopted by showing the fraction of the 391 annual values of S_{CH} that are negative, $P[S_{CH} < 0]$ which we would not expect to occur. The number of years in which this happens declines with f_{CH} and is zero for $f_{CH} > 0.32$. In fact in this regime, S_{CH} is generally too large, leading to non-zero $P[\Lambda_{SB} < 0]$ which increases with f_{CH}. The red and cyan lines in Figure 3 show the variation of S_{CH} for the optimum f_{CH} of 0.22 year^{-1}: it can be seen that negative values are seen before 1620 which are associated with initial starting conditions and the model spin-up. Subsequently, negative S_{CH} occurs just 3 times and then by small amounts which we can attribute to the numerical uncertainties in the model. Hence, although negative S_{CH} is a problem if f_{CH} is too small, it is at an acceptable level at $f_{CH} = 0.22$ year^{-1}.

4. Model Results

Setting the fractional coronal hole loss rate f_{CH} allows equations (2)–(7) and (9) to be solved for a given emergence rate variation $S(t)$. In particular, the variations in the total OSF, F_S, the coronal hole OSF, F_{CH}, the

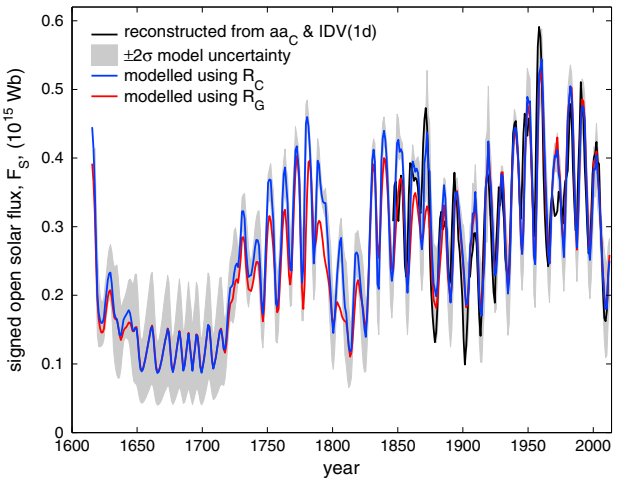

Figure 6. Modeled variations of annual means of signed open solar flux, F_S. The blue line is based on R_C, and the red line is based on R_G. (Note that for many years the red line is not visible as it is overlaid by the blue line). The black line is the reconstruction of F_S made using the aa_C and *IDV*(1d) geomagnetic indices (see Figure 1). All three series have been smoothed with a three-point running mean to improve plot clarity. The correlation between the modeled and reconstructed data is 0.876. The gray area is bounded by the maximum and minimum values for each year of all fits for which the RMS deviation is different from the optimum at the 2σ level.

streamer belt OSF, F_{SB}, and the latitudinal half width of the streamer belt, Λ_{SB}, can all be calculated. We use $S(t)$ given by equation (8) using both the corrected and extended international sunspot number, R_C and the group number, R_G. By using a range of these starting values we find that all influence on any of the modeled parameters is lost within one solar cycle, i.e., by 1620, and hence we regard the first 10 years as model spin-up time only.

4.1. Total Open Solar Flux

Figure 6 shows the variation of modeled total OSF, F_S, for R_C in blue and R_G in red. Note that these simulations do not depend on the division of OSF between the SB and CH flux, being determined by equation (2) which uses only the total fractional loss rate, f, which is set by solar cycle phase only, as demonstrated by Figure 4. Hence, these results are not dependent on the derived value for f_{CH} discussed above. The black line in Figure 6 is the reconstructed OSF variation from geomagnetic data (that derived using *IDV*(1d) and aa_C is shown, but as demonstrated by Figure 1, other combinations of indices give almost identical results). Remember that the input variation of the emergence rate S has been fine-tuned via the parameter b to give the best fit (lowest RMS deviation) which yields a correlation coefficient of 0.873. This fit was for $b = 0$. The variation of the RMS fit residual for full range of potential b values was evaluated. The probability density function (pdf) of the differences between the minimum fit residual and those at general b were evaluated and those within the 2σ significance level gave b in the range -0.05×10^{14} Wb yr^{-1} ≤ b ≤ +0.11 Wb yr^{-1}. This sets the 2σ uncertainty ranges of modeled F_S values in any year shown by the gray band. The blue and red lines show the results using R_C and R_G, respectively, and the uncertainty band for R_G (not shown for clarity) is very similar in width to that for R_C (shown in gray) in all years. The black line lies within this gray band for all but 12 of the 167 data points, the most consistent failure to fit being at the peak of solar cycle 20 (around 1970) when modeled values are consistently higher using both R_C and R_G.

The derived OSF during the Maunder minimum is interesting. Averaging over a whole number of the cycles in the MM (1655–1696) gives a mean OSF of $(0.120 \pm 0.061) \times 10^{-15}$ Wb. If we optimize b by maximizing the correlation coefficient between modeled and reconstructed OSF (rather than minimizing the RMS residual) we get a somewhat larger optimum b of 0.15 Wb yr^{-1} which reduces the MM OSF prediction somewhat to $(0.057 \pm 0.041) \times 10^{-15}$ Wb: a lower value but the uncertainty ranges of the two estimates do overlap. The value obtained by *Lockwood and Owens* [2011] from the estimate of the near-Earth IMF from cosmogenic isotopes by *Steinhilber et al.* [2010] was $(0.048 \pm 0.029) \times 10^{15}$ Wb. The modeling of OSF by *Lockwood* [2003] (running backward in time from an initial value taken from the first Ulysses perihelion pass) gave an estimate of 0.1×10^{15} Wb. *Cliver and Ling* [2011] argue that there is a floor value of the OSF of 0.080×10^{15} Wb. The minimum values of the cycles modeled here for during the Maunder minimum are 0.075×10^{15} Wb with the 2σ uncertainty limits at 0.04×10^{15} Wb and 0.13×10^{15} Wb. Thus, the modeling here suggests that the OSF during the Maunder minimum fell close to *Cliver and Ling's* proposed floor value. We would argue that this is still not a floor in any meaningful sense. From equation (1) the minimum OSF is simply the ratio of the minimum emergence rate S divided by the maximum fraction loss rate f. We here find no physical reason to set a firm minimum to S, so there is no

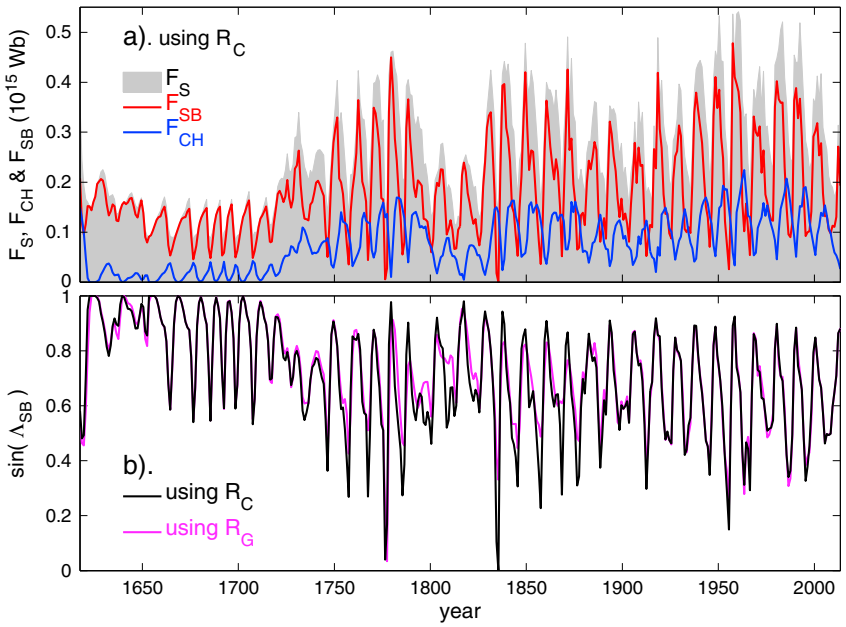

Figure 7. (a) Modeled variations of annual means of: (gray-shaded area) the total signed open solar flux, F_S; (red line) the open solar flux in the streamer belt, F_{SB}; and (blue line) the open solar flux in coronal holes, F_{CH}. These predictions all use the R_C input sunspot variation. (b.) The sine of the modeled annual mean latitudinal half-width of the streamer belt, sin (Λ_{SB}): the black and mauve lines use the sunspot numbers series R_C and R_G, respectively, as input into the model.

physical limit (a "floor"): it is simply that the probability of a given OSF is decreasingly small as OSF decreases in this tail of this distribution.

It should also be noted that the agreement between the lowest OSF estimates presented here and the proposed floor value in *Cliver and Ling* [2011] is possibly somewhat coincidental. *Cliver and Ling* argue that the near-Earth IMF comprises a 2.8 nT floor caused by magnetic field embedded in the slow solar wind (the SB flux) and a variable part that varies between 0 and 3 nT from cycle to cycle associated with the solar polar fields (the CH flux). Figure 7 shows a radically different prediction from the modeling presented here. The SB flux is far from constant; indeed, its variations are the dominant cause of the OSF variation as at most times the OSF is dominated by SB flux (the CH flux is only larger in the late declining phase and minimum of each cycle). As a result, the OSF peaks around sunspot maximum (and not near sunspot minimum when CH flux peaks). It is important to understand that B and $|B_r|$ at Earth depend on both the CH and SB flux, irrespective of whether Earth is within the streamer belt of not. This is because the latitudinal constancy of $|B_r|$ is established very close to the Sun (within $r \approx 5\ R_\odot$ [*Suess et al.*, 1996]) so that, as pointed out by *Lockwood and Owens* [2014], although the width of the streamer belt depends on the ratio of the SB and CH fluxes, $|B_r|$ (and thence B) at any heliographic latitude (inside the streamer belt or not) depends on the sum of the SB and CH fluxes.

The effect of using the group sunspot number R_G instead of R_C can be seen in Figure 6. Because they are so similar for most of the interval of the geomagnetic OSF reconstruction, the results over this interval are very similar indeed. As expected, predicted OSF values are lower on either side of the Dalton minimum if R_G is used rather than R_C. Within the Dalton minimum, cycle amplitudes are larger if R_C is used, but the cycle minima are very similar for R_C and R_G. Also, as expected from Paper 2, the fits to the reconstructed OSF are better in cycle 9 using R_G but better in cycle 11 using R_C.

4.2. Streamer Belt and Coronal Hole Flux

The top panel of Figure 7 shows the variations of the coronal hole (in blue) and streamer belt (in red) fluxes derived using R_C and the optimum f_{CH} of 0.22 year^{-1}. The area shaded gray gives the sum of the two, which is the total OSF, F_S. It can be seen that the peak F_S values (around sunspot maximum) are dominated by F_{SB}, and F_{CH} at these times is a minimum for each cycle. At the times of the minima in F_S, F_{SB} is also at a

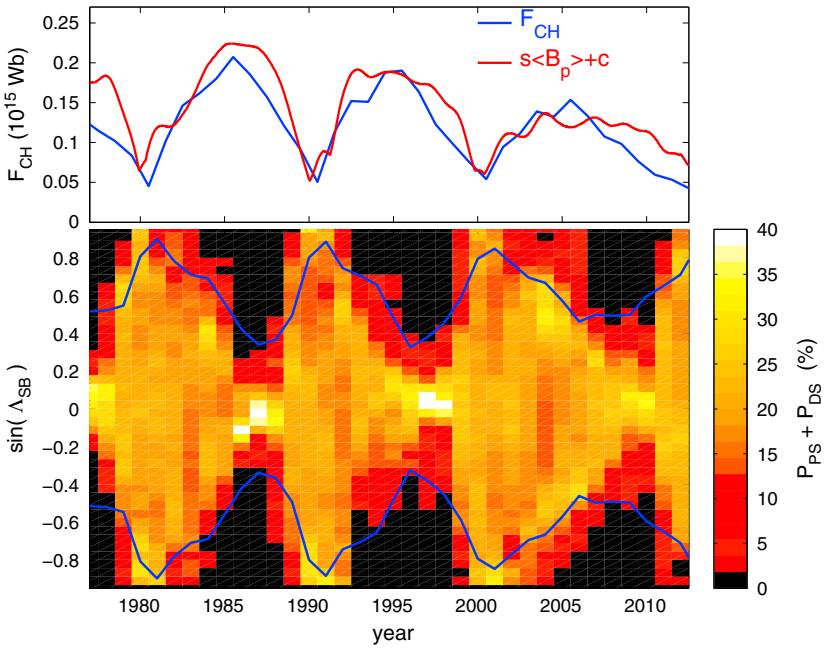

Figure 8. (Upper panel) The modeled coronal hole flux, F_{CH}, is shown in blue. The red line shows the scaled average $< B_p >$ of the polar field strength for the northern and southern hemispheres from the WSO magnetograms, B_N and B_S, respectively. The scaling coefficients $s = (1/750)$ and $c = 0.05$ (for $< B_p > = (B_N + B_S)/2$ in μT) are just to allow display on the same scale. (Lower panel) Observed and predicted streamer belt for the last three-and-a-half solar cycles. The blue line is the modeled variation shown in Figure 8. The color contours show annual means of the occurrence frequency (in %) of streamers (both dipolar and pseudostreamers from Potential Field Source Surface modeling of the corona based on magnetograms from the Wilcox Solar Observatory) as a function of heliographic latitude Λ_H and time. Streamers are defined by $\log_{10}(dS_{PH}/dS_{SS}) > 1$ where dS_{PH} is the separation of a pair of field lines in the photosphere and dS_{SS} is their separation at the coronal source surface [*Owens et al.*, 2013, 2014].

minimum and F_{CH} at a maximum giving an almost antiphase behavior of F_{SB} and F_{CH}. This behavior is reproduced for any profile of the rate of evolution of SB flux into CH flux (the $w(t\text{-}y)$ coefficients in equation (9)) that ensures the peak transfer occurs between 2 and 4 years after emergence. With this phase lag over each solar cycle, the long-term variation in F_{CH} follows that in F_S quite closely. The plot corresponding to Figure 7a generated using R_G is similar in every way, other than all fluxes are lower on either side of the Dalton minimum, as seen for F_S in Figure 6.

4.3. Streamer Belt Width

Using equation (7), the variation of the width of the streamer belt can be computed from F_S and F_{CH}. The lower panel of figure 7 shows the modeled variation of the sine of the latitudinal half-width of the streamer belt, $\sin(\Lambda_{SB})$. The black line uses R_C to quantify the OSF emergence rate whereas the mauve line uses R_G. It can be seen that the model predicts times when the coronal hole flux goes to zero ($\sin(\Lambda_{SB}) = 1$), but there are no years when the streamer belt vanished ($\sin(\Lambda_{SB}) = 0$). Very thin streamer belts ($\sin(\Lambda_{SB}) < 0.3$) in Figure 8 always precede a year of with low transfer rate (Figure 5) and arise in sunspot minimum years for which the 11 year running mean of sunspot number is high ($<R_C>_{22} > 55$). Results for R_C and R_G are very similar at each solar maximum when $\sin(\Lambda_{SB})$ peaks. However, the minima in $\sin(\Lambda_{SB})$ are less deep using R_G (compared to those for R_C) on either side of the Dalton minimum when R_G also yields lower F_S values.

The best test of the streamer belt width modeling is to compare with the streamer belt width deduced over recent solar cycles using Potential Field Source Surface (PFSS) modeling of coronal fields from magnetograph observations of photospheric magnetic fields. From the photospheric separation of adjacent field lines at the coronal source surface (at an heliocentric distance of 2.5 $R_\odot$, where $R_\odot$ is a mean solar radius), *Owens et al.* [2013] define the occurrence of dipolar streamers (DS) and pseudostreamers (PS). The colored pixels in Figure 11 show the sum of the occurrence frequencies of DS and PS ($P_{PS} + P_{DS}$) as a function of the sine of the

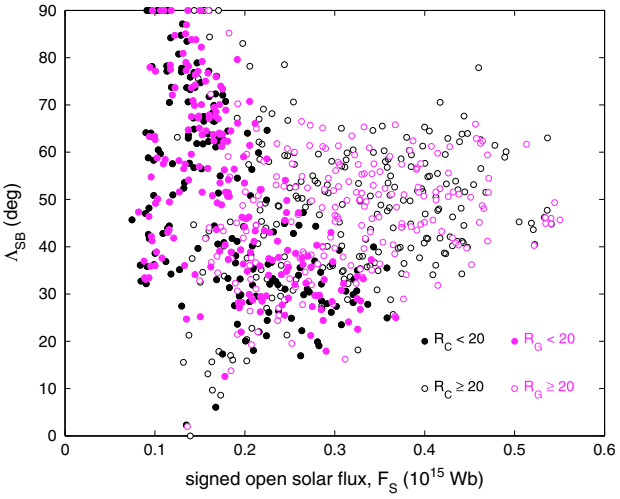

Figure 9. Modeled annual values of the streamer belt half-width, Λ_{SB}, as a function of the modeled open solar flux, F_S. Black points use the corrected and extended international sunspot number R_C as model input, whereas mauve points use the group sunspot number R_G. As in Figure 2, years with sunspot number $R_C < 20$ ($R_G < 20$ in the case of mauve points) are shown as filled circles, and years with $R_C \geq 20$ ($R_G \geq 20$ in the case of mauve points) are shown as open circles.

heliographic latitude, sin(Λ_H), and time (bins are 0.044 by 1 year in size). Black areas, where $(P_{PS} + P_{DS}) = 0$ are polar coronal holes. The occurrence of DP or PS is taken to be an indicator of the SB. The blue lines show the variation of the SB edges predicted by the model (assuming that it is symmetrical so the half-width Λ_{SB} applies in both solar hemispheres). It can be seen that the model has captured both the solar cycle variations and the cycle-to-cycle changes in the minimum SB width well.

The upper panel of Figure 7 is a check on the phasing of the derived F_{CH} variation and the use of an average transfer time of surviving SB flux to coronal holes of 2.4 years. The blue line is the modeled F_{CH} for the last three solar cycles, and the red line gives the variation of the solar polar field observed by the WSO magnetograph (scaled so it can be shown on the same axes). This field is for $< B_p > = (B_N + B_S)/2$ where B_N and B_S are the fields detected over the north and south solar poles, respectively. We would not expect an exact correspondence as F_{CH} is flux and not a field magnitude, and so at solar minimum the variations in the size of the coronal holes are a factor and at sunspot maximum coronal holes occur all latitudes. Nevertheless, the model has produced a variation in F_{CH} that is in phase with the observed $< B_p >$ variation, and the decline in the peak values of $< B_p >$ over the last three solar minima is reflected in a decline in the corresponding modeled F_{CH} peaks.

5. Discussion and Conclusions

Figure 7b shows that the model is predicting a persistently broader SB during the Maunder minimum, as suggested by *Lockwood and Owens* [2014]. Figure 9 plots the modeled latitudinal half-width of the streamer belt (SB), Λ_{SB}, as a function of the open solar flux (OSF) for 1620–2012. Black points in this plot are based on the sunspot numbers series R_C, and mauve points are based on the group sunspot number R_G. The plot is in the same format as Figure 2 with annual modeled points for the relevant sunspot number less than 20 shown by filled circles and for sunspots numbers greater than or equal to 20 shown by open circles. Figure 9 has the general approximate appearance of an inverted version of Figure 2, with the notable feature that the modeled streamer belt width at sunspot minimum increases as OSF decreases for the low sunspot number data. This gives strong support to the idea proposed by *Lockwood and Owens* [2014] that annual means of sunspot minimum solar wind velocities fall to the values of the slow solar wind when OSF is small because the streamer belt is wide and Earth remains immersed in the streamer belt and slow (or perhaps intermediate speed) solar wind flows.

The fraction of the OSF lost from coronal holes (CH) flux is best quantified by matching the predicted streamer belt width to the variations deduced for recent solar cycles using the PFSS method. The value derived sits comfortably within the allowed range between two model behaviors that are deemed unphysical: namely, persistent loss of all SB flux and net migration of CH flux back into the SB. The OSF loss rate is always dominated by the SB flux loss.

The modeled OSF is dominated at most times by the SB flux. The modeled CH flux varies in close to antiphase with the SB flux (it lags by a bit less than half a cycle) over the solar cycle and shows a cycle-to-cycle variation that mirrors in form that in the SB flux and total OSF. The modeled CH flux variation captures the phase and recent declining amplitude trend of the observed variation in solar polar fields. In the Maunder minimum

cycles of OSF are deduced despite the fact that the input open flux emergence rate is at a steady baselevel value. These cycles have been explained by *Owens et al*. [2012] as the effect of the cyclic SB loss, whereas for larger sunspot number cycles the emergence rate cycles dominate the OSF behavior. During the Maunder minimum, the modeled CH flux falls to zero at each maximum of the OSF cycles, but the CH and SB fluxes are close to equal at each of the MM minima in OSF.

The use of corrected international sunspot number R_C instead of group sunspot number R_G makes no difference in principle to the behavior but does make some difference to the detail of the variations. As would be expected, the OSF is larger on either side of the Dalton minimum if R_C instead of R_G is employed. The difference within the Dalton minimum itself is much smaller. The width of the SB reaches the same maxima for R_C and R_G at these times, but the minima seen at sunspot minimum are less deep for R_G than for R_C. This is consistent with the general trend for the sunspot-minimum SB width to increase for lower OSF values. The model outputs (the streamer-belt half-width and the coronal-hole, streamer-belt, and total open solar fluxes) are available in the supplementary data file attached to the present paper. The input sunspot number data are available in the supplementary data file attached to paper 2 [*Lockwood et al*., 2014b]. The open solar flux, solar wind speed and near-Earth IMF reconstructions from historic geomagnetic data are available in the supplementary data file attached to the paper by *Lockwood et al*. [2014d].

Acknowledgments

We thank a great many scientists who made, recorded, and digitized the solar observations employed in this paper. In particular, the international sunspot number data were obtained from WDC-SILSO, Royal Observatory of Belgium, Brussels (URL: http://sidc.be/silso/home/), the group sunspot data from NGDC, Boulder (URL: http://www.ngdc.noaa.gov/stp/solar/), and the magnetograph data from T. Hoeksema of Standford University and the staff of Wilcox Solar Observatory (URL: http://wso.stanford.edu/#Magnetograms/). We also thank the UK Natural Environment Research Council (NERC) for support under NERC grant NE/J024678/1.

Yuming Wang thanks the reviewers for their assistance in evaluating this paper.

References

Antiochos, S. K., Z. Mikic, V. S. Titov, R. Lionello, and J. A. Linker (2011), A model for the sources of the slow solar wind, *Astrophys. J.*, *731*, 112, doi:10.1088/0004-637X/731/2/112.

Arlt, R., R. Leussu, N. Giese, K. Mursula, and I. G. Usoskin (2013), Sunspot positions and sizes for 1825-1867 from the observations by Samuel Heinrich Schwabe, *Mon. Not. Roy. Astron. Soc.*, *433*(4), 3165–3172, doi:10.1093/mnras/stt961.

Beer, J., S. Tobias, and N. Weiss (1998), An active Sun throughout the Maunder minimum, *Solar Phys.*, *181*, 237–249, doi:10.1023/A:1005026001784.

Cliver, E. W., and A. G. Ling (2011), The floor in the solar wind magnetic field revisited, *Solar Phys*, *274*, 285–301, doi:10.1007/s11207-010-9657-6.

Crooker, N. U., J. T. Gosling, and S. W. Kahler (2002), Reducing heliospheric magnetic flux from coronal mass ejections without disconnection, *J. Geophys. Res.*, *107*(A2), 1028, doi:10.1029/2001JA000236.

De Toma, G. (2011), Evolution of coronal holes and implications for high-speed solar wind during the minimum between cycles 23 and 24, *Solar Phys.*, *274*(1–2), 195–217, doi:10.1007/s11207-010-9677-2.

Erdös, G., and A. Balogh (2014), Magnetic flux density in the heliosphere through several solar cycles, *Astrophys. J.*, *781*(1), 50, doi:10.1088/0004-637X/781/1/50.

Finch, I. D., and M. Lockwood (2007), Solar wind-magnetosphere coupling functions on timescales of 1 day to 1 year, *Ann. Geophys.*, *25*, 495–506, doi:10.5194/angeo-25-495-2007.

Fisk, L. A., and N. A. Schwadron (2001), The behaviour of the open magnetic field of the Sun, *Astrophys. J.*, *560*, 425–438, doi:10.1086/322503.

Fisk, L. A., T. H. Zurbuchen, and N. A. Schwadron (1999), Coronal hole boundaries and their interaction with adjacent regions, *Space Sci. Rev.*, *87*, 43–54, doi:10.1023/A:1005153730158.

Fitzenreiter, R. J., K. W. Ogilvie, D. J. Chornay, and J. Keller (1998), Observations of electron velocity distribution functions in the solar wind by the Wind Spacecraft: High angular resolution Strahl measurements, *Geophys. Res. Lett.*, *25*, 249–252, doi:10.1029/97GL03703.

Goelzer, M. L., C. W. Smith, and N. A. Schwadron (2013), An analysis of heliospheric magnetic field flux based on sunspot number from 1749 to today and prediction for the coming solar minimum, *J. Geophys. Res. Space Physics*, *118*, 7525–7531, doi:10.1002/2013JA019404.

Larson, D. E., et al. (1997), Using energetic electrons to probe the topology of the October 18– 20, 1995, magnetic cloud, *Adv. Space Res.*, *20*, 655–658, doi:10.1016/S0273-1177(97)00453-5.

Lockwood, M. (2003), Twenty-three cycles of changing open solar magnetic flux, *J. Geophys. Res.*, *108*(A3), 1128, doi:10.1029/2002JA009431.

Lockwood, M. (2013), Reconstruction and prediction of variations in the open solar magnetic flux and interplanetary conditions, *Living Rev. Solar Phys.*, *10*, 4, doi:10.12942/lrsp-2013-4. [Available at http://www.livingreviews.org/lrsp-2013-4.]

Lockwood, M., and M. J. Owens (2009), The accuracy of using the Ulysses result of the spatial invariance of the radial heliospheric field to compute the open solar flux, *Astrophys J.*, *701*(2), 964–973, doi:10.1088/0004-637X/701/2/964.

Lockwood, M., and M. J. Owens (2011), Centennial changes in the heliospheric magnetic field and open solar flux: the consensus view from geomagnetic data and cosmogenic isotopes and its implications, *J. Geophys. Res.*, *116*, A04109, doi:10.1029/2010JA016220.

Lockwood, M., and M. J. Owens (2013), Comment on "What causes the flux excess in the heliospheric magnetic field?" by E.J. Smith, *J. Geophys. Res. Space Physics*, *118*, 1880–1887, doi:10.1002/jgra.50223.

Lockwood, M., and M. J. Owens (2014), Implications of the recent low solar minimum for the solar wind during the Maunder minimum, *Astrophys. J. Lett.*, *781*, L7, doi:10.1088/2041-8205/781/1/L7.

Lockwood, M., R. Stamper, and M. N. Wild (1999), A doubling of the Sun's coronal magnetic field during the last 100 years, *Nature*, *399*, 437–439, doi:10.1038/20867.

Lockwood, M., R. B. Forsyth, A. Balogh, and D. J. McComas (2004), The accuracy of open solar flux estimates from near-Earth measurements of the interplanetary magnetic field: analysis of the first two perihelion passes of the Ulysses spacecraft, *Ann. Geophys.*, *22*, 1395–1405, doi:10.5194/angeo-22-1395-2004.

Lockwood, M., A. P. Rouillard, and I. D. Finch (2009a), The rise and fall of open solar flux during the current grand solar maximum, *Astrophys. J.*, *700*(2), 937–944, doi:10.1088/0004-637X/700/2/937.

Lockwood, M., M. J. Owens, and A. P. Rouillard (2009b), Excess open solar magnetic flux from satellite data: I. Analysis of the 3rd Perihelion Ulysses Pass, *J. Geophys. Res.*, *114*, A11103, doi:10.1029/2009JA014449.

Lockwood, M., M. J. Owens, and A. P. Rouillard (2009c), Excess open solar magnetic flux from satellite data: II. A survey of kinematic effects, *J. Geophys. Res.*, *114*, A11104, doi:10.1029/2009JA014450.

Lockwood, M., L. Barnard, H. Nevanlinna, M. J. Owens, R. G. Harrison, A. P. Rouillard, and C. J. Davis (2013a), Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr – Part 1: A new geomagnetic data composite, *Ann. Geophys.*, *31*, 1957–1977, doi:10.5194/angeo-31-1957-2013.

Lockwood, M., L. Barnard, H. Nevanlinna, M. J. Owens, R. G. Harrison, A. P. Rouillard, and C. J. Davis (2013b), Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr – Part 2: A new reconstruction of the interplanetary magnetic field, *Ann. Geophys.*, *31*, 1979–1992, doi:10.5194/angeo-31-1979-2013.

Lockwood, M., M. J. Owens, and L. Barnard (2014a), Centennial variations in sunspot number, open solar flux and streamer belt width: 1. Correction of the sunspot number record since 1874, *J.Geophys. Res. Space Physics*, doi:10.1002/2014JA019970.

Lockwood, M., M. J. Owens, and L. Barnard (2014b), Centennial variations in sunspot number, open solar flux and streamer belt width: 2. Comparison with geomagnetic data, *J. Geophys. Res. Space Physics*, doi:10.1002/2014JA019972.

Lockwood, M., H. Nevanlinna, M. Vokhmyanin, D. Ponyavin, S. Sokolov, L. Barnard, M. J. Owens, R. G. Harrison, A. P. Rouillard, and C. J. Scott (2014c), Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 years: 3. Improved representation of solar cycle 11, *Annales Geophys.*, *32*, 367–381, doi:10.5194/angeo-32-367-2014.

Lockwood, M., H. Nevanlinna, L. Barnard, M. J. Owens, R. G. Harrison, A. P. Rouillard, and C. J. Scott (2014d), Reconstruction of geomagnetic activity and near-earth interplanetary conditions over the past 167 years: 4. Near-earth solar wind speed, IMF, and open solar flux, *Annales Geophys.*, *32*, 383–399, doi:10.5194/angeo-32-383-2014.

Mackay, D. H., and A. R. Yeates (2012), The Sun's global photospheric and coronal magnetic fields: Observations and models, *Living Rev. Solar Phys.*, *9*, doi:10.12942/lrsp-2012-6. [Available at http://www.livingreviews.org/lrsp-2012-6.]

Manoharan, P. K. (2012), Three-dimensional evolution of solar wind during solar cycles 22–24, *Astrophys. J.*, *751*, 128, doi:10.1088/0004-637X/751/2/128.

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

Owens, M. J., and N. U. Crooker (2007), Reconciling the electron counterstreaming and dropout occurrence rates with the heliospheric flux budget, *J. Geophys. Res.*, *112*, A06106, doi:10.1029/2006JA012159.

Owens, M. J., N. U. Crooker, and N. A. Schwadron (2008a), Suprathermal electron evolution in a Parker spiral magnetic field, *J. Geophys. Res.*, *113*, A11104, doi:10.1029/2008JA013294.

Owens, M. J., C. N. Arge, N. U. Crooker, N. A. Schwadron, and T. S. Horbury (2008b), Estimating total heliospheric magnetic flux from single-point in situ measurements, *J. Geophys. Res.*, *113*, A12103, doi:10.1029/2008JA013677.

Owens, M. J., N. U. Crooker, N. A. Schwadron, T. S. Horbury, S. Yashiro, H. Xie, O. C. St. Cyr, and N. Gopalswamy (2008c), Conservation of open solar magnetic flux and the floor in the heliospheric magnetic field, *Geophys. Res. Lett.*, *35*, L20108, doi:10.1029/2008GL035813.

Owens, M. J., M. Lockwood, L. Barnard, and C. J. Davis (2011a), Solar cycle 24: Implications for energetic particles and the probability of a new Maunder Minimum, *Geophys. Res. Lett.*, *38*, L19106, doi:10.1029/2011GL049328.

Owens, M. J., N. U. Crooker, and M. Lockwood (2011b), How is open solar magnetic flux lost over the solar cycle?, *J. Geophys. Res.*, *116*, A04111, doi:10.1029/2010JA01603.

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.

Owens, M. J., N. U. Crooker, and M. Lockwood (2013), Solar origin of heliospheric magnetic field inversions: Evidence for coronal magnetic loop opening within pseudostreamers, *J. Geophys. Res. Space Physics*, *118*, 1868–1879, doi:10.1002/jgra.50259.

Owens, M. J., N. U. Crooker, and M. Lockwood (2014), Solar cycle evolution of dipolar and pseudostreamers belts and their relation to the slow solar wind, *J. Geophys. Res. Space Physics*, *119*, doi:10.1002/jgra.50772.

Owens, M., and M. Lockwood (2012), 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, doi:10.1029/2011JA017193.

Parker, E. N. (1958), Dynamics of the interplanetary gas and magnetic fields, *Astrophys. J.*, *128*, 664–676, doi:10.1086/146579.

Riley, P., and J. G. Luhmann (2012), Interplanetary signatures of unipolar streamers and the origin of the slow solar wind, *Sol. Phys.*, *277*(2), 355–373, doi:10.1007/s11207-011-9909-0.

Rouillard, A. P., M. Lockwood, and I. Finch (2007), Centennial changes in the solar wind speed and in the open solar flux, *J. Geophys. Res.*, *112*, A05103, doi:10.1029/2006JA012130.

Schrijver, C. J., M. L. DeRosa, and A. M. Title (2002), What is missing from our understanding of long-term solar and heliospheric activity?, *Astrophys. J.*, *577*(2), 1006–1012, doi:10.1086/342247.

Schwadron, N. A., D. E. Connick, and C. W. Smith (2010), Magnetic flux balance in the heliosphere, *Astrophys. J.*, *722*, L132–L136, doi:10.1088/2041-8205/722/2/L132.

Smith, E. J. (2011), What causes the flux excess in the heliospheric magnetic field?, *J. Geophys. Res.*, *116*, A12101, doi:10.1029/2011JA016521.

Smith, E. J., and A. Balogh (1995), Ulysses observations of the radial magnetic field, *Geophys. Res. Lett.*, *22*, 3317–3320, doi:10.1029/95GL02826.

Solanki, S., M. Schüssler, and M. Fligge (2000), Secular evolution of the Sun's magnetic field since the Maunder minimum, *Nature*, *480*, 445–446, doi:10.1038/35044027.

Steinhilber, F., J. A. Abreu, J. Beer, and K. G. McCracken (2010), Interplanetary magnetic field during the past 9300 years inferred from cosmogenic radionuclides, *J. Geophys. Res.*, *115*, A01104, doi:10.1029/2009JA014193.

Suess, S. T., E. J. Smith, J. Phillips, B. E. Goldstein, and S. Nerney (1996), Latitudinal dependence of the radial IMF component: Interplanetary imprint, *Astron. Astrophys.*, *316*, 304–312.

Svalgaard, L., and E. W. Cliver (2007), Interhourly variability index of geomagnetic activity and its use in deriving the long-term variation of solar wind speed, *J. Geophys. Res.*, *112*, A10111, doi:10.1029/2007JA012437.

Vieira, L. E. A., and S. K. Solanki (2010), Evolution of the solar magnetic flux on time scales of years to millennia, *Astrophys. J.*, *509*(1), A100, doi:10.1051/0004-6361/200913276.