



## RESEARCH ARTICLE

10.1002/2016JA022529

This article is a companion to Owens et al. [2016] doi:10.1002/2016JA022550.

## Key Points:

- The near-Earth magnetic field intensity ( $B$ ) can be reliably reconstructed back to 1750
- Geomagnetic reconstructions of  $B$  provide the most accurate method of reconstructing  $B$  back to 1845
- Uncertainty in the early sunspot record can be mitigated by using a composite approach to reconstructing  $B$

## Supporting Information:

- Supporting Information S1
- Data Set S1
- Data Set S2

## Correspondence to:

M. J. Owens,  
m.j.owens@reading.ac.uk

## Citation:

Owens, M. J., et al. (2016), Near-Earth heliospheric magnetic field intensity since 1750: 1. Sunspot and geomagnetic reconstructions, *J. Geophys. Res. Space Physics*, 121, 6048–6063, doi:10.1002/2016JA022529.

Received 19 FEB 2016

Accepted 7 JUN 2016

Accepted article online 17 JUN 2016

Published online 14 JUL 2016

©2016. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Near-Earth heliospheric magnetic field intensity since 1750: 1. Sunspot and geomagnetic reconstructions

M. J. Owens<sup>1</sup>, E. Cliver<sup>2,3</sup>, K. G. McCracken<sup>4</sup>, J. Beer<sup>5</sup>, L. Barnard<sup>1</sup>, M. Lockwood<sup>1</sup>, A. Rouillard<sup>6</sup>, D. Passos<sup>7,8,9</sup>, P. Riley<sup>10</sup>, I. Usoskin<sup>11</sup>, and Y-M. Wang<sup>12</sup>

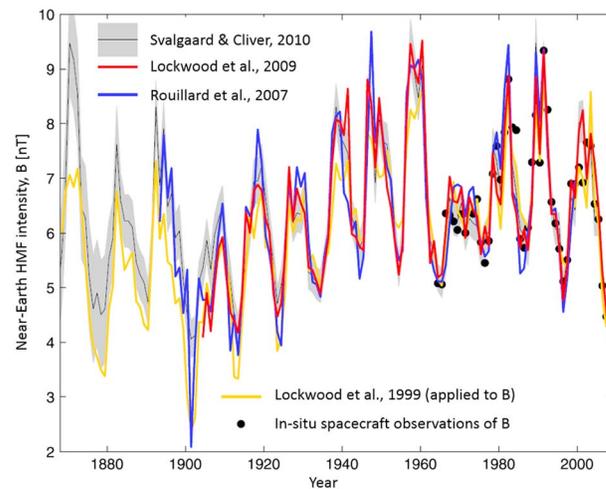
<sup>1</sup>Space and Atmospheric Electricity Group, Department of Meteorology, University of Reading, Reading, UK, <sup>2</sup>National Solar Observatory, Boulder, Colorado, USA, <sup>3</sup>Space Vehicles Directorate, Air Force Research Laboratory, Kirtland, New Mexico, USA, <sup>4</sup>100 Mt. Jellore Lane, Woodlands, New South Wales, Australia, <sup>5</sup>Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland, <sup>6</sup>IRAP/CNRS/UPS, PEPS, Toulouse, France, <sup>7</sup>CENTRA, Instituto Superior Tecnico, Universidade de Lisboa, Lisboa, Portugal, <sup>8</sup>GRPS, Department de Physique, Université de Montreal, Montreal, Quebec, Canada, <sup>9</sup>Departamento de Física, Universidade do Algarve, Faro, Portugal, <sup>10</sup>Predictive Science, San Diego, California, USA, <sup>11</sup>ReSoLVE Centre of Excellence and Sodankylä Geophysical Observatory, University of Oulu, Oulu, Finland, <sup>12</sup>Space Science Division, Naval Research Laboratory, Washington, District of Columbia, USA

**Abstract** We present two separate time series of the near-Earth heliospheric magnetic field strength ( $B$ ) based on geomagnetic data and sunspot number (SSN). The geomagnetic-based  $B$  series from 1845 to 2013 is a weighted composite of two series that employ the interdiurnal variability index; this series is highly correlated with in situ spacecraft measurements of  $B$  (correlation coefficient,  $r = 0.94$ ; mean square error,  $MSE = 0.16 \text{ nT}^2$ ). The SSN-based estimate of  $B$ , from 1750 to 2013, is a weighted composite of eight time series derived from two separate reconstruction methods applied to four different SSN time series, allowing determination of the uncertainty from both the underlying sunspot records and the  $B$  reconstruction methods. The SSN-based composite is highly correlated with direct spacecraft measurements of  $B$  and with the composite geomagnetic  $B$  time series from 1845 to 2013 ( $r = 0.91$ ;  $MSE = 0.24 \text{ nT}^2$ ), demonstrating that  $B$  can accurately be reconstructed by both geomagnetic and sunspot-based methods. The composite sunspot and geomagnetic  $B$  time series, with uncertainties, are provided as supporting information.

### 1. Introduction

Long-term solar wind variability has potentially important implications for studies of both space [Barnard et al., 2011; Usoskin, 2013] and terrestrial regional and global climate [Gray et al., 2010; Lockwood, 2012; Ineson et al., 2015]. The solar wind has only been observed directly for about 50 years [Neugebauer and Snyder, 1962]. Until recently, high-confidence time series of solar wind and heliospheric magnetic field parameters based on geomagnetic data only extended back to ~1900 [Rouillard et al., 2007; Svalgaard and Cliver, 2007b; Lockwood et al., 2009a; Svalgaard and Cliver, 2010]. In this paper we use the sunspot and geomagnetic data sets to consolidate and extend the time series for near-Earth heliospheric magnetic field (HMF) intensity,  $B$ , back to 1750. In the companion paper [Owens et al., 2016], we compare the two  $B$  series derived here (geomagnetic-based and sunspot-based) with the  $B$  series based on cosmogenic radionuclide data.

In his landmark paper on the Maunder Minimum, Eddy [1976] conclusively demonstrated that the Sun is a variable star on long time scales. Lockwood et al. [1999] provided further insight to the nature of such long-term change by using geomagnetic activity indices to show that the Sun's open magnetic flux underwent significant—factor of 2—changes during the course of the last century. The Lockwood et al. study reinvigorated the field of long-term solar variability. After ~15 years of lively research based on sunspot data [e.g., Solanki et al., 2000, 2002; Wang et al., 2005; Tapping et al., 2007; Balmaceda et al., 2009; Schwadron et al., 2010; Vieira and Solanki, 2010; Vieira et al., 2011; Owens and Lockwood, 2012; Clette et al., 2014; Lockwood and Owens, 2014; Lockwood et al., 2014c; Yeo et al., 2014; Clette and Lefèvre, 2016; Svalgaard and Schatten, 2016; Usoskin et al., 2016] and the geomagnetic activity record [e.g., Svalgaard et al., 2003; Mursula et al., 2004; Svalgaard et al., 2004; Svalgaard and Cliver, 2005; Lockwood et al., 2006; Mursula and Martini, 2006; Svalgaard and Cliver, 2006; Rouillard et al., 2007; Svalgaard and Cliver, 2007b; Lockwood et al., 2009b; Svalgaard and Cliver, 2010; Lockwood and Owens, 2011; Lockwood et al., 2009a, 2009b, 2014a, 2014b; Svalgaard, 2014], an



**Figure 1.** Annual means of  $B$  from geomagnetic reconstructions compared with direct spacecraft observations, adapted from Lockwood and Owens [2011]. Black dots show in situ spacecraft data from the OMNI near-Earth data set. The yellow line shows the Lockwood *et al.* [1999] estimate of open solar flux converted to  $B$ . The red and blue lines are reconstructions by Lockwood *et al.* [2009b] and Rouillard *et al.* [2007]. The grey line and shaded region shows the reconstruction from Svalgaard and Cliver [2010] with uncertainties calculated by Lockwood and Owens [2011].

emerging consensus on the reconstruction of  $B$  based on geomagnetic data has been forged for the last century [Svalgaard and Cliver, 2010; Lockwood and Owens, 2011]. As noted by Lockwood and Owens [2011], “This is a significant development because individually, each [method] has uncertainties introduced by instrument calibration drifts, limited numbers of observatories, and the strength of the correlations employed.”

Figure 1, adapted from Lockwood and Owens [2011], shows a comparison of previously published estimates of  $B$  from geomagnetic data (described in the figure caption) with the near-Earth spacecraft values. At the time of study, there was reasonable agreement among various geomagnetic reconstructions for the last ~100 years, though differences were still present. Subsequent versions have converged yet further. For example, Lockwood *et al.* [2014a] used four different combinations of geomagnetic indices to better separate

the effects of HMF and solar wind speed and place  $2\sigma$  uncertainties on each: for 1901 (the year with the greatest disagreement), they derived  $B = 4.7 \pm 0.5$  nT, in general agreement with the ~4.1 nT of Svalgaard and Cliver [2010]. Svalgaard [2014] has replaced Bartel’s early diurnal range values with more appropriate data before 1880. Thus, these changes have resolved many of the larger discrepancies that were present (Figure 1), as will be demonstrated in section 2.2.

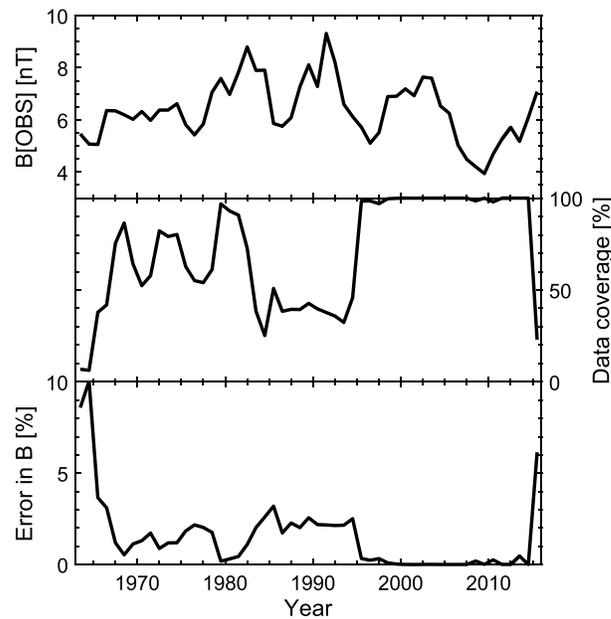
The present study was initiated as part of an International Space Science Institute (ISSI) international team organized by Leif Svalgaard, Mike Lockwood, and Jürg Beer. It extends the consensus solar wind time series for  $B$  back in time to 1750. The present century-long consensus is based primarily on geomagnetic data and, to a lesser extent, on the paleocosmic ray record, discussed in Paper 2. Here we also consider various reconstructions based on the sunspot number for the period from 1750 to the present. The era of consistent high-cadence, uninterrupted sunspot coverage began with Schwabe in 1826 while comparable geomagnetic data originated with Gauss, Weber, and the Göttingen Magnetic Union ~1830 followed by the British Magnetic program in the early 1840s [Cawood, 1979; Cliver, 1994].

In section 2, we discuss the principal data sets—in situ satellite measurements of  $B$  and the sunspot and geomagnetic proxies—and the relationships between the proxies and  $B$ . We combine the various time series to obtain the consensus record for each data type and perform the error analysis to determine the uncertainties. The results are summarized and discussed in section 3.

## 2. Data Analysis

### 2.1. In Situ Spacecraft Observations

The most direct measurement of the near-Earth HMF intensity,  $B$ , comes from in situ magnetometer measurements made by spacecraft in the solar wind. The OMNI dataset [King and Papitashvili, 2005] (see also <http://omniweb.gsfc.nasa.gov/>) collates the near-Earth solar wind measurements from numerous spacecraft, most recently IMP8, Wind, and ACE. This record begins in 1963 and continues to the present day. Figure 2 (top) shows annual means of  $B$  computed from 1 h measurements of the scalar  $B$ , referred to as  $B[\text{OBS}]$  in the remainder of this study. Figure 2 (middle) shows the data coverage for each year. Figure 2 (bottom) shows an estimate of the error in  $B[\text{OBS}]$  resulting from the incomplete data coverage. This is computed by applying the data gaps for a given year ( $Y$ ) to all other years in the OMNI interval and recalculating the yearly means.



**Figure 2.** A summary of the near-Earth spacecraft observations of  $B$  from the OMNI data set. (top) Annual means of  $B$  computed from hourly scalar  $B$  OMNI data. (middle) Percentage data coverage over the year. (bottom) The statistically estimated percentage error in  $B$  resulting from incomplete data coverage.

The error in year  $Y$  is then taken to be the average percentage error across all other years. This approach incorporates both the magnitude of the data coverage and the distribution of data gaps. For example, at the time of data access, 2015 had just over 20% data coverage, which is obviously all in the first part of the year (note that while this could be easily updated to use the most recent data, the geomagnetic and sunspot records to which we will compare are currently complete only to the start of 2014). This results in an estimated percentage error higher than for, e.g., 1985, which has similar data coverage but more evenly distributed throughout the year. The effect of the data gaps and averaging timescale on correlations with other parameters was studied by *Finch and Lockwood [2007]* who showed that optimum correlation required piecewise removal during data gaps of the data that is to be correlated with  $B[\text{OBS}]$ . This can be done to a large extent for geomagnetic data because the annual means are computed from higher resolution data. However, the same could not be done for cosmogenic isotope data because the time resolutions achieved are typically 1 year or lower.

Note also that at high time resolutions (seconds to months) geomagnetic activity depends on the southward component of the HMF in the Earth's magnetic frame rather than on the magnitude of the HMF. *Stamper et al. [1999]* showed that the orientation factor averaged to almost a constant on annual timescales so that the HMF magnitude could be inferred from the geomagnetic data on these timescales. However, even for annual means, this is not quite a constant factor and the small fluctuations place a limit on the maximum correlation between HMF and geomagnetic activity that can be obtained: for annual timescales this limit is approximately 0.95 [*Lockwood, 2013*].

As estimated errors, particularly after 1965, are relatively small, we compare the various  $B$  reconstruction methods to the whole 1963–2013 OMNI data set. For comparison, however, correlations ( $r$ ) and mean square errors (MSEs) are also computed over the 1995–2013 interval, identified throughout the paper by subscript "95," when spacecraft data coverage is essentially complete.

**2.2. Geomagnetic Data**

Geomagnetic activity provides information about both  $B$  and the solar wind flow speed [e.g., *Russell, 1975; Svalgaard, 1977; Feynman and Crooker, 1978; Cliver et al., 1998; Lockwood et al., 1999*]. Here we use two complementary approaches, both based on the interdiurnal variation (IDV) index [*Svalgaard et al., 2003; Svalgaard and Cliver, 2005, 2010*]. Quoting from *Svalgaard and Cliver [2010]*, "the interdiurnal variability (IDV) index for a given geomagnetic observatory ("station") [is] the average difference without regard to sign, from one day to the next, between hourly mean values of the horizontal component,  $H$ , and measured 1 h after midnight [as this is the time most sensitive to solar wind variations]. The average should be taken over a suitably long interval of time, such as 1 year, to eliminate various seasonal complications. IDV has the useful property of being independent of solar wind speed and is highly correlated with [in situ measurements of] the near-Earth heliospheric magnetic field (HMF) strength  $B$ ." The current version of IDV in this study is from *Svalgaard [2014]* and is termed S2014. As with earlier versions, it is based on multiple stations, where possible, to increase the signal-to-noise ratio. Figure 6 in *Svalgaard [2014]* shows that before ~1875, there were only four stations or fewer, dropping to a single station (Helsinki) for about a third of the time from 1845 to

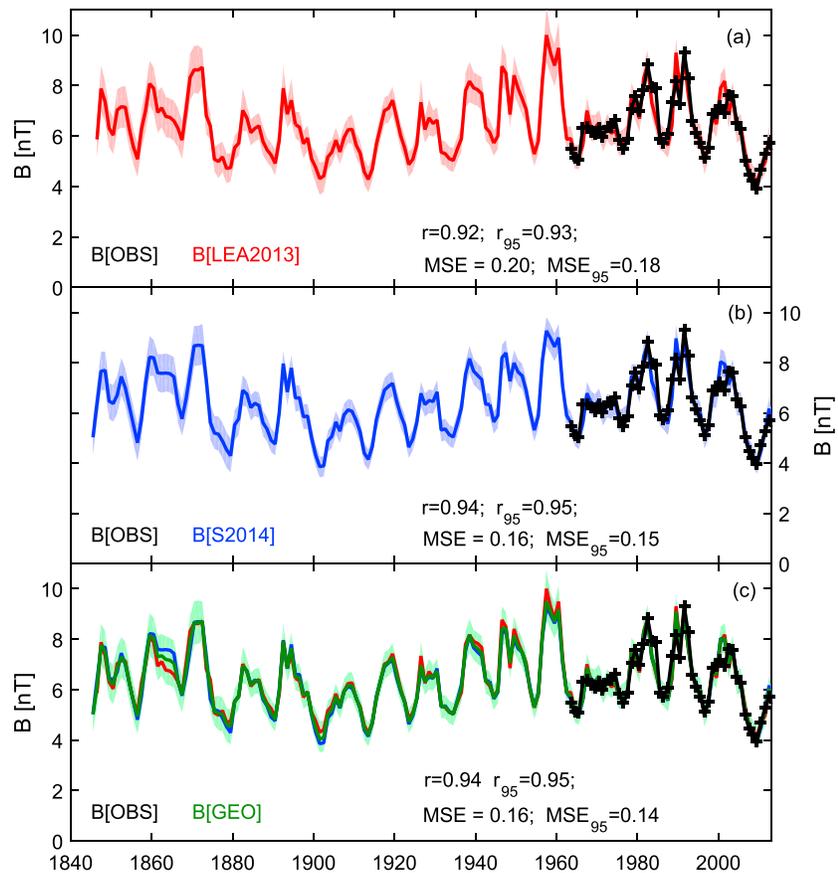
1875. After 1950, IDV is based on ~20 stations. Combining multiple observatories greatly reduces the “noise” in IDV, particularly during the space age, but this comes at the potential cost of changing characteristics of the time series into the past, as the relative weighting of stations, latitudes, quality, and indices vary with data availability.

The second approach is that recently undertaken by Lockwood and coworkers [Lockwood *et al.*, 2013a, 2013b, 2014a, 2014b], hereafter LEA2013, which uses IDV from daily means, IDV(1d), of  $H$ . This definition is equivalent to that for the Bartels’  $u$  index [Mayaud, 1980], except that IDV(1d) is based on a single station at any one time rather than on a network of stations and IDV(1d) corrects for the effects of secular change in the geomagnetic latitudes of the stations used. While IDV mitigates the effect of the regular diurnal magnetic variation by considering only data for the midnight hour, IDV(1d) does so by averaging over all 24 h a day. The IDV(1d) index is based on a single sequence of data from geomagnetically similar locations, namely, Helsinki and Eskdalemuir, with limited gaps filled by Potsdam and Seddin. The latter stations are not ideal in that they are at lower latitude than both Helsinki and Eskdalemuir; however, the magnetic latitudes of all these stations (particularly of Eskdalemuir) also change considerably with time because of the secular variation in the geomagnetic field. IDV(1d) makes allowance for this by using the IGRF and *gufm1* models of the main field (to estimate the geomagnetic latitude of each station in any one year) along with a study of the effect of geomagnetic latitude on the  $u$  values in recent decades (when many stations are available) and normalizes the whole data sequence to the geomagnetic latitude of Eskdalemuir in the year 2000. Thus, IDV(1d), unlike either IDV or  $u$  indices, makes allowance for both the secular variation in the geomagnetic field and the effects of changing the distribution of sites used in compiling the index. The advantage of this approach is that the characteristics of the time series are relatively constant with time. The potential disadvantage is that it is prone to inhomogeneities at a single station [e.g., Lockwood *et al.*, 2014b; Svalgaard, 2014] if the data are not properly vetted, though this also applies to IDV for a few intervals prior to ~1875 when this sequence also relies on just one or two stations. Because it does not use all of the available data in compiling the index, IDV(1d) can use other data as independent tests. Although no one station can provide a test of the whole period, various intervals are covered by different stations and the agreement is good in all cases (with the exception of Greenwich before 1880 for which the required temperature corrections to the magnetometer data cannot satisfactorily be made).

Comparing and combining the  $B$  estimates based on LEA2013 and S2014, denoted B[LEA2013] and B[S2014], has the potential to draw on the individual strengths of the two approaches. As we shall see, the two indices IDV and IDV(1d) are highly correlated. For both IDV and IDV(1d) the correlation coefficient with B[OBS] is very close to the maximum possible ( $\approx 0.95$ ) [Lockwood *et al.*, 2014a].

Figures 3a and 3b show the annual means of B[LEA2013] and B[S2014] time series, respectively. The black lines show annual means from the spacecraft data, B[OBS]. Figure 3 and Table 1 also give the correlation coefficients ( $r$ ) and mean square error (MSE) over the whole 1963–2013 interval, as well as over the 1995–2013 interval, indicated by the subscript “95.” The shaded areas show the uncertainty bands, determined as follows: We find the best geometric-mean fit [York *et al.*, 2004] between B[LEA2013] (or B[S2014]) and B[OBS], then introduce a random error to each point, drawn from a two-sided Gaussian distribution, before refitting the data. This is performed 10,000 times and the 90% confidence interval is taken. The best estimate of  $B$  is then taken to be the median of the 10,000 fits. This Monte Carlo method of uncertainty determination produces a similar result to the standard geometric-mean fit uncertainty [York *et al.*, 2004], but it more readily allows multiple time series to be combined to produce a weighted ensemble or composite, as discussed below.

The top lines of Table 1 list the statistical properties of the best B[LEA2013] and B[S2014] time series relative to the OMNI data over the 1963–2013 interval. (Note that here we work exclusively with whole-year averages of all data, while [Lockwood *et al.*, 2013a, 2013b] computed correlations using piecewise removal of data gaps from the OMNI and geomagnetic data and reported a correlation of 0.95 over the space age. The latter method is preferable for correctly characterizing  $B$  because otherwise the geomagnetic annual means include the influence of intervals which have no effect on the HMF data because they were not measured. In this study, however, the use of whole-year averages is necessary in order to treat all time series in the same manner; we have no means of producing piecewise removal of data gaps for the S2014 series and the sunspot-based estimates.) In general, the S2014 approach produces a better match to the whole-year



**Figure 3.** Annual means of near-Earth heliospheric magnetic field intensity,  $B$ , over the period 1845–2013. Black symbols and lines show  $B$ [OBS], from OMNI spacecraft observations. (a)  $B$ [LEA2013] geomagnetic reconstruction from Lockwood *et al.* [2013b] (red). (b)  $B$ [S2014] geomagnetic reconstruction from Svalgaard [2014] (blue). (c)  $B$ [GEO] (green), the composite resulting from a weighted average of  $B$ [LEA2013] and  $B$ [S2014], which are also plotted individually in red and blue, respectively. Correlation coefficients and mean square errors relative to  $B$ [OBS] are also shown. Shaded areas represent the 90% confidence intervals.  $r$  values show the linear correlation coefficients, MSE the mean square error in  $\text{nT}^2$  over the whole 1963–2013 interval. Subscripts “95” indicate values over the interval 1995–2013 only. See also Table 1.

**Table 1.** Linear Correlation Coefficients,  $r$ , and Mean Square Error (MSE) of Geomagnetic- and Sunspot-Based  $B$  Estimates With  $B$  from OMNI in Situ Spacecraft Measurements from 1963 to 2013<sup>a</sup>

	Linear Correlation Coefficient, ( $r$ )	Linear Correlation Coefficient, ( $r_{95}$ )	Mean Square Error, MSE ( $\text{nT}^2$ )	Mean Square Error, MSE <sub>95</sub> ( $\text{nT}^2$ )	Relative Weighting in Composite
	(1963–2103)	(1995–2013)	(1963–2013)	(1995–2013)	
$B$ [S2014]	0.94	0.95	0.16	0.15	0.58
$B$ [LEA2013]	0.92	0.93	0.20	0.18	0.42
$B$ [GEO] composite	0.94	0.95	0.16	0.14	--
$B$ [ $R_{GS}^{1/2}$ ]	0.83	0.90	0.41	0.25	0.13
$B$ [OL2012( $R_{GS}$ )]	0.88	0.93	0.33	0.17	0.15
$B$ [ $R_I^{1/2}$ ]	0.86	0.92	0.38	0.22	0.14
$B$ [OL2012( $R_I$ )]	0.87	0.94	0.34	0.16	0.15
$B$ [ $R_{IL}^{1/2}$ ]	0.86	0.93	0.38	0.22	0.14
$B$ [OL2012( $R_{IL}$ )]	0.87	0.94	0.34	0.16	0.15
$B$ [ $R_{IC}^{1/2}$ ]	0.86	0.94	0.38	0.16	0.14
$B$ [OL2012( $R_{IC}$ )]	0.87	0.94	0.34	0.16	0.15
$B$ [SSN] composite	0.89	0.94	0.29	0.17	--

<sup>a</sup>Subscripts “95” indicate values over the period 1995–2013. Also shown are the relative weightings of the individual records in the  $B$ [GEO] and  $B$ [SSN] composites.

averages of OMNI data than LEA2013 in terms of  $r$  and MSE. As discussed earlier, this is because the former averages modern geomagnetic data over many stations, thereby suppressing measurement noise. However, the problem is that the regression for this calibration data may not be as accurate for earlier times when few (in some intervals only one) stations are available. In addition, as described above, data gaps in B[OBS] will have had some effect. The LEA2013 method also produces highly accurate reconstructions over the space age, but is designed to be homogeneous over the whole interval so it reduces concerns about the effect of the changing distribution of stations as we go back in time and is not influenced by data gaps in B [OBS]. The close agreement between the two is very powerful as the S2014 approach gives correlations almost as high as they possibly could have been (considering the HMF orientation factor) and the LEA approach confirms that neither the changing availability of observing stations nor the data gaps in B[OBS] have greatly influenced the results. The green line and green-shaded area in Figure 3c.

Figure 3 shows the best and 90% confidence interval for a composite of the two geomagnetic reconstructions, B[GEO]. This is produced by randomly sampling the Monte Carlo ensembles of B[LEA2013] and B [S2014] time series in a proportion weighted by their respective inverse mean square errors relative to OMNI data [e.g., *Armstrong, 2001; Adhikari and Agrawal, 2012*]. The weighting is approximately 3:2, S2014: LEA2013, so there are 10,000 B[S2014] and 7,170 B[LEA2013] time series included in the ensemble that is used to produce B[GEO] composite. Note, however, that this combination may not be ideal, as the B[S2014] achieves a higher correlation than B[LEA2013] partly because it ignores the effect of data gaps in B[OBS]. But in practice, the two series are in such close agreement that resulting composite is unlikely to be significantly affected. The best estimate of  $B$  is the median of the 17,170 ensemble members.

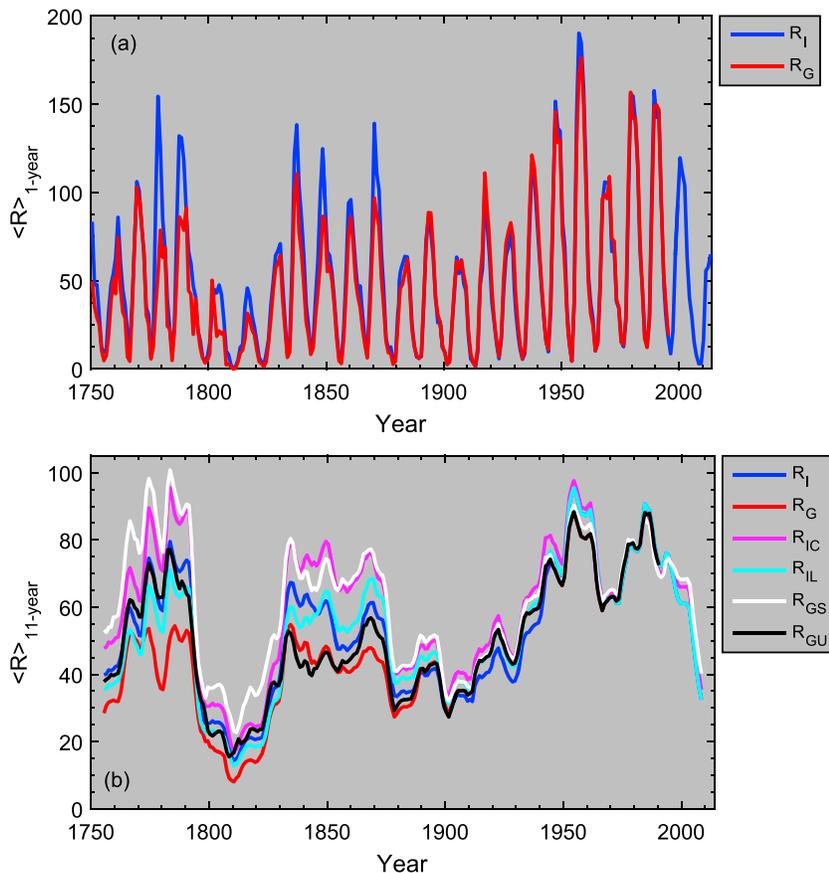
Note that the resulting best B[GEO] composite does not improve on the correlation with B[OBS] obtained with B[SC2014] alone. However, combining the two geomagnetic records ensures robustness prior to 1964, particularly to the uncertainty estimate (the effectiveness of this approach is demonstrated in section 2.3 by comparing sunspot-based estimates over the OMNI and 1845–2013 intervals). Red and blue lines in Figure 3c show the individual B[LEA2013] and B[SC2014] estimates, respectively. They are remarkably similar back to 1845. Given that B[LEA2013] is explicitly constructed to be as homogeneous as possible back in time, this suggests B[GEO] is extremely well reproduced through to the middle of the nineteenth century. The very minor differences between B[LEA2013] and B[SC2014], particularly around the 1870 cycle, are discussed in the context of the B[SSN] estimates, below.

## 2.3. Sunspots

### 2.3.1. Sunspot Records

Sunspot number (SSN) provides the longest record of direct solar observation. Thus, it is desirable to infer  $B$  on the basis of SSN. Here we briefly summarize the sunspot records themselves. The two most widely used sunspot records are the International (also referred to as the Zürich and Wolf) sunspot number,  $R_i$ , which combines the number of groups and individual spots in a particular weighting, and the *Hoyt and Schatten* [1998] group sunspot number (GSN) which only records the number of sunspot groups. GSN can be reliably extended further back in time (to the year 1610, versus 1750 or 1700 for  $R_i$ ) as sunspot groups are expected to be more readily visible than individual spots [e.g., *Usoskin, 2013*], though there may be differences in how different observers define what does or does not constitute a sunspot group. *Hoyt and Schatten* [1998] scaled their GSN up by a constant factor 12.08 to match the magnitude of  $R_i$  over the interval 1874–1976. Here to ease comparison of the relative long-term trends in the different sunspot records, we scale all records to the mean  $R_i$  over 1963–1995, the overlap between routine solar wind observations and the GSN record. As such, the *Hoyt and Schatten* [1998] GSN is here scaled up by 12.15 to yield  $R_G$ . This linear scaling has no effect on the calculation of  $B$  from the sunspot records. Note that both  $R_i$  and  $R_G$  are inherently data composites, comprising different observers, using different equipment in different locations and following different procedures. Figure 4a shows annual  $R_i$  (blue) and  $R_G$  (red) back to 1750. These two time series (and, as a result, the time series for  $B$  derived from them) differ significantly before ~1885. Thus, sunspot number, be it international or group, can become a free parameter in studies of solar and solar terrestrial physics, with the outcome depending on which indicator of solar activity is used.

A series of community meetings has been organized to review and, where appropriate, revise the sunspot number records [*Cliver et al., 2013; Clette et al., 2014; Cliver et al., 2015*]. As a result of this process, discontinuities in the international SSN in ~1946 [*Svalgaard, 2010, 2012; Lockwood et al., 2014c, 2016b*], when



**Figure 4.** Sunspot records considered in this study, covering the interval 1750–2013. To aid comparison, all records have been linearly scaled to the international sunspot number ( $R_I$ , blue line) mean for 1963–1996, the start of routine solar wind observations to the end of the group sunspot record. (a) Comparison of annual  $R_I$  (blue line) and the Hoyt and Schatten [1998] group sunspot number ( $R_G$ , red line). (b) Eleven year running means of all six sunspot records considered in this study. In addition to  $R_I$  and  $R_G$ , they are the Clette and Lefèvre [2016] correction to  $R_I$  ( $R_{IC}$ , pink line), the Lockwood et al. [2014c] correction to  $R_I$  ( $R_{IL}$ , cyan line), the Svalgaard and Schatten [2016] “backbone” group sunspot record ( $R_{GS}$ , white line), and the Usoskin et al. [2016] group sunspot reconstruction ( $R_{GU}$ , black line).

Waldmeier took over from Brunner at Zürich, and in the GSN in ~1885, reflecting a flaw in the Hoyt and Schatten [1998] construction during the time of the transition from Wolf to Wolfser [Svalgaard, 2010; Svalgaard, 2013; Clette et al., 2014; Cliver and Ling, 2016], have been identified. In addition, a post-1980 drift has been detected in  $R_I$  that has been traced to the reference station in Locarno, Switzerland [Clette et al., 2014, 2016]. Leussu et al. [2013] have reported an inhomogeneity in  $R_I$  at 1848 when Wolf replaced Schwabe as the primary observer for this time series. In the present study, we use the published “corrected” series, which are summarized here: An  $R_I$  time series proposed by Lockwood et al. [2014c], here denoted  $R_{IL}$ , employs a 12% upward correction of  $R_I$  before 1946 for the Waldmeier discontinuity and a downward adjustment of 20% before 1849 as suggested by Leussu et al. [2013]. Note that more recent analysis puts the Waldmeier correction at  $12\% \pm 3\%$ , though a simple scaling does not account for observed nonlinear effects [Lockwood et al., 2016a]. In the context of the SSN workshops, Clette et al. [2014] reported a preliminary revision of  $R_I$  (here denoted  $R_{IC}$ ). The  $R_{IC}$  series used here [Clette and Lefèvre, 2016] contains a stronger adjustment for the Waldmeier discontinuity (~18%) than that of Lockwood et al. [2016a], an alternative correction to that suggested by Leussu et al. [2013] for the discontinuity in 1848 (multiplying the original  $R_I$  values from 1849 to 1863 by 1.14) and a post-1980 correction related to the transition of the sunspot curatorship from Zürich to Brussels. It is available from the Sunspot Index and Long-term Solar Observations (SILSO) website ([www.sidc.be/silso](http://www.sidc.be/silso)) as  $S_N$ , V2.0 of the yearly sunspot record. In order to normalize  $R_{IC}$  to  $R_I$  over 1963–1995,  $R_{IC}$  is multiplied by 0.72.  $R_{IL}$  does not require normalization, as it is identical to  $R_I$  over this interval.

A new derivation of GSN was produced by *Svalgaard and Schatten* [2016], here denoted  $R_{GS}$ . The general approach Svalgaard and Schatten used was to reduce the number of “daisy chains” by identifying a limited number of contiguous principal observers. Secondary observers were scaled against these principal observers to extend the time span covered by each “backbone,” thus allowing them to be compared to each other during periods of overlap. The five backbones (designated Staudach, 1739–1822; Schwabe, 1794–1883; Wolfers-Schmidt, 1841–1944; Koyama, 1921–2000; and Locarno, 1950–2015) for the 1750–present interval are pieced together using linear regression, as well as some rather ad hoc (though likely necessary) measures to bridge the data poor interval between the Staudach and Schwabe backbones. Another independent derivation of GSN was recently produced by *Usoskin et al.* [2016] based on a calibration of solar observers to the reference dataset of RGO data for the period 1900–1976 using the active-day statistics. This series (here designated  $R_{GU}$ ) is completely free of daisy-chaining, in contrast to earlier series. These two new group series,  $R_{GU}$  and  $R_{GS}$ , are scaled by factors 12.15 and 13.88 in order to match the magnitude of  $R_1$  over the 1963–1995 interval.

Figure 4b shows 11 year running means of the six sunspot records in order to highlight the differences in the long-term trends. For  $R_{GS}$  and  $R_{IC}$ , (white and pink, respectively) the 1750–1800 period exhibits similar or higher sunspot numbers to the space age, whereas  $R_{IL}$  and  $R_{GU}$  (cyan and black, respectively) show higher values in the space age than 1750–1800. A new construction of the international/Wolf sunspot number [Friedli, 2016] shows long-term trends closer to  $R_{GU}$  than  $R_{GS}$  and  $R_{IC}$ . The differences between the new records will need to be resolved and may necessitate future revisions of SSN-based  $B$  series obtained here. At present, we take the approach that by using multiple sunspot records, we can at least constrain the time-dependent uncertainty in SSN-based estimates of  $B$ .

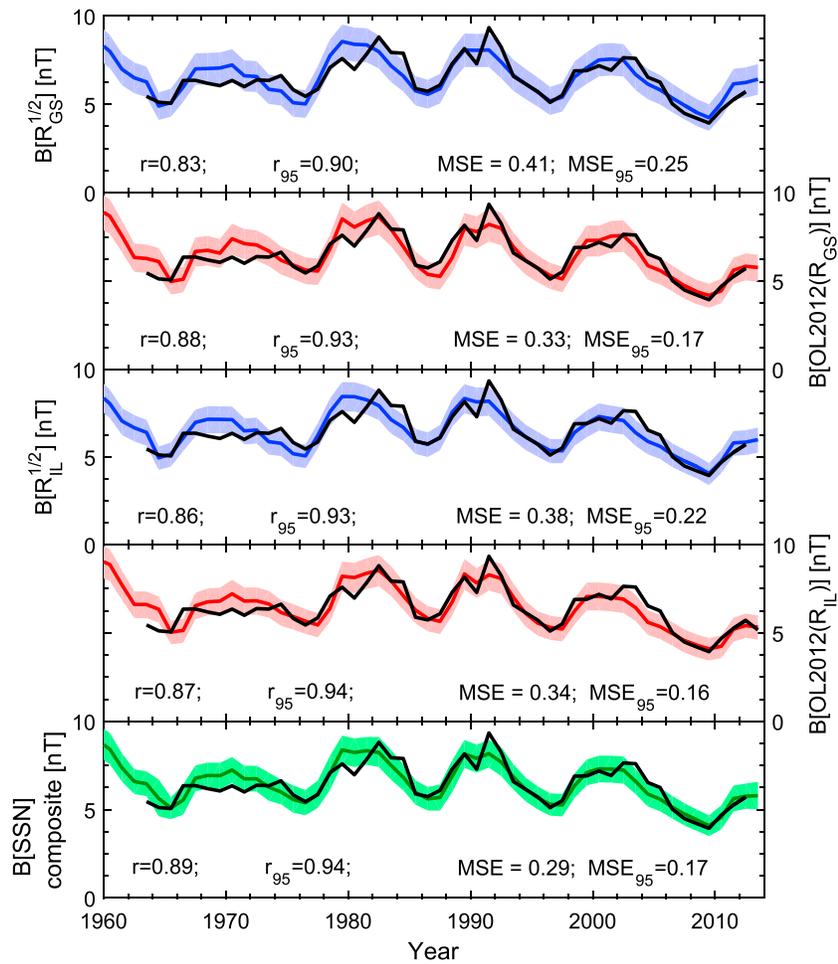
### 2.3.2. Computing $B$ From Sunspot Records

The next step is to compute  $B$  from sunspot records. Two different  $B$  reconstruction methods are used. The first, denoted  $B[SSN^{1/2}]$ , assumes  $B$  is linearly correlated to the square root of sunspot number [Svalgaard and Cliver, 2005; Wang et al., 2005]. We note that regressions between  $SSN^{1/2}$  and  $B$  produce significantly different best fit parameters at different time scales. In particular, the regression of  $SSN^{1/2}$  and  $B$  for annual means captures the general variation over the solar cycle [Svalgaard and Cliver, 2005], but it does not capture the long-term cycle-to-cycle variation seen for solar cycle means [Wang et al., 2005]. Thus, it is necessary to produce a “double regression” over both important time scales, such that  $B[SSN^{1/2}] = A <SSN^{1/2}>_{SC} + B(SSN^{1/2} - <SSN^{1/2}>_{SC}) + C$ , where  $SSN^{1/2}$  is the square root of the annual mean sunspot number,  $<SSN^{1/2}>_{SC}$  is the square root of the solar cycle mean sunspot number, using the solar cycle start/end dates defined by Owens et al. [2011b] on the basis of the average latitude of sunspots.  $A$ ,  $B$  and  $C$  are linear best fit coefficients. This purely empirical approach makes no explicit assumptions about the physical mechanism(s) by which  $B$  varies. It is single valued, in that a given sunspot number gives a single value of  $B$ , regardless of time history. Best linear fits between  $B[OBS]$  and  $SSN^{1/2}$  over the period 1963–2013 give

$$\begin{aligned} B[R_1^{1/2}] &= 0.66 <R_1^{1/2}>_{SC} + 0.27 (R_1^{1/2} - <R_1^{1/2}>_{SC}) + 1.6 \text{ nT} \\ B[R_{IL}^{1/2}] &= 0.66 <R_{IL}^{1/2}>_{SC} + 0.27 (R_{IL}^{1/2} - <R_{IL}^{1/2}>_{SC}) + 1.6 \text{ nT} \\ B[R_{IC}^{1/2}] &= 0.70 <R_{IC}^{1/2}>_{SC} + 0.26 (R_{IC}^{1/2} - <R_{IC}^{1/2}>_{SC}) + 1.2 \text{ nT} \\ B[R_{GS}^{1/2}] &= 0.69 <R_{GS}^{1/2}>_{SC} + 0.27 (R_{GS}^{1/2} - <R_{GS}^{1/2}>_{SC}) + 1.2 \text{ nT} \end{aligned}$$

Note that  $R_G$  and  $R_{GU}$  are not considered here as they are presently only available up to the end of 1995. Clearly, the solar cycle trend will be poorly constrained over these 4.5 cycles. This issue and the implications for a “floor” in  $B$  [Svalgaard and Cliver, 2007a] are addressed in section 2.3.4.

The second method derives  $B$  from an estimate of open solar flux (OSF), which is modeled as a continuity equation, where the OSF source rate,  $S$ , is assumed to vary in phase with SSN [Solanki et al., 2000; Vieira and Solanki, 2010]. This approach incorporates the effect of time history but is based on a number of assumptions about the physical processes by which OSF is produced and lost. We use the recent model of Owens and Lockwood [2012], hereafter OL2012, which uses observed empirical relations and theoretical constraints to reduce the model to a single free parameter. In brief, this model assumes OSF generation associated with coronal mass ejections (CMEs) [Owens and Crooker, 2006] and hence  $S$  varies with CME rate. Prior to the advent of coronagraph observations, sunspot number is used as a proxy for CME rate. Using the best fit between observed CME rate [e.g., Yashiro et al., 2004] and  $R_1$  gives  $S = \phi (0.234 R_1^{0.540} - 0.00153)$  Wb per Carrington rotation, where  $\phi = 10^{12}$  Wb, the average closed flux carried by a CME [Lynch et al., 2005; Owens, 2008]. From theoretical predictions, the OSF loss term ( $L$ ) is assumed to vary with the heliospheric current sheet



**Figure 5.** Time series of annual  $B$  from 1960 to 2013. Direct spacecraft measurements are shown in black. Solid blue and red lines show the results of  $B[SSN]^{1/2}$  and  $B[OL2012]$  reconstruction methods, respectively. Shaded regions show the 90% confidence interval. (first and second panels) The reconstruction methods to  $R_{GS}$  [Svalgaard and Schatten, 2016] and (third and fourth panels)  $R_{IL}$  [Lockwood et al., 2014c],  $R_{GU}$  [Usoskin et al., 2016] and  $R_{IC}$  [Clette and Lefèvre, 2016] based reconstructions, not shown, are very similar. (fifth panel) A weighted composite time series created from an ensemble of the six  $B$  time series resulting from the two  $B$  reconstruction methods applied to the three corrected sunspot number series (i.e.,  $R_{GS}$ ,  $R_{IL}$ , and  $R_{IC}$ ).

(HCS) tilt [Sheeley et al., 2001; Owens et al., 2011a]. OL2012 uses an observationally derived idealized HCS tilt which is a function of solar cycle phase and is invariant from cycle to cycle. L is related to HCS tilt through a free best fit parameter. OSF is converted to  $B$  by the observed relation between OSF and  $B$ , given by equation 12 of Lockwood et al. [2014a].

### 2.3.3. Comparing $B[SSN]$ With $B[OBS]$

The various SSN-based estimates of  $B$  are first compared with the OMNI spacecraft observations,  $B[OBS]$ . Figure 5 shows annual time series of  $B[OBS]$  (black) and the sunspot-based reconstructions using the  $SSN^{1/2}$  (blue) and OL2012 (red) methods. The 90% confidence intervals are shown as shaded areas, computed using the same method as section 2.2. Only  $R_{GS}$  and  $R_{IL}$  are shown, as  $R_I$  and  $R_{IC}$  are very similar to those of  $R_{IL}$  over the space age. Figure 5 (first to fourth panels) clearly show that both  $B$  reconstruction methods, using both group- and international-based records, track the observed solar cycle variation and the trends in cycle-to-cycle variability in  $B$ , though the full range of the solar cycle variation is generally underestimated. It is not clear if this is an effect of large hemispheric differences in sunspot number, which are particularly apparent around solar maximum [e.g., Lockwood et al., 2012, Figure 7]. Table 1 summarizes the linear correlation coefficients ( $r$ ) and mean square errors (MSEs) between  $B[OBS]$  and sunspot-based estimates over the 1963–2013 interval. Subscripts “95” refer to the interval 1995–2013. As expected, there are strong

correlations with  $B[SSN^{1/2}]$  using all four sunspot records, though these correlations are slightly lower than the equivalent  $B[OL2012]$  estimates. Similarly, for  $B[SSN^{1/2}]$  the MSE values are slightly higher than the  $B[OL2012]$  equivalent. It is important to note, however, that while the OL2012 approach provides a slightly better match to the space age observations than the  $SSN^{1/2}$  method, this does not necessarily mean it will continue to provide a better reconstruction outside this period. We note that all sunspot-based reconstructions struggle to reproduce the  $B[OBS]$  variation in cycle 20 as accurately as for other cycles. Figure 2 shows that this cycle has more  $B[OBS]$  data gaps than cycles 23 and 24 but fewer than cycles 21 and 22. Data gaps or errors in calibration of early  $B[OBS]$  data do not appear to be adequate explanations of the cycle 20 differences. Furthermore, geomagnetic reconstructions, which do not suffer from data outages in this period, do match  $B[OBS]$  in this cycle (Figure 3).

We now construct a composite  $B[SSN]$  time series based on the three corrected SSN series with complete coverage of the space age ( $R_{GS}$ ,  $R_{IL}$ ,  $R_{IC}$ ). As with the geomagnetic reconstructions, the individual time series are weighted by their “track record,” namely, the inverse of the mean square error relative to  $B[OBS]$ , as listed in Table 1. Figure 5 (fifth panel) shows the resulting composite  $B[SSN]$  time series. Note that the composite  $B[SSN]$  correlation coefficient with  $B[OBS]$  is actually higher than any individual sunspot-based estimate for the 1963–2013 interval. Similarly, the composite MSE is lower than any individual value. Thus, here the ensemble approach both improves the “best” estimate and reduces the sensitivity to any one method, meaning it is likely to be more robust outside the “training interval” of 1963–2013 than any one individual method (this is further demonstrated in section 2.3.4). More importantly, it provides an estimate of when the different reconstructions do/do not agree, allowing us to estimate uncertainty in the reconstructions resulting from both the methods and the sunspot records. Note, however, that the training interval for the models is the same as the validation interval, which is not desirable. In this instance, this approach is the best possible with the available data, as subdividing the OMNI data into a training and validation intervals leaves insufficient data to adequately constrain the models and Figure 2 shows that the effect of data gaps would be considerably higher in one half of the  $B[OBS]$  data series than the other.

#### 2.3.4. Comparing $B[SSN]$ With $B[GEO]$

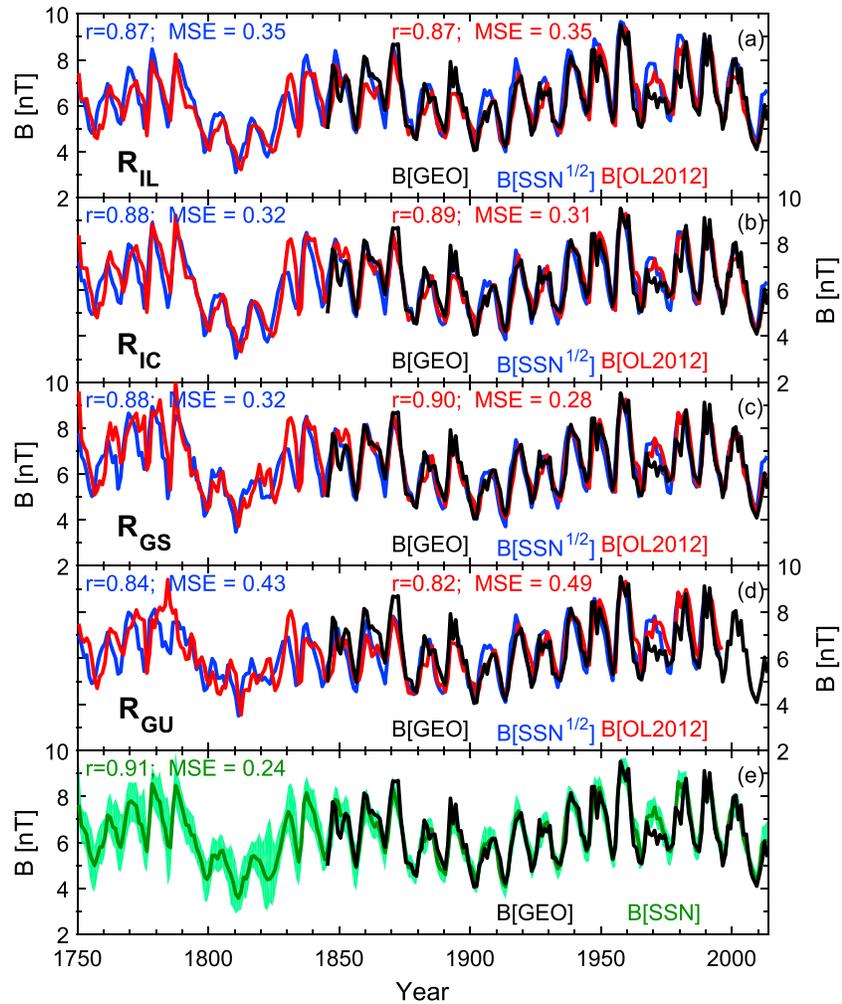
While  $B[SSN]$ , the optimum sunspot-based reconstruction of  $B$ , captures over 80% of the observed variation in  $B$  over the OMNI era, differences between the various corrected sunspot records before this time mean it can not necessarily be expected to provide the same level of accuracy prior to 1963. This same caveat does not apply to the geomagnetic reconstructions; therefore, we use the  $B[GEO]$  composite to recalibrate the weighting factors for the  $B[SSN]$  composite; essentially, we are assuming  $B[GEO]$  provides a “ground truth” for these purposes. In particular, this is desirable for constraining the cycle-to-cycle variation in the  $B[SSN^{1/2}]$  method.

The best fits of  $B[SSN^{1/2}]$  to the  $B[GEO]$  composite are very similar to those obtained by with comparison with OMNI data. Taking the interval 1845–1996, which is covered by all sunspot records, we obtain:

$$\begin{aligned}
 B[R_1^{1/2}] &= 0.48 \langle R_1^{1/2} \rangle_{SC} + 0.32 (R_1^{1/2} - \langle R_1^{1/2} \rangle_{SC}) + 3.2 \text{ nT} \\
 B[R_G^{1/2}] &= 0.44 \langle R_G^{1/2} \rangle_{SC} + 0.33 (R_G^{1/2} - \langle R_G^{1/2} \rangle_{SC}) + 3.5 \text{ nT} \\
 B[R_{IL}^{1/2}] &= 0.53 \langle R_{IL}^{1/2} \rangle_{SC} + 0.31 (R_{IL}^{1/2} - \langle R_{IL}^{1/2} \rangle_{SC}) + 2.6 \text{ nT} \\
 B[R_{IC}^{1/2}] &= 0.55 \langle R_{IC}^{1/2} \rangle_{SC} + 0.30 (R_{IC}^{1/2} - \langle R_{IC}^{1/2} \rangle_{SC}) + 2.4 \text{ nT} \\
 B[R_{GS}^{1/2}] &= 0.59 \langle R_{GS}^{1/2} \rangle_{SC} + 0.32 (R_{GS}^{1/2} - \langle R_{GS}^{1/2} \rangle_{SC}) + 2.1 \text{ nT} \\
 B[R_{GU}^{1/2}] &= 0.47 \langle R_{GU}^{1/2} \rangle_{SC} + 0.32 (R_{GU}^{1/2} - \langle R_{GU}^{1/2} \rangle_{SC}) + 3.3 \text{ nT}
 \end{aligned}$$

Note that by extending the reconstructions back through the Dalton minimum (~1800–1825), they are being used in a parameter regime slightly outside that of the training interval (1845–1996). The lowest possible  $B$  value implied by these empirical relations (i.e., the  $B$  achieved when both the solar cycle and annual means of SSN are zero) implies a “floor” in  $B$  of approximately 2–3 nT. This is well below the  $B$  “floor” value of ~4.6 nT estimated by *Svalgaard and Cliver* [2007a] on the basis of the available  $B$  reconstructions at the time. The annual mean of the observed  $B$  for 2009 was 3.9 nT, demonstrating the importance of considering both the solar cycle and the cycle-to-cycle variations in  $B$ , as well as using regressions with caution outside of the observed parameter regime. That said, there is a ~2.8 nT estimate of the floor [*Cliver and Ling*, 2011; *Cliver and von Steiger*, 2015] that falls within the range of nominal values given above.

Figures 6a and 6b show the individual  $B[SSN^{1/2}]$  (blue) and  $B[OL2012]$  (red) sunspot-based estimates over the period 1750–2013 for the four corrected sunspot records. For comparison, the  $B[GEO]$  composite is shown in



**Figure 6.** Sunspot-based reconstructions of  $B$  over the period 1750–2013 compared with the  $B[GEO]$  composite, shown in (black). Blue lines show the results of the  $B[SSN^{1/2}]$  reconstruction method, while red lines show  $B[OL2012]$ . Figures 6a–6d show the individual corrected sunspot number estimates, specifically: (a)  $R_{IL}$ , (b)  $R_{IC}$ , (c)  $R_{GS}$ , and (d)  $R_{GU}$ . Figure 6e shows a weighted composite of these eight individual  $B$  estimates, as described in the text. Note the y axes begins at 2 nT in each panel, to allow more easy comparison of the differences between the time series.

black. Table 2 summarizes the linear correlation coefficients and MSEs relative to the  $B[GEO]$  composite over the period 1845–1996 for all sunspot records.

Looking first at the statistical differences between the sunspot records listed in Table 2, the highest correlations and lowest mean square errors (for both  $B$  reconstruction methods) are generally found for  $R_{GS}$ ,  $R_{IL}$ , and  $R_{IC}$  sunspot records. The  $B[R_G]$  reconstructions provide the biggest departures from the  $B[GEO]$  composite.  $B[R_{GU}]$  is broadly similar to  $B[R_G]$ , though does provide a closer match to  $B[GEO]$ .

For both reconstruction methods,  $R_{IC}$  and  $R_{IL}$  provide a better match to  $B[GEO]$  than  $R_I$ . While the accuracy of the  $B[SSN]$  reconstructions should not be taken as a direct measure of the accuracy of the underlying SSN record, this does provide indirect support for, e.g., the correction for the Waldmeier discontinuity. Similarly,  $R_{GS}$  and  $R_{GU}$  provides a better match to  $B[GEO]$  than  $R_G$ , again indirectly suggesting that the corrections for the discontinuity circa. 1885 are merited. Both of the legacy series,  $R_I$  and  $R_G$ , have demonstrated flaws, as discussed in section 2.3.1. Thus, we do not consider  $R_I$  and  $R_G$  in our composite SSN series for  $B$ . This has the further advantages of reducing “double counting” of nonindependent data sets, since both  $R_{IC}$  and  $R_{IL}$  are based on independent corrections of  $R_I$  and of excluding the markedly low values of the early part of the  $R_G$  series.

**Table 2.** Linear Correlation Coefficients ( $r$ ) and Mean Square Error (MSE) Between Sunspot-Based Estimates of  $B$  and the Geomagnetic Composite, B[GEO], Over the Period 1845–1996<sup>a</sup>

B[SSN] Method	Linear Correlation Coefficient, $r$ (1845–1996)	Mean Square Error, MSE ( $\text{nT}^2$ ) (1845–1996)	Relative Weighting in Composite
$R_1^{1/2}$	0.87	0.38	--
OL2012( $R_1$ )	0.85	0.41	--
$R_G^{1/2}$	0.83	0.48	--
OL2012( $R_G$ )	0.80	0.53	--
$R_{IL}^{1/2}$	0.87	0.35	0.12
OL2012( $R_{IL}$ )	0.87	0.35	0.12
$R_{IC}^{1/2}$	0.88	0.32	0.13
OL2012( $R_{IC}$ )	0.89	0.31	0.14
$R_{GS}^{1/2}$	0.88	0.32	0.14
OL2012( $R_{GS}$ )	0.90	0.28	0.15
$R_{GU}^{1/2}$	0.84	0.43	0.10
OL2012( $R_{GU}$ )	0.82	0.49	0.09
B[SSN] composite	0.91	0.24	--

<sup>a</sup>Also shown are the relative weightings of the individual records in the final B[SSN] composite.

When creating the composite B[SSN] series, it is tempting to only include the individual B[SSN] records with the best match to B[GEO] (e.g., those based on  $R_{IL}$ ,  $R_{IC}$  and  $R_{GS}$ ) or to optimize the choice of individual B[SSN] records included in the composite to maximize the agreement between the composite B[SSN] and B[GEO]. But this shows this would both fail to give an accurate assessment of the uncertainty prior to the B[GEO] interval and necessarily assumes that the individual records providing the best match over the 1845–2013 interval continue to do so prior to it. (Note that on the basis of comparisons in section 2.3.3 over 1963–2013, we may have included  $R_1$  in the B[SSN] composite but excluded  $R_{GS}$ . The extended 1845–2013 interval considered here shows would have been a poor choice.) Thus, it is necessary to ensure “diversity” in the composite. We assume both reconstruction methods and all four corrected sunspot time series contain useful information on  $B$ . Thus, we produce a composite B[SSN] series in which these eight individual time series are weighted by their MSE relative to B[GEO] over 1845–1996, as shown in Table 2. The composite, however, is extended forward to 2013 using  $R_{GS}$ ,  $R_{IC}$ , and  $R_{IL}$  only. This allows a small test of the composite outside of the training interval.

The B[SSN] composite is shown in green in Figure 6e, with the shaded area showing the 90% confidence interval. The increasing uncertainty further back in time is a direct result of the divergence of the sunspot records prior to ~1870. The match between B[SSN] and B[GEO], shown in black, over the whole 1845–2013 interval is very good. Both the linear correlation coefficient and the mean square error of the composite ( $r=0.91$  and  $\text{MSE}=0.24 \text{ nT}^2$ ) are better than any one individual time series, showing the benefit of this approach. There are only three short periods where B[SSN] deviates from B[GEO] by more than the calculated uncertainty. The most recent period is around 1980, during the rise phase of cycle 21, in which B[SSN] overestimates B[GEO] by just over 1 nT. The most prominent period of disagreement, however, is the “flat”  $B$  variation in solar cycle 20, centered around 1970, which results in an approximately 1 nT B[SSN] overestimate at solar maximum. The large  $B$  in the solar maximum of solar cycle 11, occurring around 1870, is underestimated by B[SSN] by around 1 nT. We note that the geomagnetic data used in both the LEA2013 and the S2014 reconstructions were specifically corrected during this period, with the effect of raising the reconstructed B[GEO] at this time.

### 3. Discussion and Conclusions

#### 3.1. Summary

During the last ~60 years, ground- and space-based observations have provided detailed information regarding the spatial and temporal variability of the solar magnetic fields, providing insight into the operation of the solar dynamo [Charbonneau, 2010; Hathaway, 2010]. In addition, satellites have provided observations of the heliospheric magnetic fields from the vicinity of Earth, to the outer edge of the heliopause, both in and out of the ecliptic [Owens and Forsyth, 2013]. In this paper, we have compared the in situ observations of solar wind with geomagnetic observations and sunspot time series to infer the solar wind

magnetic field strength from 1750 to 2013. In the companion paper [Owens *et al.*, 2016] we compare these geomagnetic- and sunspot-based time series for  $B$  with corresponding time series based on the cosmogenic radionuclide record. All composite time series are provided in the supporting information. Our main results are summarized below where, for brevity, we refer to the estimates as B[OBS], B[GEO], and B[SSN]:

1. For the interval 1963–2013, for which satellite observations of  $B$  are available, both B[GEO] and B[SSN] exhibit excellent agreement with the annual averages of B[OBS].
2. Section 2.2 evaluated two complementary estimation procedures for B[GEO] [Lockwood *et al.*, 2013b; Svalgaard, 2014]. Both are based on the interdiurnal variation observed at geomagnetic observatories since 1845 but differ in the number of geomagnetic observatories and data utilized. The correlation coefficients with the satellite data, B[OBS], since 1963 are very high (0.92 and 0.94). A composite of the two was chosen as being the more robust outside of the 1963–2013 interval over which it was developed. For the composite B[GEO] the correlation coefficient with B[OBS] is 0.94 and the mean square error is  $0.16 \text{ nT}^2$ . This composite B[GEO] record, with uncertainties, is provided in the supporting information.
3. In section 2.3 we have evaluated two separate estimation techniques for computing B[SSN]; (a) the square root of the sunspot number [Wang *et al.*, 2005] over both solar cycle and annual time scales and (b) a continuity model of open solar flux [Solanki *et al.*, 2000] which assumes the source varies as the sunspot number and loss rate varies as tilt of the heliospheric current sheet [Owens and Lockwood, 2012]. For each method we have considered both the international sunspot number and the group sunspot number, both with and without recently proposed corrections included (for further discussion, see the Topical Issue of Solar Physics, “Sunspot Number Recalibration,” with guest editors F. Clette, E.W. Cliver, L. Lefèvre, J.M. Vaquero, and L. Svalgaard). We have defined a composite of six estimates of  $B$  based on two reconstruction techniques applied to three corrected sunspot series. Again, we argue that this composite is likely to be more robust outside the “space age” than any one individual estimate but note that reconstructions outside the parameter regime covered by the training interval, such as Grand Minima in this instance, must be interpreted with care. For the composite B[SSN], the linear correlation coefficient with B[OBS] is 0.89 and mean square error of  $0.29 \text{ nT}^2$ .
4. There are overlapping B[GEO] and B[SSN] estimates for the interval 1845–1996. In section 2.3.4, we compared estimates based on a number of different sunspot number formulations with the B[GEO] composite. Higher correlations are obtained for B[SSN] computed from corrected sunspot records (both group sunspot number [Svalgaard and Schatten, 2016; Usoskin *et al.*, 2016] and international sunspot number [Lockwood *et al.*, 2014c; Clette and Lefèvre, 2016]). A weighted composite of the eight  $B$  estimates, using the four corrected sunspot records and both B[SSN] estimation techniques, is likely to provide the most robust estimate  $B$  and its uncertainty outside of the development period (1845–1996). Compared with B[GEO], this provides a linear correlation coefficient of 0.91 and mean square error of  $0.24 \text{ nT}^2$ , an improvement over any one individual sunspot-based  $B$  estimate. This composite B[SSN] record over the period 1750–2013, with uncertainties, is provided in the supporting information.

### 3.2. Discussion

A brief review of research on the solar wind in the past is helpful to put the above results in perspective. In their pioneering paper in 1978, 16 years after the discovery of the solar wind, Feynman and Crooker [1978] used a relationship between the  $aa$  geomagnetic index and the product of the solar wind speed ( $V$ ) and the southward component of the solar wind field strength ( $B_s$ ) to bound  $V$  between  $237\text{--}395 \text{ km s}^{-1}$  and  $B_s$  between 0.65 and  $1.82 \text{ nT}$  for the solar minimum year of 1901. More than twenty years would pass before the seminal paper by Lockwood *et al.* [1999] that separated the  $V$  and  $B$  dependencies of the  $aa$  index and used the discovery by the Ulysses satellite of the constancy of the radial field with latitude to determine the open magnetic flux of the Sun on an annual basis from 1868 to 1996. The pace of research accelerated. Svalgaard *et al.* [2003] introduced the IDV index which “has the useful property that its yearly averages are highly correlated with the solar wind magnetic field strength ( $B$ ) and are independent of solar wind speed ( $V$ )”. Partly because open solar flux estimates were wrongly interpreted as being linearly proportional to  $B$  and partly because the 1901 value of  $B$  was erroneously low in some reconstructions, a vigorous debate broke out over the reconstruction of  $B$  during the past 100 years [Lockwood *et al.*, 2006; Svalgaard and Cliver, 2006]. The debate was largely settled by 2010 when a consensus emerged [Svalgaard and Cliver, 2010; Lockwood and Owens, 2011].

While the composite time series for  $B$  from 1845 to 2013 based on geomagnetic data and from 1750 to 2013 based on sunspot data presented here represent significant progress, much remains to be done. With the recent publication of the revised sunspot series of *Svalgaard and Schatten* [2016] and *Lockwood et al.* [2014c], the SSN-based  $B$  series can be extended back to 1610, although the uncertainties will be unavoidably large for the Maunder Minimum (~1650–1710) period [*Usoskin et al.*, 2015; *Vaquero et al.*, 2015]. Moreover, as noted in section 2.3.1, differences between the sunspot number time series proposed by *Svalgaard and Schatten* [2016] and *Clette and Lefèvre* [2016] with those of *Lockwood et al.* [2014c] and *Usoskin et al.* [2016] will need to be resolved and may necessitate a revision of the SSN-based  $B$  series presented here. Any such changes will not affect our geomagnetic-based  $B$  series. An improvement to B[GEO] would be a proper allowance for the effect of data gaps in the B[OBS] data series, as although that was achieved in the B[LEA2013] series it was not exploited in the present paper. An update/extension of the solar wind speed time series [*Rouillard et al.*, 2007; *Svalgaard and Cliver*, 2007a] based on the new geo-based  $B$  series is in preparation. Geomagnetic data is sporadically available after 1722, making possible further potential extensions of geomagnetic-based  $B$  and solar wind speed series in the past, though assessing the quality of such data and intercalibrating them with the current records will be challenging. In Part 2 of this study, the series derived here from geomagnetic and sunspot records are compared to cosmogenic radionuclide records [*Owens et al.*, 2016].

#### Acknowledgments

This work was partly facilitated by an International Space Science Institute (ISSI) international team selected in 2011 as number 233, "Long-term reconstructions of solar and solar wind parameters" organized by L. Svalgaard, M. Lockwood, and J. Beer. We thank ISSI for support of the team. We are grateful to the Space Physics Data Facility (SPDF) and National Space Science Data Center (NSSDC) for OMNI data. M.O., M.L., and L.B. are part funded by Science and Technology Facilities Council (STFC) grant ST/M000885/1. M.O. 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. I.U.'s contribution was made within the framework of the ReSolVE Centre of Excellence (Academy of Finland, Project 272157). Y.M.W. acknowledges support from the Chief of Naval Research. We thank both referees for useful comments.

#### References

- Adhikari, R., and R. Agrawal (2012), A novel weighted ensemble technique for time series forecasting, in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 38–49, Springer, Berlin.
- Armstrong, J. S. (2001), *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Kluwer Acad., Norwell, Mass.
- Balmaceda, L., S. Solanki, N. Krivova, and S. Foster (2009), A homogeneous database of sunspot areas covering more than 130 years, *J. Geophys. Res.*, *114*, A07104, doi:10.1029/2009JA014299.
- Barnard, L., M. Lockwood, M. A. Hapgood, M. J. Owens, C. J. Davis, and F. Steinhilber (2011), Predicting space climate change, *Geophys. Res. Lett.*, *38*, L16103, doi:10.1029/2011GL048489.
- Cawood, J. (1979), The magnetic crusade: Science and politics in early Victorian Britain, *Isis*, *70*, 493–518, doi:10.1086/352338.
- Charbonneau, P. (2010), Dynamo models of the solar cycle, *Liv. Rev. Sol. Phys.*, *7*(3), doi:10.1007/lrsp-2005-2.
- Clette, F., and L. Lefèvre (2016), The new sunspot number: Assembling all corrections, *Sol. Phys.*, in press.
- Clette, F., L. Svalgaard, J. M. Vaquero, and E. W. Cliver (2014), Revisiting the sunspot number, *Space Sci. Rev.*, *186*(1–4), 35–103, doi:10.1007/s11214-014-0074-2.
- Clette, F., L. Lefèvre, M. Cagnotti, S. Cortesi, and A. Bulling (2016), The revised Brussels-Locarno Sunspot Number (1981–2015), *Sol. Phys.*, doi:10.1007/s11207-016-0875-4, in press.
- Cliver, E. (1994), Solar activity and geomagnetic storms: The first 40 years, *EOS Trans.*, *75*, 569–569, doi:10.1029/94EO02041.
- Cliver, E., and R. von Steiger (2015), Minimal magnetic states of the Sun and the solar wind: implications for the origin of the slow solar wind, *Space Sci. Rev.*, 1–21, doi:10.1007/s11214-015-0224-1.
- Cliver, E. W., and A. G. Ling (2011), The floor in the solar wind magnetic field revisited, *Sol. Phys.*, *274*, 285–301, doi:10.1007/s11207-010-9657-6.
- Cliver, E. W., and A. G. Ling (2016), The discontinuity circa 1885 in the group sunspot number, *Sol. Phys.*, 1–22, doi:10.1007/s11207-015-0841-6.
- Cliver, E. W., V. Boriakoff, and K. H. Bounar (1998), Geomagnetic activity and the solar wind during the Maunder Minimum, *Geophys. Res. Lett.*, *25*(6), 897–900, doi:10.1029/98GL00500.
- Cliver, E. W., F. Clette, L. Svalgaard, and J. M. Vaquero (2015), Recalibrating the Sunspot Number (SN): The 3rd and 4th SN Workshops, *Cent. Eur. Astrophys. Bull.*, doi:10.1007/s11207-015-0841-6, in press.
- Cliver, E., F. Clette, and L. Svalgaard (2013), Recalibrating the sunspot number (SSN): The SSN workshops, *Cent. Eur. Astrophys. Bull.*, *37*(2), 401–416.
- Eddy, J. A. (1976), The Maunder minimum, *Science*, *192*(4245), 1189–1202, doi:10.1126/science.192.4245.1189.
- Feynman, J., and N. Crooker (1978), The solar wind at the turn of the century, *Nature*, *275*, 626–627, doi:10.1038/275626a0.
- Finch, I., and M. Lockwood (2007), Solar wind-magnetosphere coupling functions on timescales of 1 day to 1 year, *Ann. Geophys.*, *25*(2), 495–506, doi:10.5194/angeo-25-495-2007.
- Friedli, T. (2016), The construction of the Wolf Series from 1749 to 1980, *Sol. Phys.*, doi:10.1007/s11207-016-0907-0, in press.
- Gray, L. J., J. Beer, M. Geller, J. D. Haigh, M. Lockwood, K. Matthes, U. Cubasch, D. Fleitmann, G. Harrison, and L. Hood (2010), Solar influences on climate, *Rev. Geophys.*, *48*, RG4001, doi:10.1029/2009RG000282.
- Hathaway, D. H. (2010), The solar cycle, *Liv. Rev. Sol. Phys.*, *7*, 1, doi:10.1007/lrsp-2010-1.
- Hoyt, D. V., and K. H. Schatten (1998), Group sunspot numbers: A new solar activity reconstruction, *Sol. Phys.*, *181*, 491–512, doi:10.1023/A:1005056326158.
- Ineson, S., A. C. Maycock, L. J. Gray, A. A. Scaife, N. J. Dunstone, J. W. Harder, J. R. Knight, M. Lockwood, J. C. Manners, and R. A. Wood (2015), Regional climate impacts of a possible future grand solar minimum, *Nat. Commun.*, *6*, 7535, doi:10.1038/ncomms8535.
- King, J. H., and N. E. Papitashvili (2005), Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data, *J. Geophys. Res.*, *110*, A02104, doi:10.1029/2004JA010649.
- Leussu, R., I. G. Usoskin, R. Arlt, and K. Mursula (2013), Inconsistency of the Wolf sunspot number series around 1848, *Astron. Astrophys.*, *559*, A28, doi:10.1051/0004-6361/201322373.
- Lockwood, M. (2012), Solar influence on global and regional climates, *Surv. Geophys.*, *33*(3–4), 503–534, doi:10.1007/s10712-012-9181-3.
- Lockwood, M. (2013), Reconstruction and prediction of variations in the open solar magnetic flux and interplanetary conditions, *Liv. Rev. Sol. Phys.*, *10*, 4, doi:10.12942/lrsp-2013-4.
- 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 (2014), Centennial variations in sunspot number, open solar flux and streamer belt width: 3. Modeling, *J. Geophys. Res. Space Physics*, *119*, 5193–5209, doi:10.1002/2014JA019973.
- Lockwood, M., R. Stamper, and M. N. Wild (1999), A doubling of the Sun's coronal magnetic field during the past 100 years, *Nature*, *399*, 437–439, doi:10.1038/20867.
- Lockwood, M., A. Rouillard, I. Finch, and R. Stamper (2006), Comment on "The IDV index: Its derivation and use in inferring long-term variations of the interplanetary magnetic field strength" by Leif Svalgaard and Edward W. Cliver, *J. Geophys. Res.*, *111*, A09109, doi:10.1029/2006JA011640.
- Lockwood, M., A. Rouillard, and I. Finch (2009a), The rise and fall of open solar flux during the current grand solar maximum, *Astrophys. J.*, *700*(2), 937, doi:10.1088/0004-637X/700/2/937.
- Lockwood, M., A. P. Rouillard, and I. D. Finch (2009b), The rise and fall of open solar flux during the current grand solar maximum, *Astrophys. J.*, *700*, 937–944, doi:10.1088/0004-637X/700/2/937.
- Lockwood, M., M. J. Owens, L. Barnard, C. J. Davis, and S. Thomas (2012), Solar cycle 24: What is the Sun up to?, *Astron. Geophys.*, *53*(3), 3.09–03.15, doi:10.1111/j.1468-4004.2012.53309.x.
- 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., H. Nevanlinna, L. Barnard, M. J. Owens, R. G. Harrison, A. P. Rouillard, and C. J. Scott (2014a), 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, doi:10.5194/angeo-32-383-2014.
- Lockwood, M., H. Nevanlinna, M. Vokhmyanin, D. Ponyavin, S. Sokolov, L. Barnard, M. Owens, R. Harrison, A. Rouillard, and C. Scott (2014b), Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr—Part 3: Improved representation of solar cycle 11, *Ann. Geophys.*, *32*(4), 367–381, doi:10.5194/angeo-32-367-2014.
- Lockwood, M., M. J. Owens, and L. Barnard (2014c), 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*, *119*, 5172–5182, doi:10.1002/2014JA019970.
- Lockwood, M., M. J. Owens, L. A. Barnard, and I. G. Usoskin (2016a), An assessment of sunspot number data composites over 1845–2014, *Astrophys. J.*, doi:10.3847/0004-637X824/1/54, in press.
- Lockwood, M., C. Scott, M. Owens, L. Barnard, and D. Willis (2016b), Tests of sunspot number sequences: 1. Using ionosonde data, *Sol. Phys.*, *1–25*, doi:10.1007/s11207-016-0855-8.
- Lynch, B. J., J. R. Gruesbeck, T. H. Zurbuchen, and S. K. Antiochos (2005), Solar cycle dependent helicity transport by magnetic clouds, *J. Geophys. Res.*, *110*, A08107, doi:10.1029/2005JA011137.
- Mayaud, P.-N. (1980), *Derivation, Meaning, and Use of Geomagnetic Indices*, *Geophys. Monogr. Ser.*, vol. 22, 154 pp., AGU, Washington D. C.
- Mursula, K., and D. Martini (2006), Centennial increase in geomagnetic activity: Latitudinal differences and global estimates, *J. Geophys. Res.*, *111*, A08209, doi:10.1029/2005JA011549.
- Mursula, K., D. Martini, and A. Karinen (2004), Did open solar magnetic field increase during the last 100 years? A reanalysis of geomagnetic activity, *Sol. Phys.*, *224*(1–2), 85–94, doi:10.1007/s11207-005-4981-y.
- Neugebauer, M., and C. W. Snyder (1962), Solar plasma experiment, *Science*, *138*(3545), 1095–1097, doi:10.1126/science.138.3545.1095-a.
- Owens, M. J. (2008), Combining remote and in situ observations of coronal mass ejections to better constrain magnetic cloud reconstruction, *J. Geophys. Res.*, *113*, A12102, doi:10.1029/2008JA013589.
- Owens, M. J., 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.
- 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 R. J. Forsyth (2013), The heliospheric magnetic field, *Liv. Rev. Sol. Phys.*, *10*, 5, doi:10.12942/lrsp-2013-5.
- Owens, M. J., N. U. Crooker, and M. Lockwood (2011a), How is open solar magnetic flux lost over the solar cycle?, *J. Geophys. Res.*, *116*, A04111, doi:10.1029/2010JA016039.
- Owens, M. J., M. Lockwood, C. J. Davis, and L. Barnard (2011b), Solar cycle 24: Implications for energetic particles and long-term space climate change, *Geophys. Res. Lett.*, *38*, L19106, doi:10.1029/2011GL049328.
- Owens, M. J., E. Cliver, K. G. McCracken, J. Beer, L. Barnard, M. Lockwood, A. Rouillard, D. Passos, P. Riley, I. Usoskin and Y.-M. Wang (2016), Near-Earth Heliospheric Magnetic Field Intensity Since 1750. Part 2: Cosmogenic Radionuclide Reconstructions, *J. Geophys. Res. Space Physics*, doi:10.1002/2016JA022550.
- 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.
- Russell, C. (1975), On the possibility of deducing interplanetary and solar parameters from geomagnetic records, *Sol. Phys.*, *42*(1), 259–269, doi:10.1007/BF00153301.
- Schwadron, N. A., D. E. Connick, and C. Smith (2010), Magnetic flux balance in the heliosphere, *Astrophys. J. Lett.*, *722*, L132–L136, doi:10.1088/2041-8205/722/2/L132.
- Sheeley, N., Jr., T. Knudson, and Y.-M. Wang (2001), Coronal inflows and the Sun's nonaxisymmetric open flux, *Astrophys. J. Lett.*, *546*(2), L131, doi:10.1086/318873.
- Solanki, S. K., M. Schüssler, and M. Fligge (2000), Evolution of the Sun's large-scale magnetic field since the Maunder minimum, *Nature*, *408*, 445–447, doi:10.1038/35044027.
- Solanki, S. K., M. Schüssler, and M. Fligge (2002), Secular variation of the Sun's magnetic flux, *Astron. Astrophys.*, *383*(2), 706–712, doi:10.1051/0004-6361:20011790.
- Stamper, R., M. Lockwood, M. Wild, and T. Clark (1999), Solar causes of the long-term increase in geomagnetic activity, *J. Geophys. Res.*, *104*(A12), 28,325–28,342, doi:10.1029/1999JA900311.
- Svalgaard, L. (1977), Geomagnetic activity: Dependence on solar wind parameters, in *SkyLab Workshop on Coronal Holes*, edited by J. B. Zirker, pp. 371–441, Colo. Assoc. Univ. Press, Boulder, Colo.
- Svalgaard, L. (2010), Updating the historical sunspot record, in *SOHO-23: Understanding a Peculiar Solar Minimum*, edited by R. Cranmer, J. T. Hoeksema, and J. L. Kohl, pp. 297, Astron. Soc. Pac., San Francisco.
- Svalgaard, L. (2012), How well do we know the sunspot number? in *Proc. IAU Symp.*, edited by C. H. Mandrini and D. F. Webb, pp. 27–33, Cambridge Univ. Press, Mendoza, Argentina.

- Svalgaard, L. (2013), Solar activity—Past, present, future, *J. Space Weather Space Clim.*, 3, A24, doi:10.1051/swsc/2013046.
- Svalgaard, L. (2014), Errors in Scale Values for Magnetic Elements for Helsinki, *Ann. Geophys.*, 32, 633, doi:10.5194/angeo-32-633-2014.
- Svalgaard, L., and E. Cliver (2006), Reply to the comment by M. Lockwood et al. on "The IDV index: Its derivation and use in inferring long-term variations of the interplanetary magnetic field", *J. Geophys. Res.*, 111, A09110, doi:10.1029/2006JA011678.
- Svalgaard, L., and E. W. Cliver (2005), The IDV index: Its derivation and use in inferring long-term variations of the interplanetary magnetic field strength, *J. Geophys. Res.*, 110, A12103, doi:10.1029/2005JA011203.
- Svalgaard, L., and E. W. Cliver (2007a), A floor in the solar wind magnetic field, *Astrophys. J. Lett.*, 661, L203–L206, doi:10.1086/518786.
- Svalgaard, L., and E. W. Cliver (2007b), 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.
- Svalgaard, L., and E. W. Cliver (2010), Heliospheric magnetic field 1835–2009, *J. Geophys. Res.*, 115, A09111, doi:10.1029/2009JA015069.
- Svalgaard, L., and K. H. Schatten (2016), Reconstruction of the sunspot group number: The backbone method, *Sol. Phys.*, 1–32, doi:10.1007/s11207-015-0815-8.
- Svalgaard, L., E. W. Cliver, and P. LeSager (2003), Determination of interplanetary magnetic field strength, solar wind speed and EUV irradiance, 1890–2003, in *Proceedings of ISCS 2003 Symposium: Solar Variability as an Input to the Earth's Environment*, edited by A. Wilson, pp. 15–23, Eur. Space Agency, Paris, Tatranska Lomnica, Slovakia.
- Svalgaard, L., E. W. Cliver, and P. Le Sager (2004), IHV: A new long-term geomagnetic index, *Adv. Space Res.*, 34(2), 436–439, doi:10.1016/j.asr.2003.01.029.
- Tapping, K., D. Boteler, P. Charbonneau, A. Crouch, A. Manson, and H. Paquette (2007), Solar magnetic activity and total irradiance since the Maunder Minimum, *Sol. Phys.*, 246(2), 309–326, doi:10.1007/s11207-007-9047-x.
- Usoskin, I. G. (2013), A history of solar activity over millennia, *Liv. Rev. Sol. Phys.*, 10(3), doi:10.12942/lrsp-2013-1.
- Usoskin, I., et al. (2015), The Maunder minimum (1645–1715) was indeed a Grand minimum: A reassessment of multiple datasets, *Astron. Astrophys.*, 581, A95, doi:10.1051/0004-6361/201526652.
- Usoskin, I., G. Kovaltsov, M. Lockwood, K. Mursula, M. Owens, and S. Solanki (2016), A new calibrated sunspot group series since 1749: statistics of active day fractions, *Sol. Phys.*, 1–24, doi:10.1007/s11207-015-0838-1.
- Vaquero, J., G. Kovaltsov, I. Usoskin, V. Carrasco, and M. Gallego (2015), Level and length of cyclic solar activity during the Maunder minimum as deduced from the active-day statistics, *Astron. Astrophys.*, 577, A71, doi:10.1051/0004-6361/201525962.
- Vieira, L. E. A., and S. K. Solanki (2010), Evolution of the solar magnetic flux on time scales of years to millennia, *Astron. Astrophys.*, 509, A100, doi:10.1051/0004-6361/200913276.
- Vieira, L. E. A., S. K. Solanki, N. A. Krivova, and I. Usoskin (2011), Evolution of the solar irradiance during the Holocene, *Astron. Astrophys.*, 531, A6, doi:10.1051/0004-6361/201015843.
- Wang, Y.-M., J. Lean, and N. Sheeley Jr. (2005), Modeling the sun's magnetic field and irradiance since 1713, *Astrophys. J.*, 625(1), 522, doi:10.1086/429689.
- Yashiro, S., N. Gopalswamy, G. Michalek, O. C. S. Cyr, S. P. Plunkett, N. B. Rich, and R. A. Howard (2004), A catalog of white light coronal mass ejections observed by the SOHO spacecraft, *J. Geophys. Res.*, 109, A07105, doi:10.1029/2003JA010282.
- Yeo, K., N. Krivova, and S. Solanki (2014), Solar cycle variation in solar irradiance, *Space Sci. Rev.*, 186(1–4), 137–167, doi:10.1007/s11214-014-0061-7.
- York, D., N. M. Evensen, M. Lo, M. Nez, J. S. De, and B. Delgado (2004), Unified equations for the slope, intercept, and standard errors of the best straight line, *Am. J. Phys.*, 73, doi:10.1119/1.1632486.