JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 111, A09109, doi:10.1029/2006JA011640, 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

M. Lockwood,[1,2] A. P. Rouillard,[2] I. Finch,[1] and R. Stamper[1]

Received 30 January 2006; revised 2 April 2006; accepted 22 May 2006; published 21 September 2006.

**Citation:** Lockwood, M., A. P. 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.

## 1. Introduction

[1] *Svalgaard and Cliver* [2005] (hereinafter referred to as SC05) use the *IDV* geomagnetic index to infer the variation of the interplanetary magnetic field strength, $B$, since 1872. They find that "$B$ increased by ~25% from the 1900s to the 1950s" (paragraphs 1 and 24) and that this "is in contrast to the more than doubling of $B$ during the 20th century obtained from an analysis of the *aa* index by *Lockwood et al.* [1999]" (paragraph 24). We agree with neither statement. We identify a number of errors and biases in the analysis of SC05, each of which acts in the same direction, namely to reduce the true long-term drift. The key result of *Lockwood et al.* (hereinafter referred to as LEA99) is a doubling, not in $B$, but rather in the total coronal source flux which is proportional to the radial IMF component, $B_r$. This is an important distinction which we discus here in section 3. We show (in section 2) that SC05's own analysis gives a drift in decadal averages in $B$ of 38%, not the "~25%" that they quote, but we also point out (section 4) that SC05's method for dealing with data gaps gives rise to a (small) bias and that their regression procedure (which is compared to others in section 5) is not robust (sections 6 and 7), in addition to them not taking into account the distinction between $B$ and $B_r$. We here argue that a simple ordinary linear regression procedure, as used by SC05, is inferior to the method employed by LEA99 (section 9) but does nevertheless show that the *IDV* index is fully consistent with the doubling in the open solar flux found by LEA99 (section 8).

## 2. Percentage Change

[2] As used by both LEA99 and SC05, we here use 11-year running means on all IMF parameters in order to remove the solar cycle variation and reveal the underlying long-term drift. All IMF parameters that have been scaled from *IDV* and smoothed in this way will show a peak in 1956 and a minimum in 1903, as does *IDV* itself. Applying an 11-year running mean to the values given in Table 3 of SC05, we find that their regression gives a minimum of $[B]_{03} = 5.51$ nT and a maximum of $[B]_{56} = 7.58$ nT. The percentage change between 1903 and 1956 is

$$\lambda = 100\left([B]_{56} - [B]_{03}\right)/[B]_{03} \quad (1)$$

Inputting the above values of $[B]_{56}$ and $[B]_{03}$ into equation (1) yields a change of $\lambda = 38\%$.

[3] We note two points about the percentage change $\lambda$: (1) it is a sensitive function of the slope $m$ of the regression fit because an increase in $m$ increases the numerator in equation (1) but also reduces the denominator and (2) the 11-year averaging interval used to generate $\lambda$ is appropriate as it is the shortest that removes the solar cycle variation. Given that there is no doubt that the open solar flux has fallen again in magnitude in recent decades [*Lockwood*, 2003], averaging over a longer period will cause smoothing and reduce the amplitude of variations in open solar flux on timescales greater than 11 years. Hence averaging over more than 11 years would not be a test of SC05's contention that "$B$ increased by ~25% from the 1900s to the 1950s" (paragraph 1).

## 3. Radial IMF, $B_r$, and the IMF Magnitude, $B$

[4] The heliocentric radial IMF component $B_r$ is the key parameter in the analysis of LEA99 because they were interested in the (signed) open solar flux, $F_s$. (We here define $B_r = |B_X|$, where the X direction is toward the Sun in the Geocentric Solar Ecliptic frame of reference.) Using the Ulysses result that the radial field is independent of latitude and averaging over periods of more than 27 days to remove longitudinal structure, $F_s$ is related to the radial IMF at heliocentric distance $R_1 = 1$AU by

$$F_s = 4\pi R_1^2 B_r/2 = 4\pi R_1^2 |\langle B_X \rangle_T|/2 \quad (2)$$

The factor 2 arises from the fact that half the unsigned flux is inward and half is outward (i.e., assuming there is no

[1]Rutherford Appleton Laboratory, Chilton, Oxfordshire, UK.
[2]Solar-Terrestrial Physics Group, Department of Physics and Astronomy, University of Southampton, Southampton, UK.


0148-0227/06/2006JA011640$09.00

imbalance of any magnetic monopoles inside the heliocentric sphere of radius $R_1$). *Lockwood et al.* [2004] have used the two perihelion passes of Ulysses to show that using equation (2) gives values of $F_s$ that are accurate to within 5% on timescales greater than 27 days.

[5] However, the use of equation (2) raises the question of what timescale, $T$, should $B_X$ be averaged over, before the absolute value is taken, in order to best represent the solar source field. The original paper by LEA99 employed $T$ = 1 hour for which $B_r = |\langle B_X \rangle_T|$ is close to being proportional to $B$. Thus for $T$ = 1 hour, the percentage changes in $B$ and $B_r$ are almost the same and LEA99 carried out their analysis using $B_r = 0.560B$. However, subsequently, *Lockwood* [2002] pointed out the importance timescale in this relationship. If a low value of $T$ is used, variations in $B_r$ due to magnetic islands, interplanetary coronal mass ejections (ICMEs), stream-stream interactions, and Alfvén waves will all be included. However, if a larger $T$ is used, opposite polarity $B_X$ within the averaging interval (caused by these heliospheric effects rather than the polarity of the coronal source field) cancels out. On the other hand, if too large a $T$ value is used, opposite polarity $B_X$ due to genuine solar sector structure in the coronal source field will be cancelled and the true open solar flux will be underestimated. Note that the value of $B$ is not similarly influenced by $T$ because, unlike $B_X$, it is always positive. Thus $B$ will always contain disconnected magnetic flux and the effect of small-scale heliospheric structures and waves whereas $B_r$, averaged over a suitable $T$, will not. The differing effects of $T$ on $B_r$ and $B$ are illustrated by Figure 1. The heavy solid line shows the variation in annual means of $B$ (for any $T$) whereas the thin solid lines gives annual means of $B_r = |\langle B_X \rangle_T|$, which have been first averaged and then made into an absolute value on a timescale $T$. From top to bottom, the $T$ used is: 1 hour, 1 day, 2 days, 3 days, and 5 days. It can be seen that as expected from the above discussion, increasing $T$ causes the average $B_r$ values to fall but does not influence $B$.

[6] The shaded area in Figure 1 gives the radial field at Earth, $[B_r]_{PFSS}$, derived using the potential field source surface (PFSS) method of *Wang and Sheeley* [1995] from photospheric magnetogram data. The agreement between the average levels of $[B_r]_{PFSS}$ and $B_r$ is best for $T$ of 1 or 2 days. The PFSS modeling not only matches day-scale averages of the radial field at 1AU, but also a wide array of other observed heliospheric and coronal features and their solar cycle variations, such as: the coronal hole configuration from He I 10830Å images [*Wang et al.*, 1996]; the solar wind speed [*Arge and Pizzo*, 2000; *Whang et al.*, 2005] and stream-stream interaction regions [*Wang et al.*, 1997; A. P. Rouillard and M. Lockwood, The butterfly diagram of kinematic steepening of stream-stream interactions: comparing predictions and observations, submitted to *Astronomy and Astrophysics*, 2006] seen by various spacecraft in the heliosphere; and the location of the heliospheric current sheet and coronal streamers seen in white light coronagraph images [*Wang et al.*, 2000a, 2000b]. We here adopt an optimum $T$ of 1 day. This eliminates all but the largest ICMEs and magnetic islands while minimizing the risk of averaging out genuine, short-lived solar sector structure. (We gain an idea of the upper limit to the allowed range of $T$ by considering that typically, the Sun has four sectors which, on average, would take 6.5 days to pass over the Earth.) We note that $T$ = 1 day is often adopted in this context (see, for example, http://omniweb.gsfc.nasa.gov/html/polarity/polarity.html).

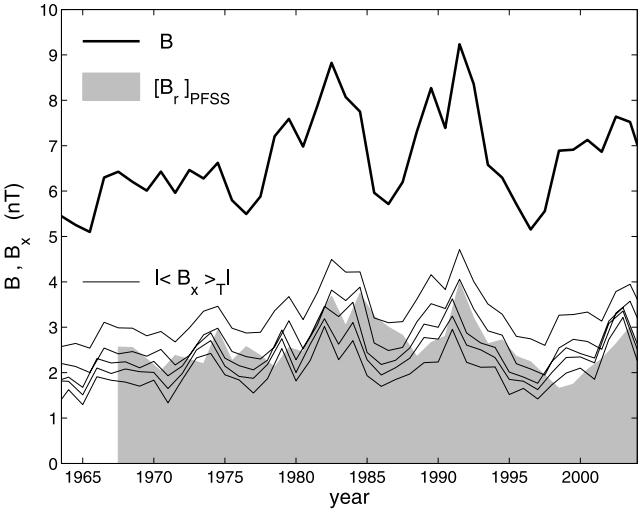


**Figure 1.** Variation of annual means of IMF parameters from the Omnitape data set. The thick solid line is the IMF field strength, $B$. The thin lines are the absolute values of the radial IMF component $B_r = |\langle B_X \rangle_T|$ for various timescales $T$ on which the data are averaged and made into absolute values before being averaged again into annual means: from top to bottom, $T$ is 1 hour, 1 day, 2 days, 3 days, and 5 days. The shaded area shows $[B_r]_{PFSS}$, the radial IMF at Earth derived from ground-based photospheric magnetograms, using the PFSS (potential field source surface) procedure, as implemented by *Wang and Sheeley* [1995].

[7] Figure 2 is a scatterplot of annual means of $B_r$, for $T$ of 1day, against $B$. The correlation coefficient is 0.90 (significance, S = 99.95%) and the best fit regression (using the residual minimization method D described in the following sections) is

$$B_r = 0.548\ B - 0.91 \tag{3}$$

Note that the residual $B$ of 1.68 nT when $B_r = 0$ is physically possible and would be disconnected flux (ICMEs, magnetic islands etc); however, it is also probable that the relationship doesn't remain completely linear down to the lowest $B_r$. Using equation (3), the values of $[B]_{03}$ and $[B]_{56}$ from SC05's own fit gives $[B_r]_{03}$ = 2.11 nT and $[B_r]_{56}$ = 3.24 nT which, by equation (1), gives a percentage change in $B_r$ (and $F_s$) of $\lambda$ = 54%.

[8] Hence SC05's own regression fit actually gives a percent change in decadal-scale means of 38% for $B$ and 54% for the open solar flux, as opposed to the ~25% change that they quote. However, SC05's regression procedure is not robust. In sections 5–7 we look at linear regressions between *IDV* and both $B$ and $B_r$, after first considering the effect of data gaps in the IMF observations in section 4.

## 4. Effect of Missing IMF Data

[9] Since the advent of the ACE spacecraft, the observations of the IMF have been almost continuous, with more

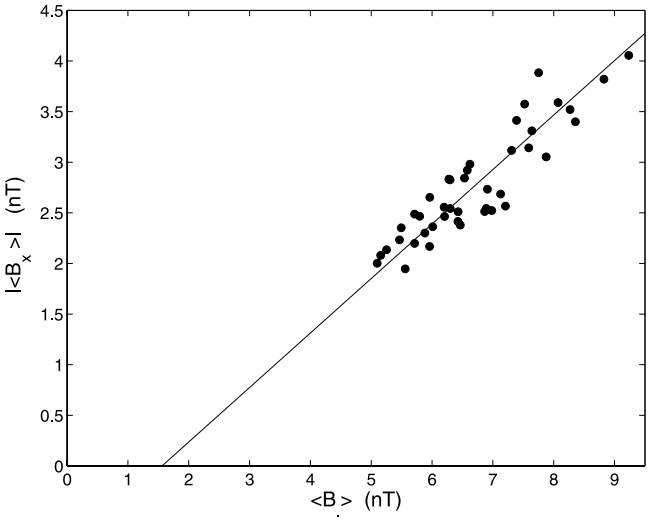


**Figure 2.** Scatterplot of annual means of the absolute value of the radial IMF component $B_r = |\langle B_X \rangle_T|$, for an averaging timescale of $T = 1$ day, as a function of the IMF field strength, $B$. The best fit regression line uses the residual minimization method D.

than 95% coverage in hourly values since 1995. Before this date, coverage was sometimes poor. We here include all years except 1963 and 1964 (for which the coverage was only 8% and 24%, respectively). The analysis of LEA99 was carried out in two ways: (1) using all available data to make annual means of both the solar wind parameters and *aa* and (2) by making annual means of *aa* from only for those times when IMF data were also available. The results for the two methods were almost identical.

[10] SC05 attempt to use a recurrence-based method to fill in the data gaps in the IMF data series. We do not believe that this is a valid technique because the autocorrelation function of the IMF at 27 days is small. Specifically, for daily means ($T = 1$day) the autocorrelation function of $B$ at 27 days is just 0.20 and it is even worse for $B_r$, being just 0.15. Hence filling in the data gaps this way will lead to errors, especially in the parameter that is relevant to the open solar flux, namely $B_r$. There is, anyway, no need to do this. We here show one can avoid this interpolation by quantifying the effect of removing the missing data on geomagnetic observations and then making allowance for that effect.

[11] Data gaps are a problem in annual mean data only if they introduce a systematic bias into the annual averages by removing data in a way that is not random. There are some intervals in the data series when the data gaps were repeatedly at certain UT which would introduce systematic effects through the variation of Earth's dipole tilt; however, mercifully, such intervals were relatively short in duration. Since 1995, data gaps have been rare but are decidedly nonrandom as they have tended to be caused by major solar-terrestrial storms. Figure 3 shows distributions of the values of the 3-hourly *am* index (which is a range-based index compiled in the same way as the *aa* index, but from 21 magnetometer stations at all longitudes and in both hemispheres) for 1965–2004. The shaded histogram is for all *am* data. The black histogram is only the data points for which all three corresponding hourly means (as given in the Omni data set) of IMF data are available, allowing for the propagation delay between the observing spacecraft and the dayside magnetosphere. The black line is the black histogram, scaled so that it has the same total number of samples as the shaded histogram. The key point is that the shapes of the two distributions are nearly identical, showing that overall the missing data introduce a very small systematic bias into the full data set.

[12] However, that there is no bias over the entire period of IMF observations does not mean that biases are not present in the annual means. Figure 4 (left) shows a scatterplot of the annual mean of the subset of *am* data for which all three hourly means of coincident IMF data are available, $\langle am \rangle_W$, as a function of the mean of all data points, $\langle am \rangle$. The dashed line is $\langle am \rangle_W = \langle am \rangle$ which would apply for 100% availability of IMF data. The solid line is the best fit regression, given by $\langle am \rangle_W = 0.927\langle am \rangle + 1.590$. The close agreement of $\langle am \rangle_W$ and $\langle am \rangle$ confirms that the data gaps are close to being random in effect, even on annual timescales, and that the geomagnetic activity driven by the unobserved IMF has only a small influence on the annual means. Nevertheless, there is a systematic effect in that when *am* is low/high, the data gaps cause a slight tendency to under/overestimate, respectively, the true value. Other definitions of the IMF availability have been used (for example the fraction of 1 min samples in the appropriate 3-hour period) and make no significant difference to the regression coefficients. Figure 4 (right) gives the

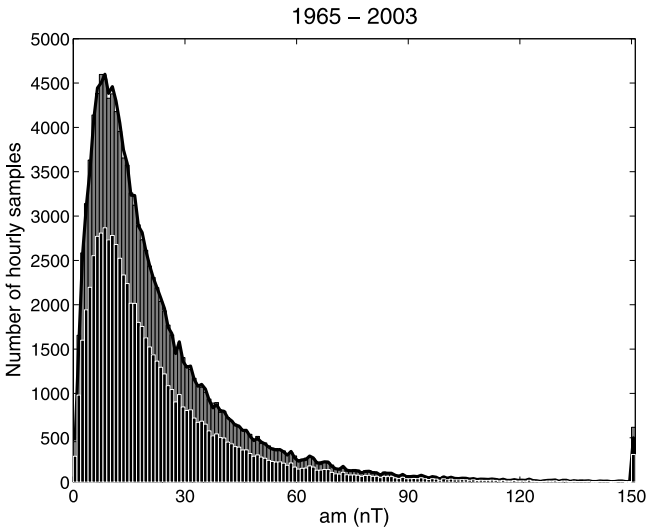


**Figure 3.** Distributions of the geomagnetic *am* index values for 1965–2004. The shaded histogram is for all the 3-hourly *am* data points, and the back histogram is for the 3-hourly *am* data for which all three coincident 1-hour means of IMF data are available in the Omni data set. The uppermost bin is for all samples giving $am > 125$. The black line is the black histogram, scaled so that it has the same total number of samples as the shaded histogram. The two distributions are almost identical in shape showing that data gaps are random and do not add systematic changes on the timescale of several decades.

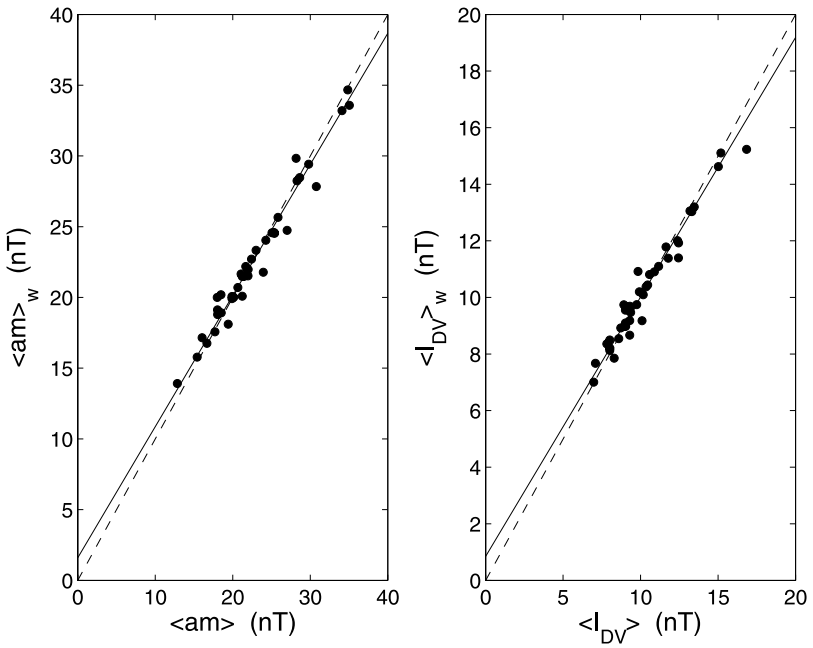


**Figure 4.** Analysis of the effect of the missing IMF data on annual timescales. (left) Annual means of the *am* geomagnetic index, for data points with coincident IMF data, $\langle am \rangle_W$, shown as a function of the mean for all data in that year $\langle am \rangle$. The solid line is the best fit regression. The dashed line is $\langle am \rangle = \langle am \rangle_W$. (right) Corresponding plot for the *IDV* index.

equivalent plot for the *IDV* index, for which the best fit regression is

$$\langle I_{DV} \rangle_W = 0.890 \langle I_{DV} \rangle + 1.103 \quad (4)$$

The regression analysis presented in section 5 has been carried out using both $\langle I_{DV} \rangle_W$ and $\langle I_{DV} \rangle$. Because the data gaps are close to being random in occurrence they introduce only small errors into the annual means and the two give similar results. However, Figure 4 shows that there is a slight effect and we here use $\langle I_{DV} \rangle_W$ to derive the regression coefficients.

## 5. Effect of Regression Technique

[13] SC05 do not make it clear which regression technique they employ. However, using the annual mean data supplied in their Table 3, we have carried out regressions using a variety of techniques. The key point here is which quantity $\Delta$ should be minimized in order to get the best fit regression line.

[14] 1. The quantity $\Delta$ minimized is the sum of the squares value of the deviation $\Delta B$ of the observed $B$ value from the best fit regression at that $I_{DV}$ ($\Delta = \Sigma_i \Delta B_i^2$). For ordinary least squares (OLS) regression, this is the method employed to predict $B$ or $B_r$ from $I_{DV}$.

[15] 2. The quantity $\Delta$ minimized is the sum of the squares value of the deviation $\Delta I_{DV}$ of observed $I_{DV}$ value from the best regression fit at that $B$ ($\Delta = \Sigma_i \Delta I_{DVi}^2$). In OLS, this is the method employed to predict $I_{DV}$ from $B$ or $B_r$. (In fact, in a subsequent publication (M. Lockwood et al., How large was the rise in open solar flux during the 20th century?—Application of Bayesian statistics, submitted to *Annales Geophysicae*, 2006, hereinafter referred to as Lockwood et al., submitted manuscript, 2006) we will show that this is also the correct minimization for the prediction of $B$ and $B_r$ from $I_{DV}$ using Bayesian statistics, with the prior knowledge that $I_{DV}$ is an approximate and indirect measurement of $B$ or $B_r$.)

[16] 3. In some statistical applications, one wishes to minimize the RMS length of the perpendiculars between each data point and the best fit regression line after both parameters have been put on the same scale (effectively, $\Delta = \Sigma_i(\Delta I^*_{DVi} + \Delta B_i^2)$, where $I^*_{DV} = m\, I_{DV} + c$). Although this major axis analysis (MAA) [*Sokal and Rohlf*, 1995] is often a valid technique, it is not the best method when we are making predictions of one parameter based on another and the scatter is arising from experimental measurement uncertainties.

[17] 4. The above regression methods are all subject to the "leverage" effect of outlying data points. Because median values are less influenced by outliers than mean values, least median of squares (LMS) regression [*Rousseeuw*, 1984; *Rousseeuw and Leroy*, 1987] is a more robust procedure that is not as greatly influenced by outliers. In this case, one minimizes the square of the deviation from median values, i.e., $\Delta = [\Delta B_i^2]_{med}$ when predicting $B$ or $B_r$ from $I_{DV}$ and $\Delta = [\Delta I_{DVi}^2]_{med}$ when predicting $I_{DV}$ from $B$ or $B_r$.

[18] At this point, it should be remembered what SC05 are attempting to do: they are fitting a geomagnetic index to interplanetary parameters for after 1965, when the space measurements became available, and then using the regression fit obtained with the longer geomagnetic data series to extrapolate back to earlier times. We have observations of the IMF which are not continuous but which, as shown in section 4, we can correct for the resulting small systematic errors caused by the data gaps. We have observations of

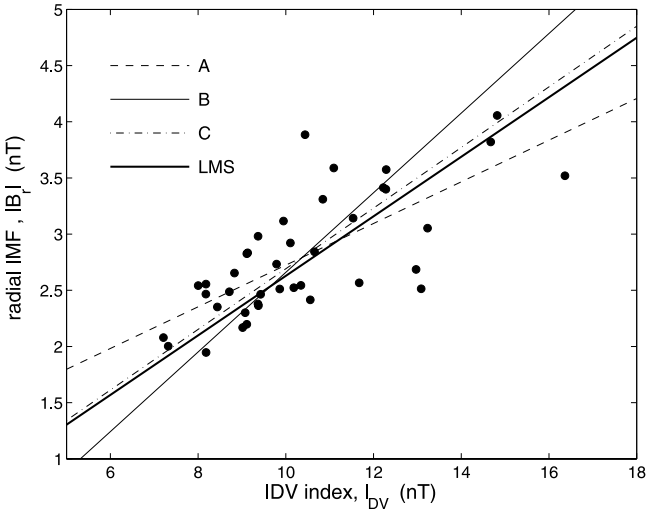


**Figure 5.** Scatterplot of annual means the radial IMF component, $B_r$, as a function of the *IDV* geomagnetic index, $I_{DV}$, demonstrating the effect of the choice of regression method used. The dashed, thin solid, dot-dashed, and thick solid lines are the best fit linear regressions for residual minimization methods A (OLS), B (required for this case by Bayesian statistics), C (MAA), and D (LMS), respectively. See section 5 for details.

geomagnetic activity, which is controlled by a number of factors, including the north-south ($B_z$) component of the IMF in the Geocentric Solar Magnetospheric (GSM) frame of reference. Thus geomagnetic activity can be no more than an approximate estimator of either the IMF strength $B$ or its radial component $B_r$, by far the dominant source of scatter in a regression (between either IMF parameter and a geomagnetic activity index) is the fact that the latter is only an approximate indicator of the former. (Were this not the case, there would be no need for expensive in situ IMF monitors on spacecraft because ground-based magnetometers would do a better job.) Thus there is an error in $B$ for a given *IDV* estimate, but only because there is an error in that *IDV* as an estimator of the IMF. For a genuine linear relationship with a normal distribution of errors, the error in $B$ for a given *IDV* would be normally distributed ($\Delta B_i = m\Delta I_{DVi}$ where $m$ is the regression slope) and there would be no problem with method A (i.e., the use of OLS to predict $B$ or $B_r$): however, if these assumptions are not valid, this causes additional problems. Note that method B avoids this by minimizing the errors in *IDV* directly and, for OLS regression, is used to predict *IDV* for a given $B$ or $B_r$ (see section 8).

[19] Figure 5 is a scatterplot of annual means the radial IMF field strength (for averaging timescale $T$ = 1 day), $B_r$, as a function of the *IDV* geomagnetic index and shows the best fit regressions obtained by the above methods. Figure 5 illustrates how sensitive the regression fit is to the minimization method used for this case. Table 1 also gives the results of regressions using these four methods. Regressions between $I_{DV}$ and $B$ and regressions between $I_{DV}$ and $B_r$ are given. Table 1 also gives the values of $[B]_{56}$, $[B_r]_{56}$, $[B]_{03}$ and $[B_r]_{03}$ derived from the regression, along with the percentage changes $\lambda$ from equation (1).

[20] Table 1 also gives the results from the regression fit given by SC05. The close similarity with our results makes it clear that they have used method A (i.e., ordinary least squares, OLS). The small difference arises from SC05's introduction of extrapolated data in the IMF data gaps, whereas we have removed the *IDV* data without simultaneous IMF data, as discussed above.

## 6. Analysis of Fit Residuals

[21] It is important to analyze any regression fit to test that the residuals are random and show no systematic behavior (either in time or with one or both of the fit parameters). For extrapolation of the IMF, the key test of the regression is how well the fitted IMF parameter, derived from the geomagnetic index, reproduces the observed IMF parameter. Figure 6 shows such an analysis for annual means of IMF strength $B$ from the fit of SC05. The difference ($B_{obs}$-$B_{fit}$) between the observed value, $B_{obs}$, and the fitted value, $B_{fit}$, is shown as a function of $B_{fit}$ in Figure 6 (left) and $B_{obs}$ in Figure 6 (right). The dashed lines are linear regression fits to the data points shown, using the

**Table 1.** Results of Correlation and Linear Regression Analyses of the *IDV* Index, $I_{DV}$, and IMF $B$ and $B_r$ for Different Residual Minimization Criteria[a]

| Parameters | Minimization Method | Minimization Δ | $r$ | $S$, % | $m$ | $c$ | $[B]_{56}$, $[B_r]_{56}$, nT | $[B]_{03}$, $[B_r]_{03}$, nT | λ, % |
|---|---|---|---|---|---|---|---|---|---|
| $I_{DV}$ & $B$ | ? | SC05 | 0.87 | ? | 0.361 | 3.04 | 7.58 | 5.51 | 38 |
| $I_{DV}$ & $B$ | A | $\Sigma_i \Delta B_i^2$ | 0.87 | 99.96 | 0.411 | 2.51 | 7.68 | 5.32 | 44 |
| $I_{DV}$ & $B$ | D | $[\Delta B_i^2]_{med}$ (LMS) | 0.87 | 99.96 | 0.448 | 2.04 | 7.67 | 5.10 | 51 |
| $I_{DV}$ & $B_r$ | A | $\Sigma_i \Delta B_{ri}^2$ | 0.72 | 99.21 | 0.185 | 0.87 | 3.20 | 2.14 | 50 |
| $I_{DV}$ & $B_r$ | B | $\Sigma_i \Delta I_{DVi}^2$ | 0.72 | 99.21 | 0.354 | −0.88 | 3.57 | 1.54 | 132 |
| $I_{DV}$ & $B_r$ | C | $\Sigma_i(\Delta I^{*2}_{DVi} + \Delta B_{ri}^2)$ | 0.72 | 99.21 | 0.270 | −0.01 | 3.39 | 1.84 | 84 |
| $I_{DV}$ & $B_r$ | D | $[\Delta B_{ri}^2]_{med}$ (LMS) | 0.72 | 99.21 | 0.265 | −0.02 | 3.31 | 1.79 | 85 |

[a]Here $r$ is the correlation coefficient and $S$ is its significance level; $m$ and $c$ are the slope and intercepts of the linear regression, ($B$ or $B_r$) = $m$ $I_{DV}$ + $c$; $[B]_{56}$ and $[B_r]_{56}$ are the peak values of the 11-year running means of the fitted $B$ and $B_r$ variations (for 1956); $[B]_{03}$ and $[B_r]_{03}$ are the minimum values of the 11-year running means of the fitted $B$ and $B_r$ variations (for 1903); λ is the percent change between these dates. A question mark indicates unknown. Note that the significance levels of correlations quoted here all make correction for the persistence of the data (from their autocorrelation functions) and hence the effective number of independent samples [*Wilkes*, 1995]. This correction was not made in the original paper by LEA99 who, as a result, quoted significance values that were too high. However, even with the correction for persistence, all the correlations presented by LEA99 remain significant at greater than the 99.9% level.

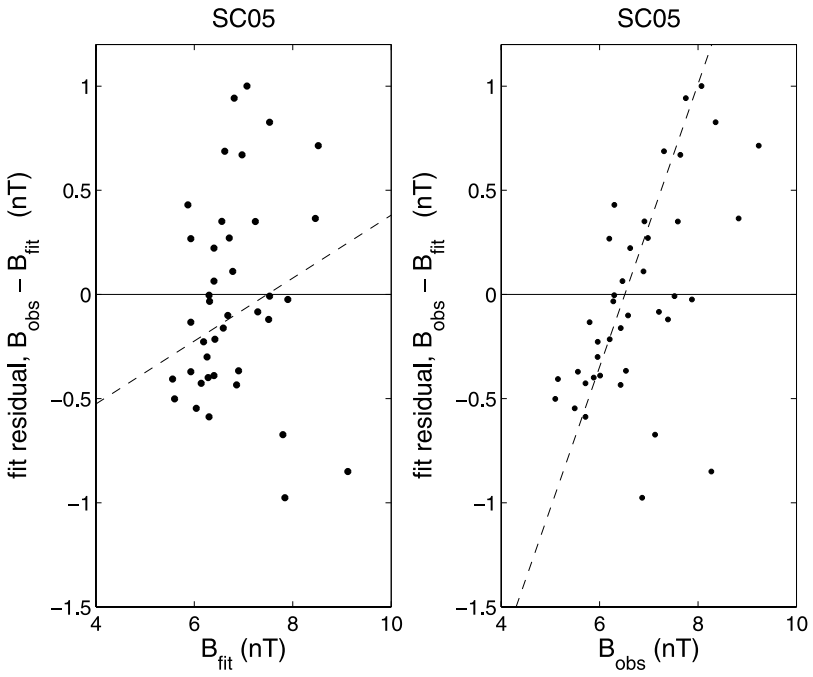


**Figure 6.** Analysis of fit residuals for the fit given by SC05. The fit residuals ($B_{obs}$-$B_{fit}$) are shown as a function of (left) $B_{fit}$ and (right) $B_{obs}$, where $B_{obs}$ is the observed value and $B_{fit}$ is the fitted value from the *IDV* index. The dashed line is the LMS linear regression fit to the points shown. Figure 6 (left) shows that the fit by SC05 does not display homoskedacity (i.e., the variance increases markedly with $B_{fit}$) and is only prevented from showing an even stronger systematic bias by any one of three outliers with excessively large $B_{fit}$ (i.e., large *IDV*). Figures 6 (left) and 6 (right) reveal that $B_{fit}$ is consistently an overestimate of $B_{obs}$ when the $B_{obs}$ or $B_{fit}$ are small and consistently an underestimate of $B_{obs}$ when the $B_{obs}$ or $B_{fit}$ is large. The slope of the LMS regressions in Figures 6 (left) and 6 (right) are $s = 0.151$, and $s = 0.679$, respectively, and the corresponding intercepts are $c = -4.42$ nT and $c = -1.13$ nT.

LMS minimization to reduce the effect of outliers. Figure 6 (left) shows that the fit of SC05 is far from obeying "homoskedacity." Homoskedacity means that the variance is constant and is assumed by least squares regression: The departure from this condition can be seen in Figure 6 (left) as an increase in the variance of the residuals as $B_{fit}$ increases. Figure 6 (left) also reveals three major outliers with large $B_{fit}$ and large and negative ($B_{obs}$-$B_{fit}$); that is, in all three cases, $B_{fit}$ (and hence *IDV*) is too large. These outliers are for 1989, 2001, and 1983, and we note that the median interhour variability derived by *Lockwood et al.* [2006] is significantly different from SC05's *IDV* for these years and does not give these outliers. The regression lines in Figure 6 employ the LMS minimization (D) that is more robust and not strongly influenced outliers: the bias in the fit of SC05 is also revealed by any of the regression procedures if one removes just one (any one) of the three largest outliers. We also wish to see if there is any systematic error in $B_{fit}$ as an indirect estimation of $B_{obs}$, before one uses it to extrapolate. These residual plots reveal systematic bias, with $B_{fit}$ consistently exceeding $B_{obs}$ when either $B_{obs}$ or $B_{fit}$ is small and with $B_{fit}$ consistently less than $B_{obs}$ when $B_{obs}$ or $B_{fit}$ is large. Thus the SC05 fit systematically overestimates low $B_{obs}$ values and underestimates high values (and so underestimates the slope $m$ and the percentage change $\lambda$). Note how strong the bias is in SC05's plot: the slope of the LMS regression fit in Figure 6 (right) is $s = 0.679$, and the intercept is $c = -4.423$ nT. If we use this regression to correct SC05's extrapolated values for $[B]_{56}$ and $[B]_{03}$ (as given in Table 1), this yields 9.84 nT and 3.38 nT, respectively, and by equation (1), a $\lambda$ of 192%. We do not suggest that this is the true value of $\lambda$ for $B$; however, this does serve to show the extreme sensitivity of the derived $\lambda$ to uncertainties and/or inadequacies in SC05's regression procedure.

[22] Figure 7 presents another problem with the simple OLS regression technique used by SC05. This is a quantile-quantile (QQ) plot and is used to test that the residuals are normally distributed, as is also assumed by least squares regression. The standardized residuals ($\Delta B_{ri}/\sigma$) are evaluated, where $\sigma^2$ is $\Sigma_i \Delta B_{ri}^2/(n-2)$ and $n$ is the number of samples, and then placed in order by size and plotted against the quantiles for a standard normal distribution. The deviations from the straight line of slope 1 reveal departures from a normal distribution [*Wilkes*, 1995; *van Storch and Zwiers*, 1999]. This is a problem for all the regressions listed in Table 1, pointing to the general unreliability of linear regressions in this context.

## 7. Effect on the Fit of Outliers

[23] It is well known that the slope of a regression fit can be greatly altered by outlying data points on the scatterplot which yield large residuals. The accepted method for analyzing the leverage effect of these outliers is to remove them one by one, until it makes no further difference to the regression fit. Figure 8 shows the various regression fits of

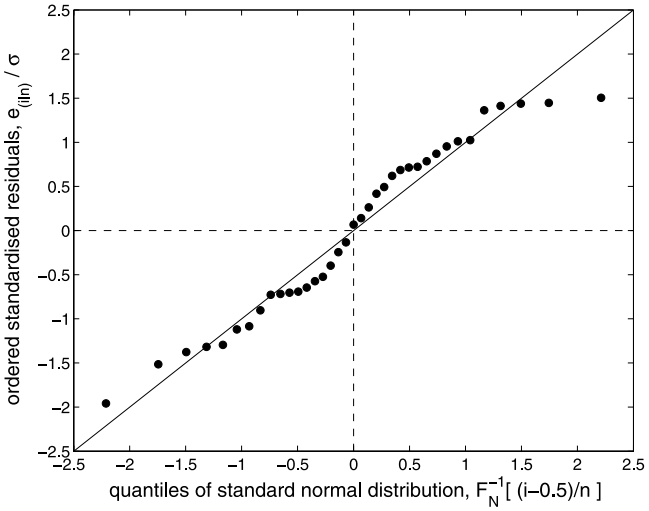


**Figure 7.** A quantile-quantile plot (QQ plot) showing that the distribution of residuals is not Gaussian. This is for the $I_{DV}$ versus $B_r$ regression. The ordered, standardized residuals $\Delta B_{ri}/\sigma$ (where $\sigma^2$ is $\Sigma_i\ \Delta B_{ri}^2/(n-2)$ and $n$ is the number of samples) are shown as a function of the corresponding quantiles of a standard normal distribution. The deviations from the line of slope 1 shown reveal departures from a normal distribution of standard distribution σ [*Wilkes*, 1995; *van Storch and Zwiers*, 1999].

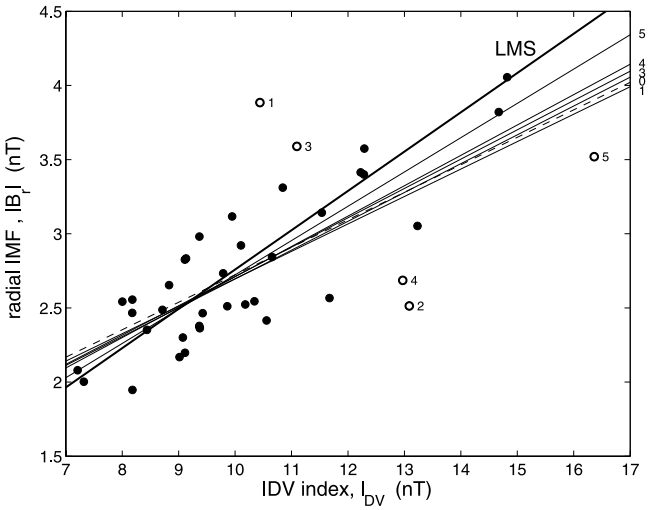


**Figure 8.** Effect of removing outliers from the fit of radial IMF $B_r$ to the *IDV* index, $I_{DV}$. The dashed and thin solid lines use minimization A ($\Delta = \Sigma_i\ \Delta B_{ri}^2$), the dashed line is for all data points, and the thin solid lines are the best fit regressions after successive removal of the worst outliers. The outliers are shown by the open circles and are numbered (1 being the largest fit residual and the first to be removed), as is the best fit regression line after the removal of that outlier. The thick solid line is for the more robust LMS minimization ($\Delta = [\Delta B_{ri}^2]_{med}$).

$B_r$ as a function of *IDV* that are obtained as successive worst outliers are removed. The points removed are also shown. It can be seen that as more outliers are removed the slope of the regression, $m$, systematically increases and is found to increase asymptotically to that for minimization method D (LMS). This is to be expected as LMS is designed to be more robust to the effect of outliers. Table 1 shows that the percentage change λ derived by LMS regression is 51% for $B$ and 85% for $B_r$.

## 8. A More Direct Test of the Results of LEA99

[24] As discussed above, by far the dominant source of scatter in a regression between the IMF parameters and a geomagnetic activity index is the fact that the latter is only an approximate indicator of the former. Minimization method B therefore has the advantage of minimizing the true errors (in *IDV*) rather than their projection onto another parameter ($B$ or $B_r$). For OLS, minimization B is used to predict $I_{DV}$ from $B$ or $B_r$. Figure 9 shows the regression fits of $I_{DV}$ as a function of $B_r$, for successive removal of outliers, and for the corresponding LMS regression. Figure 10 shows the analysis of residuals for the LMS fit (equivalent to that given in Figure 6 for the fit by SC05). Figure 9 (left) shows no systematic change in the variance, i.e., the residuals do display homoskedacity, and the LMS regression fit to the points in Figure 9 (left) reveals no systematic bias. In addition, the slope of Figure 9 (right) is small ($s$ = 0.151, with an intercept $c$ = −1.398 nT). The potential bias introduced by this $s$ is small: For example, using the values of $B_r$ deduced by LEA99 for 1903 and 1956 (1.614 and 3.441 nT, respectively, giving a percent variation of λ = 113%), the LMS fit tested by Figure 10 gives best fit *IDV* values of 6.64 nT and 12.49 nT (an 88% change in predicted *IDV*, which compares well with the 84% change in the observed *IDV* given by SC05). If we correct these estimates for the possible bias indicated in Figure 10 (right), these values become 6.17 nT and 13.07 nT (a 112% change). (To put this potential bias in context, remember that the same

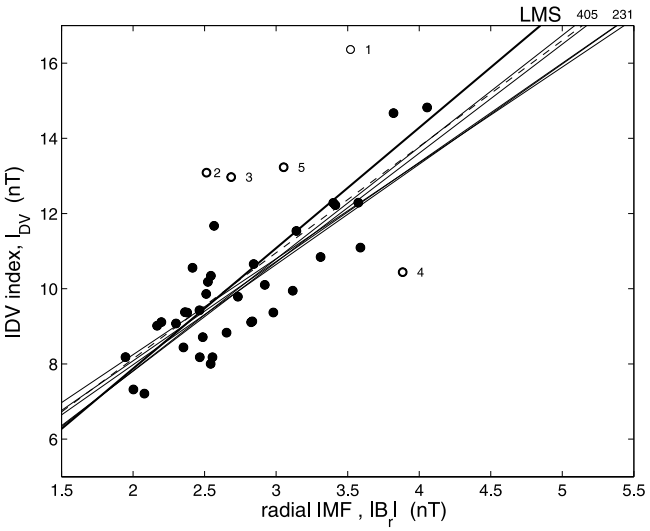


**Figure 9.** Same as Figure 8, but for the fit of the *IDV* index, $I_{DV}$, to the radial IMF $B_r$. The dashed and thin solid lines use minimization B ($\Delta = \Sigma_i\ \Delta I_{DVi}^2$). The thick solid line is for the more robust LMS minimization ($\Delta = [\Sigma_i\ \Delta I_{DVi}^2]_{med}$).

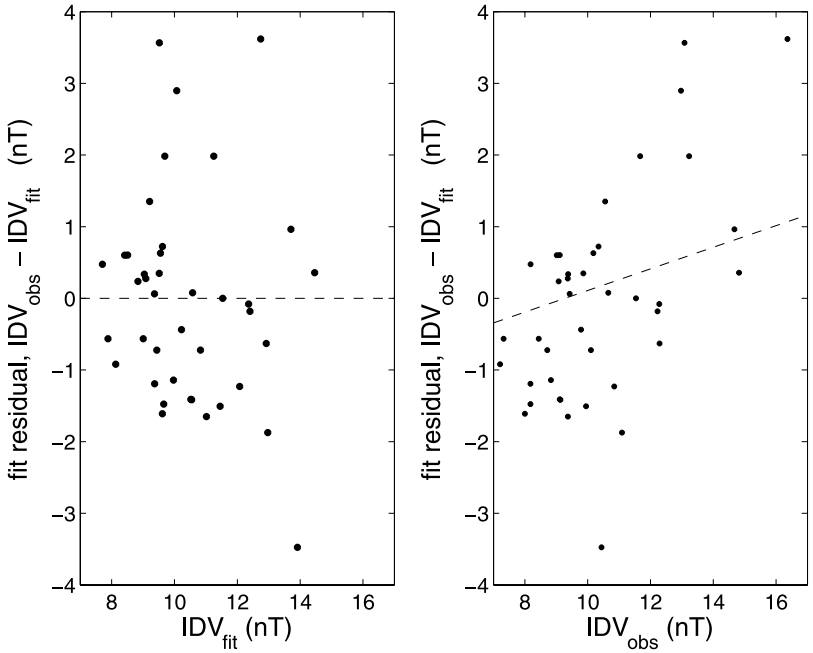


**Figure 10.** Analysis of fit residuals for the LMS fit given in Figure 9. The fit residuals ($IDV_{obs}$-$IDV_{fit}$) are shown as a function of (left) $IDV_{fit}$ and (right) $IDV_{obs}$, where $IDV_{obs}$ is the value given by SC05 and $IDV_{fit}$ is the fitted value from OLS regression with the radial IMF, $B_r$. The dashed line is the LMS linear regression fit to the annual points shown. Figure 10 (left) shows that the fit displays good homoskedacity (i.e., the variance does not increase systematically with $IDV_{fit}$) and reveals no systematic bias. The slope of the LMS regressions in Figures 10 (left) and 10 (right) are $s = -1.43 \times 10^{-14}$ and $s = 0.151$, respectively, and the corresponding intercepts are $c = 1.47 \times 10^{-14}$ nT and $c = -1.40$ nT.

procedure applied SC05's fit changed a $\lambda = 38\%$ change in extrapolated $B$ into a 192% change.)

[25] If we adopt the variation of $B_r$ since 1868 derived by LEA99 (which used $T = 1$ hour and gave a percent variation in $B_r$ of $\lambda = 113\%$) we can use the OLS and LMS regressions shown in Figure 9 to predict the variation in $I_{DV}$. The resulting 11-year running means predicted are shown in Figure 11: the thin solid lines are for the OLS fits with successive removal of outliers and the thick dashed line is for the more robust LMS regression. It can be seen that the results are not strongly influenced by outliers in this case. Figure 11 also shows (as the shaded area) the 11-year running means of SC05's *IDV* index.

[26] The major difference between observed and predicted *IDV* occurs after 1957 when *IDV* predicted from LEA99 is consistently higher than the observations given by SC05. This arises because of two factors: (1) LEA99 used hourly averages ($T = 1$ hour) and (2) there is a discontinuity in the *aa* index compared to *IDV* which SC05 attribute to a discontinuity in the calibration of *aa* when the northern hemisphere station moved from Abinger to Hartland. This discontinuity is also deduced in the analysis of median interhour variability by *Lockwood et al.* [2006], although they do not find it to be as great as in SC05's *IDV*. (In fact, *Lockwood et al.* [2006] find that the maximum possible correction required gives a variation close to *IDV* but that the optimum correction is about half of this.) The effect of these two factors is illustrated by the thick solid line which is the variation for the LMS regression with the corrected $B_r$ variation derived by applying the method of LEA99 to the

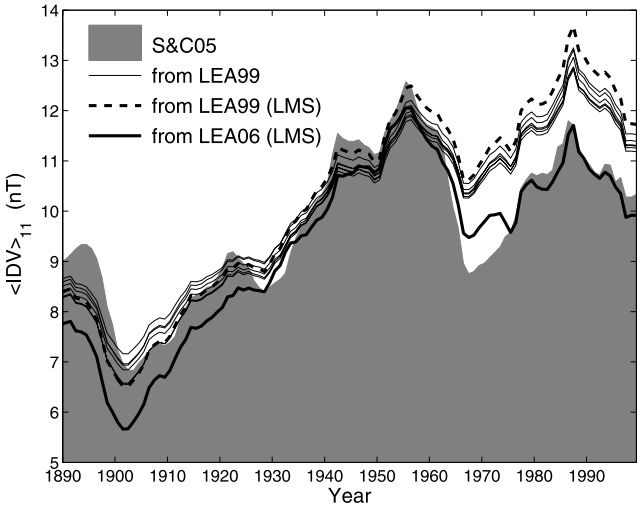


**Figure 11.** Predicted and observed 11-year running means of the *IDV* index using the regressions of *IDV* as a function of $B_r$ shown in Figure 9. The thick dashed line is the predicted variation using the LEA99 estimate of the open solar flux variation (which varies by $\lambda = 113\%$ between 1903 and 1956) using the LMS regression. The thin solid lines are the OLS fits for successive removal of outliers. The thick solid line is the predicted *IDV* variation using the LMS regression for the upper limit of uncertainty of the revised *aa* index by *Lockwood et al.* [2006] (LEA06), which gives a variation of $\lambda = 140\%$ in the open solar flux between 1903 and 1956 for $T = 1$ day. The observed variation of *IDV* given by SC05 is shown by the shaded area.

*aa* variation with the maximum possible (as opposed to the optimum) correction, and with an averaging timescale $T = 1$ day. For this $B_r$ variation, the percent change in the open solar flux between 1903 and 1956 is $\lambda = 140\%$.

[27] The key point is that before 1960, the variation of observed *IDV* and that predicted from the open flux variation of LEA99 are very similar indeed (for either OLS or LMS regression). Thus this analysis shows that *IDV*, as derived by SC05, is indeed fully consistent with the 113% change in open solar flux between 1903 and 1956 reported by LEA99. For the maximum possible correction factor to *aa* deduced by *Lockwood et al.* [2006], the open flux variation between 1903 and 1957 is 140%, and this does predict a slightly larger change in *IDV* than is reported by SC05.

[28] The OLS regression prediction of $I_{DV}$ from $B_r$ presented in this section is much more reliable and robust than the corresponding OLS prediction of $B_r$ (or $B$) from $I_{DV}$. This is because (1) we are minimizing the true errors (in *IDV*) in the regressions; (2) the residuals do not display heteroskedacity; (3) the residuals do not show any systematic bias with fitted values; (4) the residuals show a much smaller bias against the observed values; and (5) the fits are much less influenced by outliers.

## 9. Analysis Procedure

[29] We do not consider any of the regression analyses presented here to be optimum, because we have only corrected the simple ordinary linear least squares regression between the IMF parameters and *IDV* that was used by SC05. None of the fits presented here (nor that by SC05) have a Gaussian distribution of the residuals, indicating that *IDV* varies with other related parameters and/or that variations are nonlinear. This violates a key assumption of OLS regression. The correlation coefficient with $B$ is high at $r = 0.87$, but even this means that only $r^2 = 0.76$ of the variation in *IDV* is explained by $B$ alone. The correlation coefficient between $B_r$ and *IDV* is $r = 0.72$, meaning that only $r^2 = 0.51$ of the variation in *IDV* is explained by $B_r$ alone. This is not surprising because neither $B$ nor $B_r$ is directly related to energy coupling from the solar wind to the magnetosphere, the controlling IMF component being $B_z$ in the GSM frame of reference. In addition, the energy density available in the solar wind depends on its speed, $V_{sw}$, and number density, $N_{sw}$. These factors are not accounted for in the correction we have presented here of SC05's simple linear, single-parameter statistical OLS regression. However, these factors were accounted for in LEA99's nonlinear, detection/correlation procedure, based on dimensional analysis and known physics. LEA99 explained all but 2% of the variation in the *aa* index by applying the physics-based dimensional analysis of *Vasyliunas et al.* [1982], as implemented by *Stamper et al.* [1999]. They obtained a correlation coefficient of $r = 0.97$, meaning that $r^2 = 0.94$ of the variation in *aa* is explained by the combination of $B$, IMF orientation and solar wind speed and number density derived by LEA99. Finch et al. (Solar wind-magnetosphere coupling functions on timescales of 1 day to 1 year, submitted to *Annales Geophysicae*, 2006) show that most of the remaining variation in *aa* actually arises from data gaps in the IMF data series which mean that the highest $r$ that could have been obtained is 0.98 ($r^2 = 0.96$).

## 10. Conclusions

[30] There a number of problems with the analysis of SC05:

[31] 1. They underestimate the long-term change in their own results.

[32] 2. They do not make the important distinction between the average radial IMF and the average IMF magnitude (thereby neglecting disconnected flux and small-scale heliospheric structure).

[33] 3. The simple ordinary linear least squares (OLS) regression they employ yields residuals that show heteroskedacity and which do not have a Gaussian distribution, violating key assumptions of ordinary linear regression.

[34] 4. The regression results of SC05 are strongly influenced by outliers, which apply great leverage to their regression fit. More reliable regressions are obtained by least median squares (LMS) regression and from Bayesian statistics.

[35] 5. SC05 have attempted to fill in the data gaps in the IMF data using a 27-day recurrence technique, despite the very low autocorrelation functions of the IMF at 27 day lags. We here show that the missing IMF data has only a small effect as they are largely random in their effect on geomagnetic activity: however, neglecting the data gaps in annual means or attempting to fill them in using SC05's procedure causes one to (slightly) underestimate the long-term drift.

[36] All of the above act to reduce SC05's estimate below the real drift in the open solar flux. Using the more robust least median squares (LMS) regression to reduce the dependence on outliers, we find the percentage change in decadal means, $\lambda$, is 51% for $B$ and 85% for $B_r$ (the latter being similar to the $\lambda = 84\%$ change in *IDV* itself). In a future paper (Lockwood et al., submitted manuscript, 2006), we will show Bayesian statistics calls for the use of minimization method B in the prediction of $B_r$ from *IDV*, which we here show gives a percentage change of $\lambda = 132\%$ for $B_r$ and the open solar flux.

[37] However, a more reliable and direct test of the results of LEA99 is to predict the *IDV* index from the variation in open solar flux that LEA99 reported. This test shows that although there is indeed a greater decrease in geomagnetic activity after 1960 than predicted for the *aa* index data used by LEA99, the variation in open solar flux derived by LEA99 for before 1960 (giving a 113% change) is fully consistent with the *IDV* index derived by SC05.

[38] **Acknowledgments.** The authors are grateful to Yi-Ming Wang for provision of the PFSS data. We also thank Goeff Daniell of Southampton University for checking the statistical methods used here and for the development of a Bayesian statistics analysis which will be reported in a future publication.

[39] Amitava Bhattacharjee thanks the reviewer for the assistance in evaluating this paper.

## References

Arge, C. N., and V. J. Pizzo (2000), Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates, *J. Geophys. Res.*, *105*, 10,465–10,479.

Lockwood, M. (2002), Relationship between the near-Earth interplanetary field and the coronal source flux: Dependence on timescale, *J. Geophys. Res.*, *107*(A12), 1425, doi:10.1029/2001JA009062.

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

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.

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

Lockwood, M., D. Whiter, B. Hancock, R. Henwood, T. Ulich, H. J. Linthe, E. Clarke, and M. A. Clilverd (2006), The long-term drift in geomagnetic activity: Calibration of the *aa* index using data from a variety of magnetometer stations, *Ann. Geophys.*, in press.

Rousseeuw, P. J. (1984), Least median of squares regression, *J. Am. Stat. Assoc.*, *79*(388), 871–880.

Rousseeuw, P. J., and A. M. Leroy (1987), *Robust Regression and Outlier Detection*, John Wiley, Hoboken, N. J.

Sokal, R. R., and F. J. Rohlf (1995), *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd ed., W. H. Freeman, New York.

Stamper, R., M. Lockwood, M. N. Wild, and T. D. G. Clark (1999), Solar causes of the long-term increase in geomagnetic activity, *J. Geophys Res.*, *104*, 28,325–28,342.

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.

van Storch, H., and F. W. Zwiers (1999), *Statistical Analysis in Climate Research*, Chapter 8, Cambridge Univ. Press, New York.

Vasyliunas, V. M., J. R. Kan, G. L. Siscoe, and S.-I. Akasofu (1982), Scaling relations governing magnetospheric energy transfer, *Planet. Space Sci.*, *30*, 359–365.

Wang, Y.-M., and N. R. Sheeley Jr. (1995), Solar implications of Ulysses interplanetary field measurements, *Astrophys. J.*, *447*, L143.

Wang, Y.-M., S. H. Hawley, and N. R. Sheeley Jr. (1996), The magnetic nature of coronal holes, *Science*, *271*, 464–469.

Wang, Y.-M., N. R. Sheeley Jr., J. L. Phillips, and B. E. Goldstein (1997), Solar wind stream interactions and the wind speed expansion factor relationship, *Astrophys. J.*, *488*, L51–L54.

Wang, Y.-M., N. R. Sheeley Jr., and N. B. Rich (2000a), Evolution of coronal streamer structure during the rising phase of solar cycle 23, *Geophys. Res. Lett.*, *27*(2), 149–152.

Wang, Y.-M., N. R. Sheeley Jr., D. G. Socker, R. A. Howard, and N. B. Rich (2000b), The dynamical nature of coronal streamers, *J. Geophys. Res.*, *105*(A11), 25,133–25,142.

Whang, Y. C., Y.-M. Wang, N. R. Sheeley Jr., and L. F. Burlaga (2005), Global structure of the out-of-ecliptic solar wind, *J. Geophys. Res.*, *110*, A03103, doi:10.1029/2004JA010875.

Wilkes, D. S. (1995), *Statistical Methods in the Atmospheric Sciences*, Chapter 6, Elsevier, New York.

I. Finch, M. Lockwood, and R. Stamper, Rutherford Appleton Laboratory, R25, Room 1-04, Chilton OX11 0QX, UK. (m.lockwood@rl.ac.uk)

A. P. Rouillard, Solar-Terrestrial Physics Group, Department of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK.