Submitted version
File LockwoodandMcWilliams_2021b_S1.doc
submitted to: J. Geophys. Res. 9th September 2021

On optimum solar wind – magnetosphere coupling functions for transpolar voltage and planetary geomagnetic activity

Mike Lockwood[1] and Kathryn A. McWilliams[2]

[1] *Department of Meteorology, University of Reading, Earley Gate, Reading, RG6 6BB, UK*

[2] *Institute of Space and Atmospheric Studies, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 5E2, Canada*

Abstract. We use 65,133 hourly averages of transpolar voltage (Φ_{PC}) from observations made over 25 years by the SuperDARN radars, along with simultaneous interpolated *am* geomagnetic index values, and study their optimum interplanetary coupling functions. We find lags of 18min. and 31min. for Φ_{PC} and *am*, respectively, and fit using a general coupling function with four free fit exponents. To converge to a fit, we need to average interplanetary parameters and then apply the exponent which is a widely-used approximation: we show how and why this is valid for all interplanetary parameters, except the factor quantifying the effect of the clock angle of the interplanetary magnetic field, $sin^{d}(\theta/2)$, which should be computed at high time resolution and then averaged. We demonstrate the effect of the exponent *d* on the distribution, and hence weighting, of samples and show it causes *d* to be best determined from the requirement that the coupling function is a linear predictor, which yields d=2.50±0.10 for Φ_{PC} and d=3.00±0.22 for *am*. To check for overfitting, fits are made to half the available data and tested against the other half. Ensembles of 1000 fits are used to study the effect of the number of samples on the distribution of errors in individual fits and on systematic biases in the ensemble means. We find only a weak dependence of solar wind density for Φ_{PC} but a significant one for *am*. The optimum coupling functions are shown to be significantly different for Φ_{PC} and *am*.

Plain Language Abstract. Coupling functions are mathematical combinations of variables observed in the solar wind, just before it impacts near-Earth space. They are used to predict the

effect that the solar wind will have (or, for retrospective studies, will have had) on the space-weather environment of the Earth. There is a very wide variety of proposed optimum forms for coupling functions in the literature, some of which work better than others and we show which performs best depends on which terrestrial disturbance indicator we are trying to predict and on what timescale. We look at the validity of some commonly-used assumptions made when compiling a coupling function and, using an unprecedentedly large data set of two different types of terrestrial space weather disturbance indicator, we derive the optimum coupling functions and their statistical uncertainties. We show that that the required coupling functions are significantly different in the two cases. The results establish some important principles for the development of these coupling functions and show they need to be tailored to the specific space weather disturbance indicator and timescale that they aim to predict.

Main points

• 1. Using a very large dataset we analyze the sources and effects of noise in correlation studies used to derive solar wind coupling functions

• 2. We study effects of weighting by the distribution of samples which varies with the choice of IMF orientation factor and averaging timescale

• 3. The optimum coupling functions for transpolar voltage and planetary geomagnetic activity are significantly different.

1. Introduction.

Coupling functions are combinations of interplanetary parameters that are used to quantitatively predict terrestrial space weather indicators and indices. Ideally, they should have a linear relationship with the index or measured parameter that they aim to predict. There are a huge number of combinations that have been proposed and tested since correlations between interplanetary parameters measured by spacecraft and terrestrial disturbance indices became possible (*Arnoldy*, 1971). The concept of a combination of parameters capturing their net influence (i.e., a coupling function) grew out of the PhD studies of *Perreault* (1974). An excellent review of the development of coupling functions, the theories behind them and the empirical fits, has been given by *McPherron et al.* (2015).

Coupling functions have generally taken the basic mathematical form of the product of measured parameters, each to the power of an exponent. Parameters used have been the interplanetary magnetic field (IMF), $B = |\vec{B}|$ or its transverse component perpendicular to the Sun-Earth line, $B_{\perp}$; the solar wind speed, V_{SW}; the solar wind number density N_{SW} or its mass density $\rho_{SW} = m_{SW}N_{SW}$ (where m_{SW} is the mean ion mass); and (for timescales shorter than about 1 year), a factor to allow for the orientation of the IMF in the Geocentric Solar Magnetospheric (GSM) frame of reference, such as the clock angle in GSM, θ. We here denote magnetic field exponents by a, mass density or number density exponents by b, solar wind speed exponents by c and IMF orientation factor exponents by d.

Some improvements to this basic multiplicative form have been suggested in the form of additive terms. For example, *Newell et al.* (2008) proposed adding to a term designed to predict the dayside magnetopause reconnection voltage with a smaller term to predict the voltage generated by non-reconnection "viscous-like" interaction. *Lockwood* (2019) proposed a development to energy-transfer coupling functions whereby, in addition to the energy extracted from the dominant energy flux in the solar wind (namely the kinetic energy flux of the particles), the smaller one due to the solar wind Poynting flux is added. Given that the Poynting flux in the solar wind is two orders of magnitude smaller than the particle kinetic energy flux, this appears an unnecessary complication: however, the Poynting flux enters the magnetosphere without the relative inefficiency with which kinetic energy of the solar wind is converted into Poynting flux by currents flowing in the bow shock, magnetosheath and magnetopause (*Cowley*, 1991; *Lockwood*, 2004; *Ebihara et al.*, 2019).

Other, more complex, forms with combinations of additive and multiplicative terms have been proposed (e.g., *Borovsky*, 2013; *Luo et al*. 2013). The formulation of *Luo et al*. (2013) aims take account of daily and seasonal variations in the terrestrial space weather index predicted (that are due to station locations and orientation of the Earth's dipole) and non-linearities caused by the expansion and contraction of the polar cap as solar wind driving varies. It also removes rapid fluctuations using low-pass filters. The result is that it is highly complex and, as noted by *McPherron et al.* (2015), it is unclear how many free parameters are present in this coupling function, but they estimate that it is of order 35. Because these more complex formulations add to the number of free fit parameters, this greatly increases the problem of statistical "overfitting" (*Chicco*, 2017). Overfitting occurs when a fit has too many degrees of freedom and it can start to fit to the noise in the training data subset, which is not the same as the noise in the test or operational data. As a result, the fit has reduced predictive accuracy. This is a recognized pitfall when signal-to-noise ratio in the data is low, as is usually the case in disciplines such as climate science (*Knutti et al.,* 2006) or population growth (*Knape & de Valpine*, 2011), but has not often been considered in space physics in the past. However, this is now changing with the advent of systems analysis of the magnetosphere and the application of machine-learning techniques to space weather data (e.g., *Camporeale*, 2019; *Stephens et al.,* 2020).

Hence the effect of adding more terms, even if based on sound physical theory, is not always a positive one. For example, *Lockwood* (2019) showed that although adding the solar wind Poynting flux term does increase the correlation with the geomagnetic *am* index and that the increase for daily or shorter timescales is a small but statistically significant improvement (at over the 3-σ level), the improvement for annual or Carrington rotation means was not statistically significant: hence in the latter cases no statistically significant improvement was achieved, despite the number of free fit variables being doubled from 1 to 2 and the additional term being based on sound theory.

Table 1 lists a number of coupling functions that have been developed, based on theory and/or empirical fitting (*Balikhin et al.* , 2010; *Bargatze et al* , 1986; *Borovsky,* 2013; *Burton et al.* , 1975; *Cowley* , 1984; *Feynmann & Crooker,* 1978; *Finch & Lockwood* , 2007; *Kan and Lee* , 1979; *Lockwood* , 2019; *Lockwood et al.* , 2014; *Lockwood et al* , 2019a; *Luo et al.* , 2013; *McPherron et al.* , 2015; *Milan et al* , 2012; *Murayama* , 1982; 1986; *Newell et al.*, 2007; *Perreault & Akasofu* , 1978; *Scurry and Russell* , 1991; *Siscoe et al.*, 2002; *Svalgaard & Cliver,* 2005; *Temerin & Lee* , 2006; *Vasyliunas et al* , 1982; *Wang et al.*, 2013; *Wygant et al.*, 1983). This list is very far from complete, but examples have been chosen to illustrate both the variety and the similarities, and also some of the principles of the physical theories used to develop them.

Table 1 gives the timescale τ on which each coupling function was derived and/or has been tested and deployed. It is noticeable that at larger τ, simpler coupling functions have been very successful in yielding very high correlations (*Finch and Lockwood*, 2007). The averaging timescale of the interplanetary and the terrestrial data that are compared is a crucial consideration because solar wind parameters have a variety of autocorrelation times which means that their distributions of values change with τ in different ways (*Lockwood et al.*, 2019a; 2019b). However, this is not often considered when compiling a coupling function and τ is not even explicitly defined in several of the publications (in several cases in Table 1, τ could only be defined from the data plots presented).

One idea that has been proposed is that there is a "universal coupling function" that best predicts all terrestrial space weather indices (*Newell et al.*, 2007, 2008). This idea runs counter to the method now routinely used to reconstruct interplanetary parameters from historic observations of geomagnetic activity. These reconstructions exploit the finding that different geomagnetic indices have different responses to interplanetary parameters and so combinations of them can be used to infer the separate interplanetary parameters. This was inherent in the reconstruction of open solar flux from historic observations of geomagnetic activity by *Lockwood et al* (1999) but first explicitly pointed out and used to extract more than one parameter by *Svalgaard et al.* (2003), who noted that on annual timescales the IMF B and solar wind speed V_{SW} could both be derived from any combination of geomagnetic indices that had different dependencies on these two parameters (i.e., different optimum coupling functions). This has been exploited by *Svalgaard and Cliver* (2007), *Rouillard et al.* (2007), *Lockwood et al.* (2009), *Lockwood and Owens* (2011), and *Lockwood et al.* (2014). These methods and results have developed from simple single fits to large ensembles of fits allowing for uncertainties and been reviewed by *Lockwood* (2013). If different indicators of geomagnetic activity have different optimum coupling functions, it means that other space weather activity indicators, such as transpolar voltage, cannot share the same optimum coupling as all, if any, of the geomagnetic activity indices. We here investigate the differences between the optimum coupling functions for transpolar voltage Φ_{PC} and the global *am* geomagnetic index which has been shown to have the most uniform response to solar wind forcing with Universal Time and time of year by virtue of the relative uniformity of the observing network and its use of area-based weighting functions (*Lockwood et al*., 2019c).

Table 1 shows that many of the proposed coupling functions predict a role of solar wind number density N_{SW} or mass density $\rho_{SW} = m_{SW}N_{SW}$ (where m_{SW} is the mean ion mass) as contributing to solar wind energy coupling and/or to the driving of magnetospheric convection. For energy considerations, this is mainly because ρ_{SW} and N_{SW} control the dominant (kinetic) energy flux in the solar wind ($\frac{1}{2}\rho_{SW}V_{SW}^3$) but it has been proposed that solar wind dynamic pressure ($P_{SW} = \rho_{SW}V_{SW}^2$) also has an effect. This is through altering the cross-sectional area that the magnetosphere presents to the solar wind flow (*Vasyliunas et al*, 1982) and also via the compression of the near-Earth tail, which enhances the magnetic energy density stored there for a given open magnetospheric flux, thereby enhancing the current in the auroral electrojet of the substorm current wedge when that stored energy is released during a substorm expansion phase (see review by *Lockwood*, 2013). Such a dependence of geomagnetic disturbance in the substorm current wedge region was isolated and identified by *Finch et al.* (2008). This would be in addition to the dependence on ρ_{SW} and V_{SW} due to the energy flux in the solar wind and/or due to the magnetic reconnection which generates the open flux. *Lockwood et al.* (2020a; b) show that P_{SW} has an effect on geomagnetic activity that is explicitly distinct from that of power input into the magnetosphere. In addition, the squeezing of the near-Earth tail by P_{SW} would elevate the magnetic shear across the cross-tail current sheet, and hence the total current in that sheet. This could enhance the nightside reconnection voltage Φ_N that closes open field lines. The expanding contracting polar cap (ECPC) model predicts that this would elevate the transpolar voltage Φ_{PC} which is influenced at any one instant by the reconnection voltages in both the dayside magnetopause Φ_D and the cross-tail current sheet Φ_N (*Lockwood, 1991; Cowley and Lockwood,* 1992, *Lockwood and McWilliams*, 2021). However, we need to consider the averaging timescale used, τ. If τ is short compared to the substorm cycle duration we would expect Φ_{PC} to reflect the enhanced Φ_N, and so show some dependence on P_{SW} from this effect of squeezing the tail. On the other hand, if τ is long compared to the substorm cycle duration, the average Φ_N tends to Φ_D and we would therefore expect Φ_{PC} to show only any dependence that Φ_D has on P_{SW} which appears to be considerably smaller (*Lockwood and McWilliams*, 2021). However, we note that it has long been proposed that P_{SW} has an effect on Φ_D through increasing the magnetic shear across the dayside magnetopause during southward IMF (e.g., *Scurry and Russell*, 1991).

This discussion of the role of solar wind dynamic pressure is just one example of an important general point – namely that there are a great many processes simultaneously at play in driving the terrestrial space weather response. To allow for these, solar wind coupling functions have evolved away from having theoretically-derived exponents *a*, *b*, *c* and *d* (which were often integers or ratios of integers) to empirically-fitted non-integer values. Hence for the example of P_{SW} effects on the near-Earth tail we do not complicate the coupling function with an additional term or weighting branching ratio, rather we allow the exponents *b* and *c* (in the terms ρ_{SW}^{b} and V_{SW}^{c}) to vary to allow for such an effect and we would expect such an effect of P_{SW} to raise the exponent *b* and raise *c* by twice as much. Hence combinations of mechanisms can be allowed for as long as their effects are multiplicative. To bring theoretical and empirical approaches together, *Borovsky* (2013) used the approach of making a complex theoretical derivation and the reducing to a simple multiplicative form with approximations to derive exponents; however, the uncertainties introduced by any one approximation are not always apparent.

There is one last important point to note about coupling functions that is discussed further in the final section of the present paper. None of the forms listed in Table 1 allow for the pre-existing state of the magnetosphere. There are many reasons to expect non-linear magnetospheric responses. For example, the response to a given solar wind forcing quantified by a coupling function will depend on how much open magnetospheric flux already exists at the time but in addition is very likely to also depend on how enhanced the ring current is at the time and/or the state of the mid-tail plasma sheet and cross-tail current sheet. These effects all depend upon the prior history of solar wind-magnetosphere coupling. There are also regular diurnal and annual effects to consider such as dipole tilt effects and seasonal effects in the ionosphere. If they are neglected, all these factors are a source of noise for correlation studies between interplanetary coupling functions and terrestrial disturbance indices.

In this paper, we do not attempt to compare the performance of this large number of proposed coupling functions. Such test have been carried out in the past, often as part of an evaluation of a newly-proposed function (e.g., *Newell et al.*, 2007). Detailed tests against model output were carried out for three coupling functions by *Spencer et al* (2009) and the performance of seven coupling functions in predicting mid-latitude geomagnetic range indices was compared for a range of timescales τ between 1 day and 1 year by *Lockwood and Finch* (2007). *Newell et al.* (2007) compared 20 coupling functions against 10 terrestrial indices at hourly resolution. Rather, we here establish some general principles and apply a generalized common form of coupling function to an unprecedently large dataset containing two different indicators of terrestrial space weather disturbance (the transpolar voltage and a planetary geomagnetic index) to see if they are significantly different or can be predicted by a common "universal" coupling function.

1-i. Coupling functions based on energy considerations

Lockwood (2019a; b) have shown that the *am, AL* and *SML* geomagnetic indices, which all respond primarily to the substorm current wedge, are well predicted over a range of timescales by the estimated power input to the magnetosphere, P_α (*Vasyliunas et al*., 1982). This coupling function is given by the product of the dominant energy flux in the solar wind (due to the kinetic energy flux of the particles), the cross-sectional area of the magnetosphere it is incident upon, and a dimensionless transfer function (t_r the fraction of the incident power that crosses the magnetopause into the magnetosphere).

$$P_\alpha \;=\; (\rho_{sw} V_{sw}^2)/2) V_{sw} \times (\pi L_o^2) \times t_r \qquad (1)$$

where L_o is the radius of cross-section of the magnetosphere presented to the solar wind flow.

The dayside magnetosphere is assumed to be constant in shape so that $L_o = cL_s$ where $c = L_o/L_s$ is the dayside magnetopause shape factor (assumed constant) and L_s is the stand-off distance of the nose of the magnetosphere which is derived from pressure balance between the geomagnetic field and dynamic pressure of the solar wind, P_{SW} (*Farrugia et al*., 1989):

$$L_o \;=\; cL_s \;=\; ck_1 (M_E^2 /\, P_{sw} \mu_o)^{1/6} \qquad (2)$$

where k_1 is the pressure factor for shocked supersonic flow around a blunt nose object, M_E is the magnetic moment of the Earth and μ_o is the permeability of free space (the magnetic constant) *Vasyliunas et al.* (1982) use a dimensionless transfer function of the form:

$$t_r \;=\; k_2 \;\, M_A^{2\alpha} \, sin^d(\theta/2) \qquad (3)$$

where the solar wind Alfvén Mach number is $M_A = V_{SW}\,(\mu_o \rho_{SW})^{1/2}/B$, and k_2 is a constant and α is called the "coupling exponent" that arises from the unknown dependence of t_r on M_A and is the one free fit parameter. θ is the IMF clock angle in the GSM frame of reference. The dependence of t_r on M_A arises from the fact that the dominant energy flux in the undisturbed solar wind, the kinetic energy flux of the particles, is converted into the Poynting flux that enters the magnetosphere by the currents that flow in the bow shock and magnetosheath (*Cowley*, 1991, *Lockwood*, 2004; 2019; *Ebihara et al*., 2019). From (1), (2) and (3)

$$P_{\alpha} = k\, B^{2\alpha} \rho_{sw}^{(2/3-\alpha)} V_{sw}^{(7/3-2\alpha)} sin^{d}(\theta/2) \quad (4)$$

Where $\{M_E^{2/3} c^2 k_1 k_2 \pi / (2\mu_o^{(1/3-\alpha)})\}$ are rolled into the constant k. However, note that the secular variation in M_E, and hance k, can be allowed for from models of the intrinsic geomagnetic field in long-term reconstructions of space weather conditions (*Lockwood et al.*, 2017). Despite allowing for B, ρ_{SW}, V_{SW} and θ, the coupling function P_{α} has only the one free fit parameter, the coupling exponent α that arises from an unknown dependence of the transfer function on the solar wind Mach number. This means that P_{α} is much less prone to overfitting than functions that have separate exponents for the parameters. (Essentially, the exponents of B, ρ_{SW}, V_{SW} are related by the theory, and all are determined by just α).

The IMF orientation factor $sin^{d}(\theta/2)$ was not treated as an independent variable by *Vasyliunas et al*. (1982). However, these authors did outline a test which was used to find that d = 2 was the required factor for the optimum (best-fit) α. The same test for other applications of the formulation by *Lockwood et al.* (2019a; b) found a slightly different α (and that it varies with timescale) and this made d = 4 marginally better . Table 1 shows that $sin^{d}(\theta/2)$ is a commonly-used IMF orientation factor for low τ, particularly with d = 4. However, a range of d between 1 and 6 has been proposed in the literature. We here note that the test by *Vasyliunas et al*. (1982) has the very important implication that the optimum d is not independent of the other parameters in the coupling function.

In their paper, *Vasyliunas et al.* (1982) are somewhat uncertain as to whether they should employ the transverse component of the IMF, $B_{\perp}$ (the magnitude in the GSM YZ plane) or the full IMF magnitude $B = (B_X^2 + B_{\perp}^2)^{1/2}$. They found it made only a minor difference in practice but opted to use $B_{\perp}$ in their text and equations. Their argument was that B_X is not relevant because the field was draped over the nose in the magnetosheath. However, this choice is somewhat inconsistent theoretically because the IMF enters into their coupling function only through the Alfvén Mach number M_A in the interplanetary (unshocked) field and that depends on B and not on $B_{\perp}$. On the other hand, $B_{\perp} sin^{d}(\theta/2)$ is physically meaningful as a way of quantifying the southward component if the IMF in GSM coordinates.

1-ii. Coupling functions based on voltage considerations

In addition to planetary geomagnetic activity, we are aiming to predict transpolar voltage Φ_{PC}, we would expect a coupling function based on the interplanetary magnetic field to be more appropriate. Many studies (e.g., *Cowley*, 1984; *Reiff and Luhmann*, 1986), suggest that the transpolar voltage Φ_{PC} is well predicted by the dawn-to-dusk interplanetary electric field

$$E_{sw} = V_{sw} B_S \approx B_\perp V_{sw} sin^d(\theta/2) \quad (5)$$

Because the voltage applied by the solar wind across the diameter of the magnetosphere is $2L_o E_{SW}$, we can define the reconnection efficiency (the fraction of incident interplanetary field lines captured by magnetopause reconnection) η as

$$\eta = \Phi_{PC}/(2L_o E_{sw}) \quad (6)$$

We can then make the same assumption about the dayside magnetopause as was used to generate P_α and again use pressure equilibrium with the solar wind dynamic pressure (*Siscoe et al.*, 2002)

$$\Phi_{PC} = 2\eta\, c L_s E_{sw} = 2\eta\, c E_{sw}\{2kM_E^2/(\mu_o \rho_{sw} V_{sw}^2\}^{1/6} = \eta\, E_{sw} \kappa \{\rho_{sw} V_{sw}^2\}^{-1/6} \quad (7)$$

where $\kappa = 2c\{2kM_E^2/\mu_o\}^{1/6}$. From (5), (6) and (7) we have a theoretical prediction of Φ_{PC}, which we term Φ_{SW} (the predicted value of Φ_{SW} from solar wind parameters)

$$\Phi_{sw} = \eta\, \kappa B_\perp \rho_{sw}^{-1/6} V_{sw}^{2/3} sin^d(\theta/2) \quad (8)$$

Note that the reconnection efficiency η is very unlikely to be a constant (*Borovsky and Birn*, 2014), rather it is likely to show some dependence on the interplanetary parameters. For example, increased solar wind dynamic pressure may increase the magnetic shear across the relevant current shear and various factors may vary the fraction of the dayside magnetopause covered by the magnetopause reconnection X-line (or X-lines) (*Walsh, et al*., 2017). Hence, we might expect the optimum exponents for B, ρ_{SW} and V_{SW} to differ somewhat from the 1, −1/6 and 2/3, respectively, predicted by Equation (8).

1-iii. Coupling functions from empirical fits

Like many of the papers listed in Table 2, we here make empirical fits using a general form of coupling function C_f, given by

$$C_f = B_\perp^a \rho_{sw}^b V_{sw}^c \sin^d(\theta/2) \qquad (9)$$

This general form which can reproduce P_α (for $a = 2\alpha$, $b = 2/3-\alpha$, and $c = 7/3- 2\alpha$), E_{SW} (for $a = 1$, $b = 0$ and $c = 1$) and Φ_{SW} (for $a = 1$, $b = -1/6$, and $c = 2/3$). As shown by Table 1, this form also encompasses a wide variety of the proposed empirical coupling functions. Note that this form could also reproduce the often-used "epsilon" factor, ε, (for which $a = 2$, $b = 0$ and $c = 1$) but that is not considered further in this paper because ε is based on the incorrect assumption that the relevant energy flux in the solar wind is the Poynting flux (see *Lockwood*, 2013; 2019) and, although this can be made consistent with other energy coupling functions such as P_α (that is correctly based on the dominant solar wind kinetic energy flux) this is only achieved using an extreme value of unity for the coupling exponent α, and this does not agree at all with experimental estimates. This is the reason why ε performs considerably less well than P_α on all averaging timescales (see *Finch & Lockwood*, 2007).

It should be noted that not all proposed coupling functions, not even all the simple ones, fit the general formulation given in Equation (9), particularly those that employ additive terms. For example, *Boyle et al* (1977) propose the use of $10^{-4}V_{SW}^2 + 11.7B \sin^3(\theta/2)$ to predict Φ_{PC}, which it does exceptionally well: the reasons for its success will be analyzed later in this paper. In general, the problem with additive terms is that, unless each term is describing a distinct physical mechanism, they are purely numerical fits to the available data. Adding terms until a fit is achieved without a theoretical basis does makes the risk of overfitting considerably greater: essentially one can fit any time series with combinations of other time series if one is free to select enough of them until a fit is obtained. Physics-based coupling functions are usually fundamentally multiplicative in form although some factors can be broken down into the sums of additive terms for theoretical reasons (e.g., *Borovsky*, 2013; *Lockwood*, 2019; *Newell et al*, 2008).

The next section describes how there are a number of procedural issues to resolve for studies using even the relatively simple form of coupling function generalized by Equation (9). For this reason, in the present paper we do not extend the present study to formulations involving additive terms.

1-iv. Frequently neglected factors in deriving coupling functions

There are a number of factors that have often been neglected when deriving coupling functions, the most important being: (i) the effect of data gaps; (ii) the effects of data averaging; (iii) the effect of the number of datapoints available; (iv) the differences between the various terrestrial space weather indicators; (v) overfitting; (vi) non-linearity and pre-conditioning of the magnetosphere; (vi) other sources of noise such as measurement errors, propagation lags, spatial structure in interplanetary space (which can mean that the solar wind hitting Earth differs from that measured at the upstream spacecraft), seasonal and other dipole tilt effects. We address just some of these in this paper. The effect of data gaps was studied by *Lockwood et al.* (2019a) who introduced synthetic gaps at random (but to give the same distribution of durations as has occurred for early interplanetary observations) into continuous and near-continuous data and studied the errors introduced. These errors were not only in the greater uncertainty of one individual fit, but also in systematic deviations in the means and modes of the distributions of ensembles of many fits. It is often assumed that the effect of data gaps averages out, but this is not the case: data gaps introduce noise into the correlation studies and fitting procedures which generate both random and systematic errors.

Correlations of coupling functions with terrestrial space weather indicators naturally increase with increased averaging timescale τ because the noise in both time series is increasingly averaged out (*Finch and Lockwood*, 2007). However, there are problems associated with averaging high-resolution interplanetary field data in relation to the IMF orientation and these are often not addressed. *McPherron et al* (2015) correctly used hourly data which they obtained by passing 1-minute data through low-pass filter by taking a 61-point running average and resampled every hour to obtain centered hourly averages. They note that this improves the hourly-average coupling functions by eliminating nonlinearities resulting from the use of hourly averages of IMF components in calculating the transverse component $B_\perp$ and the clock angle θ. This is certainly true and in the next section we investigate how good this procedure is and why it is needed. We also point out there is a second issue to consider about the effects of data averaging.

1-v. The effect of averaging procedure

The magnetosphere responds to integrated forcing (*Lockwood et al.*, 2016). For example, if we have a terrestrial indicator that responds to the energy input into the magnetosphere and a coupling function that quantifies that energy input, over a period τ we require the total of that energy input. Similarly, for any empirical coupling function C_f (equation 9) we want the integrated solar wind forcing over the time. By the definition of the arithmetic mean, this means we need a coupling function for the interval τ given by

$$(1/\tau)\int_0^\tau C_f\,dt \;=< C_f >_\tau \;=< B_\perp^a\,\rho_{sw}^b V_{sw}^{\,c}\,sin^d(\theta/2)\;>_\tau \qquad (10)$$

Where the values C_f, $B_\perp$, ρ_{SW}, V_{SW} and θ are all values from high-time resolution measurements. However, this has usually in the past been approximated using the seemingly similar value

$$[C_f]_\tau \;=< B_\perp >_\tau^{\,a}.< \rho_{sw} >_\tau^{\,b}.< V_{sw}\; >_\tau^{\,c}.< sin(\theta/2) >_\tau^{\,d} \qquad (11)$$

And in many cases the average clock angle has been computed from the means of the IMF *Y* and *Z* components so $[\theta]_\tau$ is used for θ and $< B_\perp >_\tau$ is replaced by $[B_\perp]_\tau$, where

$$[\theta]_\tau \;=\; tan^{-1}\,(|< B_Y >_\tau|/< B_z >_\tau) \qquad (12)$$

as is the transverse IMF component

$$[B_\perp]_\tau \;=\; (< B_Z >_\tau^2 + < B_Y >_\tau^2)^{1/2} \qquad (13)$$

This generates two separate problems, the first of which was addressed by the averaging procedure for $B_\perp$ and θ that was adopted by *McPherron et al.* (2015) who evaluated both at high time resolution before averaging and avoided using wither $[\theta]_\tau$ and $[B_\perp]_\tau$ (this is hereafter referred to as the *MEA15* procedure and is what we will use in later sections). In Figure 1 we highlight its importance but also deconvolve it from a second effect. Note that same operations are used in generating $< C_f >_\tau$ and $[\,C_f\,]_\tau$ and the difference between the two is purely the order in which they are carried out: $< C_f >_\tau$ can be characterized as the parameters being "combined-then-averaged" , whereas for $[\,C_f\,]_\tau$ they are "averaged-then-combined".

Figure 1a demonstrates that it is not a valid assumption to take $< C_f >_\tau$ and $[C_f]_\tau$ to be the same, especially when using $[\theta]_\tau$ and $[B_\perp]_\tau$ (i.e., not using the *MEA15* procedure). Figure 1 is for the example of the *Vasyliunas et al.* (1982) energy transfer coupling function P_α for a coupling exponent $\alpha = 1/3$ (hence this P_α is an example of C_f with $a = 2/3$, $b = 1/3$, $c = 5/3$ and we here use $d = 4$). The raw data in Figure 1 are all the 9,930,183 valid 1-minute resolution values of P_α and all the 11,646,678 valid 1-minute resolution values of the IMF clock angle θ and tangential field $B_\perp$ available from the Omni2 dataset for 1995-2020, inclusive (*King and Papitashvili*, 2005). This interval is used because data gaps are both much rarer and shorter than before 1995 because of the advent of the Wind, Advanced Composition Explorer (ACE) and Deep Space Climate Observatory (DSCOVR) spacecraft (*Lockwood et al.,* 2019a). The averaging time in this example is $\tau = 1$ hr. Figure 1a compares $< P_\alpha >_\tau$ and $[P_\alpha]_\tau$ and the linear correlation coefficient between the two is very poor indeed, being just 0.26. Note in Figure 1a both $< P_\alpha >_\tau$ and $[P_\alpha]_\tau$ have been normalized by dividing by P_o, the overall mean of P_α: this has the advantage of cancelling out all the constants in the theoretical derivation of P_α. Rather than presenting scatter plots with massively overplotted points, Figure 1 employs data density plots with the fraction of samples, $n/\Sigma n$, color-coded with n being the number of sample pairs in small bins. In Figure 1a there are 100 bins of width 0.08 for both axes. Figure 1b identifies why the agreement in Figure 1a is so poor: it is for G, which is C_f (in this case is P_α) without the IMF orientation term, i.e.

$$G = C_f/F(\theta) = C_f/sin^4(\theta/2) = B_\perp^a \, \rho_{sw}^b \, V_{sw}^c \qquad (14)$$

This is a factor that we will use again later in deriving optimum values for d. Figure 1b compares the combine-then-average values and the average-the-combine values for G (for the same example as shown in Figure 1a and in the same format), $< G >_\tau$, with a corresponding average-then-combine value $[G]_\tau = <B_\perp>^a <\rho_{SW}>^b <V_{SW}>^c$: again, all values have been normalized by dividing by the overall mean, G_o. Note that we here use $<B_\perp>^a$ and not $[B_\perp]_\tau^a$ (where $[B_\perp]_\tau$ is defined by Equation 13) – in other words we have moved to the *MEA15* procedure in order to remove the component-averaging effect on $B_\perp$ (and θ is not a factor in

G). The agreement is here is very good indeed, with values close to the diagonal line.

However, the agreement in Figure 1b is still not quite perfect. Small differences remain because of the difference between "Hölder means" (or a "power means") $[<X^p>_\tau]^{1/p}$ of a general variable X and the corresponding arithmetic means $<X>_\tau$ and hence between $<X^p>_\tau$ and $<X>_\tau^p$. Figure 1b shows these differences are very small indeed for the variables X, the exponents p and the timescales τ involved in G and can be neglected. However, in general, arithmetic and Hölder means are related by what is called the "Hölder path" which results in the Hölder mean increasing with p (the arithmetic mean being the Hölder mean for the special case of $p = 1$). From comparison of Figures 1a and 1b, we know that the poor correlation in Figure 1a must be arising from the IMF orientation term, $F(\theta)= sin^4(\theta/2)$ and/or not using the *MEA15* procedure to averaging of $B_\perp$. Figure 1c compares the combine-then-average values of the clock angle θ, $<\theta>_\tau$ with the average-then-combine value, $[\theta]_\tau$, given by equation (12), in the same format as Figure 1a (for bins of 2°×2°) and although a great many points line up along the diagonal, there is considerable spread, especially at θ near zero or 180° (strongly northward and strongly southward IMF, respectively). Figure 1d makes the same comparison for the transverse field estimate, $B_\perp$. Note that if we use the IMF magnitude B instead of $B_\perp$ in the coupling function, this effect does not arise; however, as found by *Vasyliunas et al* (1982), tests show that using $B_\perp$ usually results in somewhat higher correlations. Figure 1e is for the same comparison for $sin^4(\theta/2)$ and the spread is greatest at the southward IMF end of the range.

Figure 1f demonstrates that the *MEA15* averaging essentially removes all problems associated with $B_\perp$ by avoiding $[B_\perp]_\tau$. However, Figure 1g shows that a problem still remains with the clock angle term $sin^4(\theta/2)$. This is because the arithmetic and Hölder means are appreciably different for this parameter. There is still a good correlation in Figure 1g and many of the points line up along the ideal diagonal: hence it is tempting to say this is just one more (small) source of noise and so it is valid to use $<sin(\theta/2)>^d$ instead of $<sin^d(\theta/2)>$. However, there is a subtle point here: the spread shown in Figure 1g increases with d because the difference between arithmetic and power means increases with exponent. Hence using $<sin(\theta/2)>^d$ discriminates against higher d by introducing more noise and so such studies will tend to derive a value for d that is too low.

We can understand why the IMF orientation term is so different to the other three by looking the variability of the various factors within the averaging period. *Lockwood et al.* (2019a) showed that the autocorrelation time of the IMF orientation is considerably shorter than for the other parameters and so most of the variability of P_α on sub-hour timescales originates from the IMF orientation term. This is true for all coupling functions. If a parameter X is constant over the averaging time, then both the Hölder mean $[<X^p>_\tau]^{1/p}$ and the arithmetic mean are equal to that constant value of X and $<X^p>_\tau = <X>_\tau^p$. On the other hand, if X varies a great deal during the averaging interval, then the Hölder mean is greater/smaller than the arithmetic mean for p greater/smaller than unity. Hence the much greater variability in the IMF orientation is the reason why it behaves so differently. (However, note that if we increase the averaging timescale τ, the other parameters will also start to suffer from the same problem as the clock angle term).

We can conclude, the often-used average-then-combine procedure generates large errors for the IMF orientation terms in deriving an empirical coupling function C_f, even for $\tau = 1$ hr. The *MEA15* averaging procedure removes a great deal of the problem (at last for $\tau = 1$ hr), but a second error remains for the clock angle term. This generates a problem when using an iterative procedure, such as the Nelder-Mead simplex search method used here (*Nelder and Mead*, 1965; *Lagarias et al.*, 1998) to fit the exponents a, b, c and d. This is because of the need to compute the mean of the combination of the samples (and in the dataset used in Figure 1 there are 9,930,183 valid 1-minute samples of P_α) at the start of every round of the iteration. We have achieved this in some cases, but it takes enormous amounts of computer time and sometimes fails to converge. Fortunately, Figure 1 points to a compromise. It suggests we can use a hybrid approach of using $<B_\perp>^a$, $<\rho_{SW}>^b$, and $<V_{SW}>^c$, but must use $<sin^d(\theta/2)>$ for the IMF orientation term. This yields a mean coupling function for averaging time τ of

$$[C'_f]_\tau = <B_\perp>_\tau^a . <\rho_{sw}>_\tau^b . <V_{sw}>_\tau^c . <sin^d(\theta/2)>_\tau \qquad (15)$$

Figure 1h compares $<P_\alpha>_\tau$ and $[P'_\alpha]_\tau$ and it shows that agreement is very good with all points lying close to the diagonal line and the correlation coefficient is 0.997. We have repeated this test for all permutations of the maximum and minimum estimates of the exponents a, b, c and d derived here and it is always valid to this level for $\tau = 1$ hr. Equation (15) is practical for use in an iterative fit procedure because for a given d we can compute $<B_\perp>_\tau$, $<\rho_{SW}>_\tau$, $< V_{SW}>_\tau$, and $< sin^d(\theta/2)>_\tau$ just once before each iteration and then readily iterate a, b, and c to the optimum fit using the Nelder-Mead simplex search. This can then be repeated for different values of d. We have carried out some sample tests of our analysis that compared the results of fits using the ideal mean $< C_f >_\tau$ and our pragmatic hybrid solution, $[C'_f]_\tau$ and the results were almost identical. However, we were limited in the number of these tests that we could carry out by the extremely large compute time caused by the need to average the whole dataset at each iteration step to define the exponents when using $< C_f >_\tau$. We have repeated all calculations using the average-than-combine procedure, $[C_f]_\tau$ (but using the *MEA11* procedure for $B_\perp$ and θ to avoid $[\theta]_\tau$ and $[B_\perp]_\tau$) and, as described below, the fits obtained were always poorer because of the effect highlighted in Figure 1g.

2. Data Employed

We use the dataset of hourly mean transpolar voltage Φ_{PC} observed over the years 1995-2020 (inclusive) by the northern-hemisphere SuperDARN array of coherent-scatter HF radars, as described by *Lockwood and McWilliams* (2021). These hourly data are means of 30, 2-minute integrations. We adopt the requirement that the hourly mean of the number of radar echoes available, n_e, exceeds a minimum value $n_{lim} = 255$. This threshold was adopted by *Lockwood and McWilliams* (2021) as a compromise between having enough echoes that the influence of the model used in the “map-potential” data-assimilation technique is small, but not so large that the distribution of Φ_{PC} values is greatly distorted by the loss of low-flow, low-n_e samples. *Lockwood and McWilliams* (2021) also found that this threshold gave peak correlation between the radar Φ_{PC} estimates and those from nearby passes of low altitude polar-orbiting spacecraft. The condition that $n_e > n_{lim} = 255$ yields a total of 65,133 Φ_{PC} samples in the dataset.

We wish to compare the optimum coupling function for the global parameter Φ_{PC} with that for global geomagnetic activity. We here use the *am* geomagnetic index (*Mayaud*, 1980). This has the most uniform network, in both hemispheres, of observing stations and uses weighting functions to yield the most uniform response possible to solar wind forcing with Universal Time and time of year (*Lockwood et al*., 2019c). The *am* index is based on the range of variation of the horizontal field component in 3-hour windows. To get a data series that is simultaneous with the Φ_{PC} data, we here linearly interpolate the 3-hourly *am* values to the mid-points of the hours used to generate the Φ_{PC} data. This is only done for the Φ_{PC} samples that meet the $n_e > n_{lim} = 255$ criterion and so we end up with a dataset of 65,133 interpolated *am* samples that are simultaneous with the Φ_{PC} data. The advantage of using *am* is that it is the geomagnetic index that is by far the most free of seasonal and hemispheric effects which introduce noise in correlation studies, and it is genuinely global. The disadvantage is that it is 3-hourly and the interpolated values will reflect this timescale.

To derive the coupling functions, we use 1-minute resolution averages of the Omni dataset of near-Earth measurements of interplanetary space (*King and Papitashvili*, 2005). From this we generate running means using one-hour (61-point) boxcar averages of $B_\perp$, ρ_{SW}, V_{SW}, and $sin^d(\theta/2)$ for the value of d we are investigating (the using the *MEA15* averaging procedure). Mean values are only considered valid when the number of samples is large enough to make the error in the mean less than 5%, thresholds that were determined by *Lockwood et al.* (2019a) for each parameter by the random removal of 1-minute samples from hourly intervals for which all 60 samples were available: because of its very low acfs, the most stringent requirement is set by the IMF orientation factor which requires 82% of samples (i.e., 43 out of the 60). The averaging generates a sequence of hourly running means that are 1 min apart. We combine these into mean coupling function $[C'_f]_{1hr}$ using our hybrid averaging formula (Equation 15). For test purposes only we also generate $[C_f]_{1hr}$ using the average-then-combine procedure (equation 11, with *MEA15* averaging to generate hourly means of θ and $B_\perp$). We then select the value at each time of the transpolar voltage and *am* dataset, allowing for the appropriate propagation lag, δt_p.

To determine the required propagation lags we make the initial assumption that the IMF orientation factor is $sin^3(\theta/2)$ (i.e., $d = 3$), although this is refined in Section 3 of this paper. We have carried out a sensitivity test to show that this choice does not influence the optimum derived lags. The Omni data have been propagated from the point of observation to the nose of the magnetosphere (*King and Papitashvili*, 2005): any variable error in that propagation will be a source of noise in our correlation studies. We then add a lag δt to allow for propagation across the magnetosheath to the dayside magnetopause and then to the relevant part of the ionosphere. We then vary δt between −60 min (unphysical) and +120 min and for each lag evaluate the linear correlation coefficients between Φ_{PC} and *am* and the optimum coupling function, C_f (for the assumed value for *d* of 3). Note that here and hereafter we refer to the hourly coupling function generated by our hybrid averaging procedure, $[C'_f]_{1hr}$ as just C_f, unless we are making a comparison with the results of the often-used average-then-combine procedure, in which case we distinguish between $[C'_f]_{1hr}$ and $[C_f]_{1hr}$. We want C_f to be linearly related to the terrestrial activity indicator and so we maximise the linear correlation coefficient, *r*. The exponents *a*, *b*, and *c* at each δt were determined using the Nelder-Mead simplex method to minimize (1-*r*) (*Nelder and Mead*, 1965; *Lagarias et al.*, 1998). From this the optimum exponents *a*, *b*, and *c* (for the assumed *d* = 3) and the correlation coefficient *r* were determined at each lag δt.

The lag correlograms, $r(\delta t)$ obtained this way are shown in the top panel of Figure 2: mauve is for Φ_{PC} and blue is for *am*. The vertical dashed lines mark the lags δt_p giving peak correlation. The bottom panel shows the best-fit exponents *a*, *b*, and *c* as a function of lag δt: it can be seen that they do vary somewhat with δt but only to a small extent around the optimum lags. δt_p. From Figure 2, we determine the optimum lags are δt_p = 18.5 min for Φ_{PC} and δt_p = 30.5 min for *am*. Note the much greater persistence in the plot for *am*, because of it is interpolated from 3-hourly data, and this makes the peak for *am* lower and broader. The survey of the Φ_{PC} dataset by *Lockwood and McWilliams* (2021) demonstrates how Φ_{PC} responds to both the reconnection rate at the dayside magnetopause Φ_D and reconnection in the cross-tail current sheet tail Φ_N (a good proxy for which is the *AL* auroral electrojet index), as predicted by the ECPC model (*Lockwood*, 1991; *Cowley and Lockwood*, 1992). Indeed, in the approximation that the polar cap remains circular at all times, Φ_{PC} is the average of Φ_D and Φ_N (*Lockwood, 1991). Lockwood and McWilliams* (2021) show that for low $-AL$, the lag of Φ_{PC} after solar wind forcing is about 5 min, which is consistent with the expected response delay of Φ_D, but the lag of the *AL* response (and hence inferred Φ_N) is 35 min, similar to the lag for *am* that is derived here. Hence we would expect the average lag for Φ_{PC}, which is generated by a combination of Φ_D and Φ_N, to be around 20 min., as is indeed found to be the case in Figure 2. However, we note that there is considerable variability in the lags connected with Φ_N, partly because of the variability in substorm growth phase duration (*Freeman and Morley*, 2004; *Li et al*., 2013) but also because, depending on the onset location, the precipitation in the initial part of the expansion phase can suppress ionospheric flow by enhancing conductivity, giving an addition delay in the appearance of the full voltage due to Φ_N (*Grocott et al*., 2009).

The optimum coupling exponents at these lags are a = 0.672, b = 0.017 and c = 0.561 for Φ_{PC} and a = 0.802, b = 0.360 and c = 2.566 for *am* (for this d of 3). The uncertainties in these values and their dependence on d will be evaluated later. The gray areas in Figure 2 define the 1-σ, 2-σ and 3-σ uncertainties in the δt_p estimates. These are evaluated by looking at the significance S of the difference between the correlation at a general lag $r(\delta t)$ and its peak value at the optimum lag δt_p (where $r = r_p$) where $S = 1-p$, and p is the probability of the null hypothesis that r and r_p are the same. S is computed using the Meng-Z test (*Meng et al.,* 1992) for the significance of the difference between correlation r_{AB} (between two variables A and B) and r_{AC} (between A and C) allowing for the fact that B and C may be correlated ($|r_{BC}| > 0$). S is, by definition, zero at the optimum lag δt_p, and the 1-σ, 2-σ and 3-σ uncertainties are the lags at which S has risen to 0.68, 0.95 and 0.997, respectively. For Φ_{PC} the 2-σ uncertainty band is between 17.2 min. and 19.8 min.; for *am* it is between 26.5 min. and 34.5 min. Note that these uncertainties are small because the number of samples is so large. Because Figure 2 was generated using an assumed value of d = 3, it was repeated for a range of selected values of d between 1 and 7 (which section 3-ii shows covers the range of interest), the differences between the derived optimum lags were always smaller than the above 2-σ uncertainties.

3. The IMF orientation factor

As discussed by *Vasyliunas et al*. (1982), the optimum IMF orientation factor is not independent of the other fit exponents. In addition, Section 1-v has described how, because its much greater rapid variability, we have to deal with it differently when generating average coupling functions. Section 3-i discusses the effect of the distribution of IMF orientation factors before in Section 3-ii we evaluate the optimum values of d for Φ_{PC} and for *am*.

3-i. Occurrence distributions of IMF orientation factors and the effect of averaging timescale

Figure 3 shows the distributions of various parameters relevant to the IMF orientation factor, all panels being for 1-minute integrations of data and in the Geocentric Solar Magnetospheric (GSM) frame of reference. This Figure is for 11,646,678 1-minute Omni data samples from 1995-2020, inclusive. The vertical axis is the fraction of samples $n/\Sigma n$ in 100 bins of width that are 1% of the range of the horizontal axis. The sequence of Figures 3a-3e are from *Lockwood et al.* (2019b) and explain how strange, highly-asymmetric distributions of 1-minute samples of the various coupling functions come about from a near-Gaussian distribution of the IMF B_Z component, which is very close to symmetric around zero, and a double-peaked distribution of the IMF B_Y component, which is also very close to symmetric around zero. As discussed above, the most commonly-adopted form of the IMF orientation factor has been $sin^d(\theta/2)$ with $d = 4$ although a range of d from 1 to 6 has been proposed. Figure 3f shows that $d = 2$ yields a symmetric distribution around an average of 0.5 with dominant isolated peaks in the bins closest to 0 and 1. On the other hand, Figure 3(g) shows that $d = 4$ yields a highly asymmetric distribution with an even-larger isolated peak in the bin nearest 0 and only a very small one in the bin nearest 1. The peak in the lowest bin is even larger for $d = 6$, shown in mauve in Figure 3(h) and larger again for two other commonly used IMF orientation factors B_S in green (where $B_S = -B_Z$ for $B_Z < 0$ and $B_S = 0$ for $B_Z \geq 0$) and $U(\theta)cos(\theta)$ in blue (where $U(\theta) = 0$ for $\theta < 90°$ and $U(\theta) = -1$ for $\theta \geq 90°$). The distributions for B_S and $U(\theta)cos(\theta)$ are very similar because $U(\theta)cos(\theta) = B_S/B$ and the factor 4.5 is used to display B_S on the same scale in Figure 3h because it makes the mean value the same as for $U(\theta)cos(\theta)$ and very similar to that for $sin^6(\theta/2)$.

These strange distributions of IMF orientation factors have great significance for statistical studies of the performance of a proposed coupling function because they determine the weighting given to a given clock angle θ in a correlation study. This means that when we alter d, we are not just investigating the how the IMF orientation influences solar wind-magnetosphere coupling, we are also changing the weighting given to certain IMF orientations in our correlation studies. For B_S and $U(\theta)cos(\theta)$ the value is zero for 50% of the dataset (for $B_Z > 0$) and so the coupling function is strongly weighted to accurate prediction of quiet times, which is probably not what is wanted in many applications. Figure

3h shows the distribution is not quite so extreme for $sin^6(\theta/2)$, but it has the same basic form. As we reduce d , that weighting shifts until for d = 2 the distribution is dominated by two equal peaks close to due northward and close to due southward IMF. For d = 1 (Figure 3e) it is dominated by close to purely southward IMF. The key point is that the choice of the IMF orientation factor is also setting the weighting given to certain data in the statistical fit of the coupling function if we use a fit metric such as correlation coefficient or root-mean-square deviation.

Figure 1 of *Lockwood et al.* (2019a) shows why the IMF orientation factor has a key role in setting the variability of a coupling function. It is because its autocorrelation function (acf) falls much more rapidly with time lag for any other solar wind parameter. For a lag of 1 hour, the acf for $sin^4(\theta/2)$ in near-Earth space is 0.45, whereas for the solar wind number density N_{SW} it is 0.88, for the IMF B it is 0.93 and for the solar wind speed V_{SW} is 0.99. Hence short-term variability of a coupling function is set by that in the IMF orientation factor whereas, as shown below, this factor essentially becomes constant at timescales of a year or more. This exemplifies the general fact that the IMF orientation factor distribution depends critically on averaging timescale which is here illustrated by Figure 4 for the commonly adopted $sin^4(\theta/2)$ factor. We take running boxcar (running) means of the 1-minute data over intervals τ and deal with data gaps by only retaining averages that are made up of a fraction of the potential maximum number samples that exceeds $f(\tau)$, the minimum needed to keep errors due to data gaps below 5%. The minimum fractions $f(\tau)$ needed were computed by introducing random synthetic data gaps into continuous IMF data, computing the error caused and repeating 1000 times, as carried out for τ = 1hr by *Lockwood et al.* (2019a). For example, Figure 1b of *Lockwood et al.* (2019a) shows that we require $f(\tau) > 0.82$ to keep errors in the hourly mean IMF orientation factor to below 5%. At very large τ it becomes very hard to find intervals with no data gaps; however, $f(\tau)$ falls with τ and so for τ > 1 day we use the $f(\tau)$ value for 1 day.

As τ is increased, the central limit theorem (*Fischer*, 2010) applies and the distribution of any parameter narrows towards a delta function at the overall mean (i.e., the value derived for a τ equal to the duration of the whole dataset). However, because of the unusual form of the distribution at τ = 1min., the distribution for $sin^4(\theta/2)$ evolves through a series of forms and how it does so is determined by the timescales of the variability in the IMF orientation. For τ

= 15min. the distribution is quite similar to that for τ = 1 min., but the peak at $sin^4(\theta/2) = 0$ has diminished and more samples occur at larger values. For τ = 1 hr (the timescale used in this paper), this results in a near-linear distribution, but still with a pronounced peak at 0. By τ = 6 hr the distribution has evolved into very close to a lognormal form and by τ = 1 day it is close to a Gaussian form that is symmetrical about the overall mean value (the mauve vertical dashed line). Further increases in τ cause the width of the distribution about the overall mean to decrease. For τ = 1 year, the distribution is narrow and hence the IMF orientation factor can, to within a reasonably small error, be taken to be constant. This is why successful coupling functions at annual timescales usually do not contain a factor that allows for IMF orientation. Note that all parameters in a coupling function, not just the IMF orientation, follow the central limit theorem, but the other factors tend to start (for 1-minute observations) from a log-normal form and then evolve into the narrowing Gaussian and do not start from the unusual distributions for the IMF orientation factors (*Lockwood et al.,* 1999a; b).

Figure 5 is the same as Figure 4, but for another value for d that has been proposed in the literature, namely d = 2 (e.g., *Kan and Lee,* 1979; *Borovsky,* 2013. *Lyatsky et al.,* 2007). This reveals the $sin^2(\theta/2)$ has very different behavior to $sin^4(\theta/2)$. At all τ, the distribution is symmetric about 0.5 and the mean value (vertical dashed line) and the value for in-equatorial field (vertical green line) are both always at 0.5. For τ up to about 15 min., this yields a uniform distribution with $sin^2(\theta/2)$ with just small peaks at zero and unity that decay as τ is increased. This even distribution makes $sin^2(\theta/2)$ a very attractive choice if studying timescales up to about 15 min. However, for τ =1 hr and above the distribution takes on some undesirable characteristics, with most samples coming from near-in-equatorial field and fewer from the extremes near 0 and 1. As discussed below this has some consequences

In the literature values for d between 1 (*Fedder et al.,* 1991, *Borovsky* 2008) and 6 (*Temerin and Lee*, 2006; *Balikhin et al.*, 2010) have been proposed and used. From the above, the choice of IMF orientation factor and of the averaging timescale both have a subtle effect on the coupling function fitting by changing the weighting given to the data samples. The central limit theorem means that the same effect applies to other factors in the coupling function, but the effects are less marked because they do not start from as extreme a distribution for 1-min values as does the IMF orientation factor. One key insight here is that we should not expect a coupling function that works well at one timescale to be equally effective at another. Hence some of the differences between the coupling functions proposed in Table 1 will have arisen from the different averaging timescales used.

The behavior in Figures 4 and 5 is very different to that obtained by an average-then-combine procedure given by equation (12) (not shown). In these cases, the distribution tends to maintain its high-resolution form up to τ of about 1 day when it starts to narrow under the central limit theorem. However, as τ is further increased it gets noisy and the broadens again as the means of both the Y and Z components of the IMF tend to zero. The key point is that this behavior is purely an artefact of the average-then-combine procedure, and the combine-then average is what mimics the physics of the magnetosphere.

3-ii. Optimum exponent *d* of the IMF orientation factor

In section 2 we defined the optimum lags for the interplanetary data, δt_p, and found that they were not significantly influenced by the choice of the exponent d in the $sin^d(\theta/2)$ IMF orientation factor. In this section, we define the optimum d using those lags. We vary d over the full proposed range (we used values from 1 to 6.5 in steps of 0.01) and using the optimum lags δt_p, we optimized a, b and c to maximize the correlation r at each d. The results are shown for Figure 6, using the same format as Figure 2.

The top panel of Figure 6 shows that for both Φ_{PC} and *am* the correlation has a peak at low d, specifically d = 2.1 for Φ_{PC} and d = 1.3 for *am*. The bottom panel shows how the other exponents (a, b and c) depend slightly on d. Note that we have used two averaging methods to generate hourly coupling functions C_f in Figure 6: the solid lines are for our hybrid procedure $[C'_f]_{1hr}$ (Equation 15) and the dashed lines are for the average-then-combine procedure, $[C_f]_{1hr}$ (Equation 11). As expected from Figure 1g, the correlations for $[C_f]_{1hr}$ are lower than for $[C'_f]_{1hr}$. For a few sample values of d (specifically 2, 3, 4 and 6) we also repeated the computation using $<C_f>_{1hr}$ (Equation 10): in each case, iteration took over a thousand times longer than the corresponding fit using $[C'_f]_{1hr}$, but the results for a, b, c and r were all the same for $<C_f>_{1hr}$ and $[C'_f]_{1hr}$ to within the estimated uncertainties. From Figure 6a, it appears that the $sin^2(\theta/2)$ IMF orientation factor performs best for Φ_{PC} and that an even lower d is best for *am* because they yield higher correlation coefficients. However, as discussed above some of this is the favorable distribution of samples that averaging brings about and the subsequent weighting of IMF orientations in deriving the correlation coefficient. Hence there is second test that we need to carry out and that is presented in the next section.

3-ii. Test of the IMF orientation factor and linear regression coefficients

Vasyliunas et al., (1982) provide a test for the optimum form of the IMF orientation factor $F(\theta)$. This is based on the fact that we want the coupling function C_f to be linearly related to the terrestrial response at all activity levels. To evaluate this, we use the function G (i.e., C_f without the $F(\theta)$ factor, defined by Equation 14). We want C_f to vary linearly with the terrestrial index T (which is either Φ_{PC} or *am* in the current paper). Hence we want

$$T = s_T C_f + i_T = s_T GF(\theta) + i_T \qquad (16)$$

which yields a requirement that

$$F(\theta) = (1/s_T) \times (T - i_T)/G \qquad (17)$$

which we can test for. The supplementary material to *Lockwood et al* (2019b) showed that this test yields $F(\theta) = sin^4(\theta/2)$ for a T of *am* and a coupling function C_f of P_α. We here repeat that test for *am* with the optimum form of G that is derived here for *am* and also apply it to Φ_{PC}. For any proposed form of $F(\theta)$, one can divide the data up into averaging bins of $F(\theta)$ for which we evaluate the mean of $(T-i_T)/G$ for each bin, where i_T is the intercept of the linear regression fit of C_f to T. If the means for the bins $<(T-i_T)/G>$ are proportional to the means $<F(\theta)>$, then Equation (16) applies, and we know that $F(\theta)$ is of the correct form for the proposed G to give a linear coupling function. The value of the intercept i_T needed makes the best-fit linear regression pass through the origin: the slope of that regression is $(1/s_T)$, by Equation (17). Hence this procedure also yields the coefficients s_T and i_T such that the terrestrial space weather index predicted from the interplanetary data is:

$$T_{pred} = s_T C_f + i_T = s_T \{< B >^a < \rho_{sw} >^b < V_{sw} >^c < sin^d(\theta/2) >\} + i_T \qquad (18)$$

Figures 7 and 8 plot the results of this test of $F(\theta)$ for Φ_{PC} and *am*, respectively. Parts (a), (b) and (c) of Figure 7 are examples of plots of $< (\Phi_{\mathrm{PC}} - i_{\Phi})/G>$ against $< F(\theta)>$ for $F(\theta) = sin^{\mathrm{d}}(\theta/2)$ for three different values of *d*. Parts (a), (b) and (c) of Figure 8 are the corresponding plots of $< (am - i_{am})/G>$, as a function of $<F(\theta)>$. In all cases we use the derived optimum G for the value of *d* in question (i.e., using the coefficients *a*, *b* and *c* given in Figure 6). Averaging is carried out over 25 bins of $F(\theta)$ of width 0.04, covering the full range of 0 to 1. Parts (a), (b) and (c) are, in both Figures, for *d* below, equal to and above the optimum value which is derived below: they show that the best fit quadratic polynomial (the red line) and this is not linear in parts (a) or (c) of either figure (the green line gives the best linear regression). For the Parts (a) of the Figures, the coefficient of the power-2 term in the best fit quadratic polynomial is positive, whereas for the Parts (c) it is negative - i.e., the curvature of the best fit of the polynomial is in the opposite sense to part (a). For the Parts (b) of both Figures the fit is linear, and this is what makes the *d* used in these cases the optimum value as it means the coupling function is linearly related to the terrestrial index.

The derivation of the optimum value of *d* is shown in the Parts (d) of Figures 7 and 8 which plot the power-2 term coefficient in the best fit-quadratic (a_{Φ} for Φ_{PC} and a_{am} for *am*) as a function of the exponent *d* over the full range of values proposed in the literature. The uncertainty band of this coat the 1-σ, 2-σ and 3-σ levels are plotted in shades of gray in both figures (but only easily discerned in Figure 8). The optimum *d* for Φ_{PC} and *am* are the values that make a_{Φ} and a_{am} (respectively) zero, for which the variation is linear. The 1-σ, 2-σ and 3-σ uncertainties in *d* are where the edges of the uncertainty bands in a_{Φ} and a_{am} fall to zero and this yields the vertical uncertainty bands in the optimum *d* that are shown.

Figure 7 shows that the required *d* is 2.50±0.07 (at the 2-σ uncertainty level) for Φ_{PC} and Figure 8 shows that it is 3.00±0.22 for *am*. Hence the optimum IMF orientation factors for Φ_{PC} and *am* are not the same within 2-σ uncertainties. These uncertainty bands for Φ_{PC} and *am* are also both shown in Figure 6 which reveals that they do not overlap even at the 3-σ uncertainty level. Hence the value for *d* is significantly lower for Φ_{PC} than *am* and both are lower than the commonly-used $d = 4$. The value for *am* is closer to the value of 3.67 derived by *McPherron et al.* (2015) for a very large dataset of 1-hour *AL* index values using the linear prediction filter technique. Some agreement is to be expected because both *am* and *AL* are dominated by the effect of the substorm current wedge and show considerable agreement (*Adebesin*, 2016; supplementary information to *Lockwood et al.*, 2019a,). However, they are different indices, so we should not expect them to show exact agreement in the *d*. The values of s_T and i_T for the optimum *d* are given in the Parts (b) of Figures 7 and 8.

The question then arises as to why the correlations *r* at these optimum *d* are slightly lower than found at lower *d*. The answer can be found by referring back to the analysis of the $d = 2$ case and the $F(\theta) = sin^2(\theta/2)$ factor presented in Figure 5. This series of distributions shows that the dataset becomes weighted towards the middle of the range of $sin^2(\theta/2)$ values as the timescale is increased and there are fewer data constraining the large and low values. This is clearly demonstrated by the distribution for these data with $\tau = 1$hr in Figure 5c. Hence although $sin^2(\theta/2)$ gives very slightly higher r_p, it is only because the dataset becomes weighted towards the center of the distribution with weaker weighting given to the extremes of low and high $F(\theta)$. To test this conclusion, we carried out correlations where the data were divided into 25 bins of $F(\theta)$ and for each bin, samples were selected at random such that all the $F(\theta)$ bins contained the same number of samples (the number that were in the least-populated bin), thereby removing the sampling bias at the expense of losing data. The peak correlations were indeed shifted to larger *d*: in the case of Φ_{PC} the peak *r* was for $d = 2.55$ and so closely matched the value derived in Figure 7. For *am*, the peak was shifted to 2.10 but was very flat, making the peak only marginally greater than for the *d* of 3.4 found from Figure 8. These correlation tests are still not bias-free because reducing the samples to the minimum number is any one bin means that fits for some *d* have systematically higher sample numbers than others. Nevertheless, this test is enough to confirm that the choice of *d* does influence the correlation coefficients by preferentially weighting certain clock angles.

In contrast, in fitting the quadratic polynomial to the bins in parts (a), (b) and (c) of Figures 7 and 8, equal weight is given to the data points for the different $F(\theta)$ bins, despite the fact that there are different numbers of samples in those bins. Hence, unlike the correlation coefficient *r*, these fits are not influenced by the distribution of samples. Hence they provide a better test of the optimum form of $F(\theta)$ that best describes the solar-wind magnetosphere coupling than do the correlation coefficients.

From Figure 7b, for Φ_{PC} the best-fit linear regression slope is $s_\Phi = 0.223$ for B in nT, ρ_{SW} in kgm^{-3} and V_{SW} in km s^{-1} and the intercept is $i_\Phi = 10$kV. From Figure 8b, for *am* the linear regression slope is $s_{am} = 0.542\times10^{-3}$ (again for B in nT, ρ_{SW} in kgm^{-3} and V_{SW} in km s^{-1}) and the intercept is $i_{am} = 6.6$nT.

It can be seen from the bottom panel of Figure 6 that, in general, the uncertainty in d introduces only small changes in the best-fit exponents a, b and c. However, the changes across the uncertainty bands are not zero. Hence when we compute the uncertainties in a, b and c we need to fold in both the fit uncertainties at the optimum d and effect of the uncertainty in that optimum d.

4. First-order check for overfitting

We here fit with three free fit parameters (a, b and c), we are pre-determining two others (d and the optimum lag, δt_p) which can influence the results and hence, even for such a large dataset, overfitting could be a problem. We here test for that in the most straightforward way by dividing the data into just two "folds" (whilst noting that machine-learning techniques often use several more folds for different tasks) and then fitting to the one half and the testing against the independent second half. Note also that testing also raises another set of complications with a variety of performance metrics available for consideration (*Liemohn et al.*, 2018), and the most appropriate one (or ones) for the application in question should be deployed, especially in the context of forecasting (*Owens*, 2018).

We here use the optimum lags δt_p and d exponents derived above and consider only linear correlation coefficient and root mean square (rms) error as test metrics. The results are demonstrated in Figures 9 and 10. The fit dataset used to define exponents a, b and c was for 2012-2019, inclusive and the resulting values are given in Table 2. The same exponents and regression coefficients (s_Φ and i_Φ for Φ_{PC} and s_{am} and i_{am} for *am*) were then applied to generate the predicted values (Φ_{pred} for Φ_{PC} and am_{pred} for *am*) for both the fit and the test subsets (1995-2011) using Equation (18). The test and fit intervals were chosen to give roughly equal numbers of samples. Because there are so many datapoints, information is lost in a scatter plot because so many points are overplotted: Figures 9 and 10 are therefore presented as datapoint density plots which give the fraction of samples (colour coded on a logarithmic scale) in bins of $\Phi_{\rm pred}$ and $\Phi_{\rm PC}$ in Figures 9a and 10a that are 1kV by 1kV in size and bins of $am_{\rm pred}$ and *am* in Figures 9b and 10b that are 1nT by 1nT in size. Comparing Figures 9 and 10 there are no obvious differences in behavior, which is quantified by the correlation coefficients *r* and the rms deviations Δ between observed and predicted values. For $\Phi_{\rm pred}$ and $\Phi_{\rm PC}$, *r* is 0.856 and 0.887 for the fit and test sets, respectively, and Δ is 12.8 kV and 12.3 kV. Hence by both metrics the test set is actually performing slightly better than the fit set. For $am_{\rm pred}$ and *am*, *r* is 0.812 and 0.822 for the fit and test sets, respectively, and Δ is 10.18 nT and 10.75 nT. Hence in this case the correlation is very slightly higher for the test set, but the rms deviation is slightly lower for the fit set. In both cases, the performance of the fits on the test set is essentially the same as for the fitting set and there is no doubt that the coupling functions have predictive power. Table 2 shows the best fit coupling functions are explaining 75-79% of the variation in $\Phi_{\rm PC}$ and 66-68% of the variation in interpolated *am*.

5. Estimation of uncertainties and the influence of the number of samples

Figure 11 presents distributions of fitted values of the other 3 exponents *a*, *b* and *c* for 3 subsets of the transpolar voltage data and compares them to the value for the full set of N = 65133 samples (given by the vertical dashed line in each case). The distributions are generated by taking 1000 random selections of N samples (from the total of $N_{\rm T}$ = 65133 samples with $n_{\rm e} > n_{\rm min}$ = 255 available): the values of N used were $N_{\rm T}/25$ = 2606 (on average, corresponding to 1 yr of data); $N_{\rm T}/10$ = 6513 (on average, corresponding to 2.5 yr of data) and $N_{\rm T}/2.5$ = 26503 (on average, corresponding to 10 yr of data). The fraction of samples $n/\Sigma n$ are plotted in bins of width (1/30) of the maximum range of each exponent shown. In each case, three histograms are shown: the light grey histogram bounded by the mauve line is for $N_{\rm T}/25$ samples, the darker grey bounded by the blue line is for $N_{\rm T}/10$ and the darkest grey bounded by the black line is for $N_{\rm T}/2.5$. The distributions are generally symmetric about the optimum value for the whole dataset, but not always so for the smallest N and, as expected, they narrow down toward the value for the full dataset as N is increased. The standard deviations of the distributions are given in each case on the plot. Figure 12 shows the corresponding distributions for *am*. Distributions are broader and peaks lower for *am* than for $\Phi_{\rm PC}$, which is expected because all plots presented thus far have had greater noise and larger uncertainties for the fits to the *am* data. Figures 11 and 12 stress how much in error an individual fitted value can be if smaller datasets are used. However, that both the mean and the mode of some of the distributions are shifted from the value for the whole dataset when N is low, meaning that there are systematic errors as well as random errors when N is very low.

Figures 13 -15 present comparisons of the errors in the exponents for the case of a very small dataset of N = 255 samples (just 0.4% of the full dataset, again chosen at random) and for the full dataset of $N = N_T$ = 65133 samples. The left-hand panels are presented to stress how large uncertainties can be if datasets are small. Figures 13, 14 and 15 are for exponents a, b, and c, respectively and the left-hand panels are for the same random selection of samples in each case. In each of the Figures, one of the three exponents is assigned a fixed value and the other two are fitted using the same Nelder-Mead simplex search procedure that was used to fit all three exponents in previous plots (again, we are using the optimum d and lag δt_p defined previously). The fixed value is then varied over the range shown. In each panel, results for both Φ_{PC} (in mauve) and *am* (in blue) are presented. The top rows give the correlation coefficients, r, with peak values, r_p, shown by the vertical dashed lines. The middle panel shows the significances S (S_Φ for Φ_{PC} and S_{am} for *am*) of the difference between the correlation at a general lag δt and its peak value at the optimum exponent. As before, we evaluate S using the Meng-Z test and the 1-σ, 2-σ and 3-σ uncertainty bands around the optimum value are shaded in dark gray, lighter gray and lightest gray, respectively. The bottom panel looks at the probability that the peaks for Φ_{PC} and *am* are the same, which is $(1-S_\Phi)(1-S_{am})$, and is discussed in the next section.

Figure 14 shows the results for exponent a. As expected, the uncertainties for the full dataset (right-hand plots) are very small, but the same is not true for the small dataset. Figure 14 shows the results for exponent b and Figure 15 shows the results for exponent c.

6. Significance of the differences between fits for transpolar voltage and geomagnetic activity

A notable feature established earlier is that the optimum d for Φ_{PC} and *am* are not the same: the shaded areas of Figure 6 show that the uncertainties do not overlap for even the 3-σ level.

Using Figures 13, 14 and 15 we can similarly compare the exponents *a*, *b*, and *c* (of *B* , ρ_{SW}, and V_{SW} respectively) for Φ_{PC} and *am*.

From the bottom, right-hand panel of Figure 13, it appears that that the dependence of Φ_{PC} and *am* on *B* could be the same for $N = 255$. The 2σ uncertainty bands of *a* overlap and the peak of $(1–S_{\Phi})(1– S_{am})$ is about 0.5. In other words, from those data there was about a 50% chance that Φ_{PC} and *am* shared the same dependence on *B*. However, when we go to the full dataset, there is no detectable chance that they are the same as for $N = 65133$ the peak $(1–S_{\Phi})(1– S_{am})$ is less than the resolution of 10^{-20}. Figure 14 shows that even in a small dataset we can detect the difference in the optimum *b* (the dependence on solar wind mass density, ρ_{SW}), the peak probability that the optimum *b* is the same for Φ_{PC} and *am*, $(1–S_{\Phi})(1–S_{am})$ is smaller than 5×10^{-6} for the data subset of 522 samples so that even the 3σ uncertainty bands do not quite overlap. The maximum $(1–S_{\Phi})(1– S_{am})$ is smaller than 10^{-20} for the full dataset. Lastly Figure 16 shows that for both the subset and the full dataset the dependence of Φ_{PC} and *am* on V_{SW} is different, with peak values of $(1–S_{\Phi})(1– S_{am})$ of 5×10^{-5} and $<10^{-20}$, respectively.

The derived optimum values and 2-σ uncertainties in *a*, *b*, *c* and *d* are listed in Table 2. The uncertainties *a*, *b*, and *c* in Figures 13-15 and are combined with those caused by the uncertainties in the optimum *d* derived from Figure 6. The same procedure was followed for the fit data subset (2012-2019) shown in Figure 9 and the derived exponents and their uncertainties are also given in Table 2.

7. Discussion and Conclusions.

We have analyzed the optimum coupling functions for a dataset of 65133 hourly mean transpolar voltage estimates Φ_{PC} observed between 1995 and 2020 by the northern-hemisphere SuperDARN radar network and a matching set of fully-simultaneous *am* index values, linearly interpolated to the center times of the radar data hours from the 3-hourly index. We have fitted using a generalized mathematical function that encompasses many proposed coupling functions and have carried out only a 2-fold test for overfitting (i.e., dividing the data into a fitting and a test data set roughly equal sample sizes). In future, machine learning techniques could be used to carry out full multi-fold fitting.

Our aim in this paper has been to establish some important principles concerning how the data can be averaged and how to ensure the IMF orientation term used does not bias the data in a way that does not match the physics of solar wind-magnetosphere coupling and also to ensure that the coupling functions derived are linear predictors of Φ_{PC} and *am*.

Table 2 gives the results for fitting the whole dataset (1995-2020) and for the fit data subset (2012-2019). We find the derived exponents for the full datasets and the subset agree to within the 3-σ uncertainties in every case. Table 2 gives the 2-σ uncertainties and these overlap in all cases for exponents *a*, *c* and *d*. Although they are close for *b* they do no overlap as we would wish. It appears that the 2-σ uncertainties for *b* have been underestimated. We have searched to see if there is a source of systematic error. One possibility that we note is that for the pre-2012 data we need to do more interpolation of the solar wind mean ion mass m_{SW} than we do in the test interval when we generate 1-minute values of $\rho_{SW} = m_{SW}N_{SW}$. We will investigate this possibility in a later study.

The correlation coefficient obtained for the *am* geomagnetic activity index is $r = 0.814$ which means that the coupling function is explaining $100r^2 = 66.3\%$ of the variation in *am*. This is good but not exceptional. For example, *McPherron et al.* (2016) explained 43.7%, 61.2%, 65.6%, and 68.3% of the variance in the hourly *AL* index using, respectively, epsilon ε (*Perrault and Akasofu*, 1978), $V_{SW}B_s$, the universal coupling function (*Newell et al*., 2007) and the optimum coupling function that they had derived which was $B_\perp^{0.79} N_{SW}^{0.10} V_{SW}^{1.92} \sin^{3.67}(\theta/2)$ (i.e., $a = 0.79$, $b = 0.10$, $c = 1.92$ and $d = 3.67$). Unfortunately, *Newell et al.* (2007) did not test the 20 coupling functions they considered against the *am* index. The closest they used to *am* was the *kp* index for which the main coupling functions correlation gave $100r^2$ that ranged from 30% for ε to 58% for their universal coupling function. However, we note that there is a ±20% peak-to-peak "McIntosh" pattern in *am* caused by dipole tilt effects (*Lockwood*, 2020a) which our optimum coupling function does not attempt to allow for with a dipole tilt term. This makes predicting 66.3% of the variation in *am* without it very encouraging.

The correlation for our transpolar voltage coupling function is $r = 0.865$ which means we are predicting $100r^2 = 75\%$ of the variation in Φ_{PC}. This is as high as has any that has been reported previously and is for a much larger dataset. An early study by *Wygant et al.* (1983)

from a limited number of satellite passes explained 55% of the variation in Φ_{PC} with the coupling function $BV_{SW}\sin^4(\theta/2)$ (i.e., a = 1, b = 0, c =1, d = 4). *Mori and Koustov* (2014) surveyed the effectiveness of different coupling functions in predicting a Φ_{PC} values from 1 year of SuperDARN radar data. They found percentages of the variance explained ranging from 13% for ε in equinox up to 61% (for $B_\perp^{1/2}V_{SW}^{1/2}\sin^2(\theta/2)$; i.e., a = 0.5, b = 0, c = 0.5 and d = 2). However, the benchmark test in transpolar voltage prediction is set by the coupling function of *Boyle et al.* (1977) who reported correlations of up to 0.87, explaining 75% of the variance of Φ_{PC}, from observations from a number of Low-Earth Orbit satellites over a three-year interval. The coupling function they derived was $10^{-4}V_{SW}^2 + 11.7B\sin^3(\theta/2)$ (where V_{SW} is in kms^{-1} and B is in nT). A concern of any additive fit of this kind is that it may be open to overfitting. However, we can now check for that by testing it against the fully independent SuperDARN Φ_{PC} data used here. The correlation we obtain is r = 0.831, and so 69% of the variance in our Φ_{PC} data is explained. This is not quite as high as *Boyle et al.* (1977) reported for their fit dataset, nor quite as high as the correlation we have found here; however, it is only just smaller and it is robust and does not suffer from overfitting to any serious extent. If we take the two terms in the Boyle function separately, we find the correlation with V_{SW}^2 is very low with r = 0.2 ($100r^2$ = 4%) but that with $B\sin^3(\theta/2)$ is 0.828 ($100r^2$ = 68.6%), and so almost as good as for the combination of terms (r = 0.831, $100r^2$ = 69.1%). Hence, the key part of the Boyle et al. function has exponent $a = 1$, b =0, c = 0 and d = 3. We note that taking the dominant term in the Boyle et al function, $B\sin^3(\theta/2)$, means there were 2 free fit parameters and the variance explained was 68.6% and by adding a third free parameter *Boyle et al.* (1977) raised this to 69.1% we have used 4 free fit parameters and have raised the variance of transpolar voltage explained to 75%.

7-i. The IMF orientation factor

As shown in Table 1, exponents d of an IMF orientation factor $sin^d(\theta/2)$ of between 2 and 6 have been suggested from empirical studies and simulations with numerical global MHD models have suggested d as low as 1.5 (*Hu et al*., 2009) or even 1 (*Fedder et al.*, 1991; *Borovsky*, 2008). For both the transpolar voltage Φ_{PC} and the *am* geomagnetic index, we find that the IMF orientation factors in the coupling function for all suggested d between 1 and 6 all perform reasonably well in terms of the correlation coefficient. We find that marginally higher correlations for hourly averages for the low d exponents, the best correlations being for

Φ_{PC} at d = 2.1 and for *am* at d = 1.3. However, we have shown that the distributions mean that these low d values are favoured mainly because they weight the statistics towards near θ = 90° and against data for strongly northward IMF (θ approaching 0) and strongly southward (θ approaching 180°). The latter bias is, of course, particularly undesirable because periods of large θ drive the strong space weather which is often what we want the coupling function to predict and quantify. From the requirement of linearity across all clock angles we find the optimum exponents d are 2.50±0.07 for Φ_{PC} and 3.00±0.22 for *am*.

As shown by Table 1 a great many studies have used $sin^{d}(\theta/2)$ with d = 4 and this exponent has also been found for energy transfer across the magnetopause in MHD simulations of global energy transfer across the magnetopause (e.g., *Laitinen et al.*, 2007). However, we here find the optimum clock angle term is $sin^{2.5}(\theta/2)$ for the transpolar voltage, which is very close to the $\sin^{2.70}(\theta/2)$ found by numerical global MHD simulations of energy transfer by *Wang et al.* (2014) and the $\sin^{2.67}(\theta/2)$ found by *Newell et al.* (2007) for their coupling function aimed at predicting the magnetopause reconnection voltage. By comparison, the exponent d in the successful equation of *Boyle et al*. (1977) is d = 3 and *Mori and Koustov* (2014) found the best of the coupling functions they tested as a predictor had d = 2. Hence our value of d for Φ_{PC} is close to others and in the middle of the range of previous proposed coupling functions for predicting Φ_{PC}.

Newell et al. (2007) argued that their coupling function was universal and could be used to predict a variety of geomagnetic indices, but we here find a simultaneous dataset of the *am* index is best fitted with the significantly larger d of 3.00. This is lower than the 3.67 derived by *McPherron et al*. (2015) from hourly *AL* index data using the linear prediction filter technique. This is a valuable comparison as the *McPherron et al*. (2015) study was also using a very large hourly dataset and because both *am* and *AL* are dominated by the effect of the substorm current wedge and show considerable agreement (*Adebesin*, 2016; supplementary information to *Lockwood et al.,* 2019a). The key point we wish to stress is that our uncertainty analysis confirms that different values of d are needed for the simultaneous transpolar voltage and the *am* geomagnetic index datasets.

7-ii. Other coupling function exponents

The values of the other exponents *a*, *b* and *c* (of B , ρ_{SW}, and V_{SW} respectively) do, in general, depend on the exponent *d* used in $sin^{d}(\theta/2)$. Some empirical fit studies have derived values for *d* that are not within the optimum range derived here, and the concern is that the associated *a*, *b* or *c* have also been shifted from optimum values to compensate. However, we also stress that the best option for *d* depends on averaging timescale τ because both alter the distribution of samples and hence the weighting given to the various θ values. The effect of timescale can be seen in extreme form for annual mean data, for which the IMF orientation is almost completely irrelevant as the value of *d* falls to zero. Coupling exponents are also are very likely to depend on which terrestrial parameter the coupling function is designed to predict. Hence we do not make detailed comparisons with previous estimates for *a*, *b* and *c* as they will be the optimum values for the particular *d* that was used.

For the optimum *d* we find the best fit to the hourly transpolar voltage Φ_{PC} is given by a = 0.642±0.019, b = 0.018±0.008 and c = 0.550±0.047. These are somewhat different to the values of $a = 1$, b = –0.167, and c = 0.667 expected for the theoretical coupling function Φ_{SW} based on the interplanetary electric field (Equation 8) and the differences imply that the reconnection efficiency η has quite considerable dependencies on all three parameters. Specifically, from our results and Equation (8) η appears to vary as $B^{-0.358}$, $\rho_{SW}^{0.185}$ and $V_{SW}^{-0.117}$. Work is needed to see if these inferred external influences are consistent with the analysis of *Borovsky and Birn* (2014) who concluded that the reconnection voltage is not just a function of the interplanetary electric field which implies other factors are influencing the reconnection efficiency.

On the other hand, the optimum exponents for the *am* geomagnetic index are a = 0.802±0.022, b = 0.360±0.012 and c = 2.560±0.072. *Lockwood et al.* (2019a) found an optimum coupling exponent α = 0.44±0.04 for the estimated power input into the magnetosphere and the *am* index for the 3-hr timescale of the *am* index. From equation (4) this predicts a = 0.88±0.08, b = 0.23±0.04 and c = 1.53±0.08. Hence although *a* is within the expected range, *b* and *c* are larger than predicted by P_{α}. It is interesting to compare with the exponents a = 0.79, b = 0.10, and c = 1.92 found by *McPherron et al* (2015) from their study of 19 years' AL data as the value of *a* is very similar but, again, *b* and *c* are somewhat lower than for our study of am.

From energy coupling into the magnetosphere from numerical MHD simulations *Wang et al.* (2014) derive $a = 0.86$, $b = 0.24$ and $c = 1.47$ (with a *d* of 2.7, similar to the 3.0 found here) which is extremely close to the above exponents for P_α with $\alpha = 0.44$. Together with our results, this strongly suggest the *am* index has an additional dependence on $\rho_{SW}^{0.13}$ and $V_{SW}^{1.03}$ for a given power input into the magnetosphere. *Lockwood et al*. (2020b) find that 75% of the variation in *am* is explained by the estimated power input and that some of the remaining variance is associated with the solar wind dynamic pressure combined with the dipole tilt. They argue this is the effect of squeezing the near-Earth tail, an effect *Lockwood et al*. (2020b) show is found in both global MHD simulations and in the inference of an empirical model of the magnetopause location. If the additional dependence of $\rho_{SW}^{0.13}$ was all due to the dynamic pressure effect in squeezing the tail (i.e., factor of $P_{SW}^{0.13} = \rho_{SW}^{0.13} V_{SW}^{0.26}$), this would leave a further dependence of $V_{SW}^{0.77}$ in the results presented here that remains unexplained.

7-iii. Universality of coupling functions

We have demonstrated that although the coupling functions for Φ_{PC} and *am* could appear to have the same exponents if we use small datasets, when we use a very large one, as in this paper, the differences are shown to be highly significant and real. We have also been careful to avoid artefacts, such as those caused by averaging and overfitting. This implies that there is no such thing as a universal coupling function that can optimally predict both voltage disturbances in the magnetosphere and geomagnetic disturbances and the coupling function needs to be tailored to the terrestrial disturbance indicator of interest in each case. This has implications for how we might allow for "preconditioning" of the magnetosphere which is discussed in the next section.

7-iv. Preconditioning

One major limitation of all the coupling functions discussed in this paper is that they assume that the terrestrial space weather index predicted is determined by the prevailing near-Earth interplanetary conditions only (allowing for the required propagation lag). This means that any preconditioning of the magnetosphere-ionosphere system is neglected and will contribute to the noise in the fits. To start to make allowance for preconditioning we have to make a distinction between two types: (i) preconditioning caused by the Earth's dipole tilt; and (ii) preconditioning that depends on the prior history of the solar wind.

7-iv-i. Preconditioning by dipole tilt

Preconditioning by the dipole tilt can change the state of the magnetosphere, giving a larger or smaller response to a given solar wind forcing. This is an external factor depending on Earth's orbital characteristics which means it should be highly predictable. Studies show that genuinely global geomagnetic activity indices show a pronounced "equinoctial" (a.k.a. "Mcintosh") pattern with time-of-year and Universal Time, associated with the tilt of Earth's magnetic dipole axis (see reviews by *Lockwood et al*., 2020a; 2021). Attempts to expand the coupling function with a factor to allow for the effect of the dipole tilt were made by *Svalgaard* (1977), *Murayama et al*. (1980), and *Luo et al*. (2013) and dipole tilt effects have been included in the filters used in the linear prediction filter technique (*McPherron et al*., 2013).

However, how this should best be done does depends on the mechanism that is responsible for the effect and there are a large number of postulated mechanisms aimed at explaining the Mcintosh effect. One invokes the dipole tilt influence on ionospheric conductivities within the nightside auroral oval and postulates that the electrojet currents are weaker when conductivities caused by solar EUV radiation are low in midnight-sector auroral ovals of both hemispheres (*Lyatsky et al.,* 2001; *Newell et al.,* 2002). Other proposals invoke tilt influences on the dayside magnetopause reconnection voltage (*Crooker & Siscoe*, 1986; *Russell et al.,* 2003) or the effect of tilt on the proximity of the ring current and auroral electrojet (*Alexeev et al.,* 1996) or tilt effects on the stability of the cross-tail current sheet through its curvature (*Kivelson & Hughes*, 1990; *Danilov et al.*, 2013; *Kubyshkina et al*., 2015). All of these effects have the potential to reproduce the McIntosh dipole tilt pattern, but which if any, are effective remains a matter of debate. Recently, strong observational (*Lockwood et al.,* 2020b) and modelling (*Lockwood et al.,* 2020c) evidence argues that the amplitude of the McIntosh pattern increases with solar wind dynamic pressure, suggesting that the dipole tilt influences the degree of squeezing of the near-Earth tail by solar wind dynamic pressure. Given that dynamic pressure effects are included in most coupling functions via the ρ_{SW}, and V_{SW} terms, and that the effect is reasonably simultaneous with other solar wind effects, we might expect this effect to influence best-fit coupling exponents by raising *b* and *c* for geomagnetic activity but not for transpolar voltage. Thus, this mechanism has some relevance to understanding why the coupling function for transpolar voltage may be so different from that for the *am* index.

7-iv-ii. Preconditioning related to prior solar wind history

The storage-release system manifest in substorms shows the response of the magnetosphere is non-linear: the effect of a given burst of southward-pointing IMF, for example, is different at the start of the growth phase (when the open magnetospheric flux is flow) compared to at the end of the growth phase (when it is high). Hence the response that depends on the state of the magnetosphere is in at the time, and that is set by the prior history of solar wind magnetosphere coupling. One technique to allow for the non-linearity of response caused by this type of preconditioning is local linear prediction [*Vassiliadis et al*., 1995; *Vassiliadis, 2006*]. In this technique, moving average filters are continually calculated as the system evolves and these are used to compute the output of the system for this filter. The filter used is derived or selected according to the state of the system. Another way of dealing with this non-linearity is by using neural networks (e.g., *Gleisner and Lundstedt*, 1997). Our finding that the coupling function is significantly different for transpolar voltage and geomagnetic activity is significant in this respect. It means that if, for example, we wanted at allow for preconditioning due to the open flux in the magnetosphere, we would want to look at the prior history of an optimum coupling function for dayside reconnection voltage but would need to use a different coupling function to best predict the geomagnetic disturbance.

A number of other physical mechanisms have been proposed as ways of further preconditioning the magnetosphere. They include: mass loading of the near-Earth tail with ionospheric O^+ ions from the cleft ion fountain (*Yu and Ridley*, 2013); the formation of thin tail current sheets (*Pulkkinen and Wiltberge*r, 2000); the development of a cold dense plasma sheet (*Lavraud et al*., 2006).

Another effect could be the effect on the reconnection rate in the cross-tail current sheet of enhanced ring current, as has been proposed by *Milan et al*. (2008; 2009) and *Milan* (2009). The magnetosphere sometimes responds to continued solar wind forcing (over a period of tens of minutes) by generating a substorm, or a string of substorms and sometimes with a steady convection event (e.g., *Kissinger et al*, 2012). Studies (e.g., *Gleisner and Lundstedt*, 1999) have demonstrated that the response of the auroral electrojet indices depends on the current *Dst* value. *O'Brien et al.* (2002) studied two intervals in which the solar wind coupling function was similar, one of which resulted in an isolated substorm and the other in a steady convection event. They noted the main difference was the pre-existing state of the magnetosphere in that prior to the substorm, the magnetosphere was quiet but whereas before the steady convection event the magnetosphere was already undergoing enhanced activity. *McPherron et al.* (2005) estimate that about 80% of steady convection events are associated with a substorm onset but thereafter the magnetospheric behavior diverges. The work of *Juusola et al*. (2013) strongly suggests that enhanced ring current is the reason that a steady convection event forms as opposed to a substorm, quite possibly through the mechanism proposed by Milan and co-workers.

Hence preconditioning of the magnetosphere undoubtedly occurs through at least one mechanism, and this will be a major noise factor in the derivation of a simple coupling function and hence a major limitation on the performance of that coupling function. The problem is that not only are the effects of the various mechanisms on the response different, the time constants of the prior activity that is influencing the response will be different in each case. This means that the time profiling of any preconditioning quantification factor in a coupling function using the prior history of the interplanetary parameters will depend on the mechanism.

To underline this point about the importance of the mechanism that is causing pre-conditioning, note that some mechanisms, such as the cold dense plasma sheet, would emphasize prior periods of quiet, northward IMF conditions as giving higher activity for a given input (*Borovsky & Denton*, 2006; 2010; *Lavraud et al.*, 2006), whereas others, such as the ring current enhancement mechanism would emphasize prior periods of enhanced solar wind magnetosphere coupling. The time constants for forcing in the build-up to ring current enhancements (*Lockwood et al.*, 2016) are different to those for the development of a cold, dense plasma sheet (*Fuselier et al.*, 2015). Yet another proposed preconditioning mechanism involves the effect of solar wind dynamic pressure and thus would introduce yet another different precursor time profile (*Xie et al*., (2008). Some of these preconditioning effects have been predicted by numerical modelling (e.g.. *Lyon et al*., 1998; *Wiltberger et al*., 2000) and it is quite possible that we may need numerical simulations to isolate the preconditioning effects and determine how best to allow for them.

However, if we are to make these improvements to coupling functions to allow for preconditioning, we will need to remember that they will, inevitably, introduce more free fit parameters, making tests to guard against overfitting ever more important.

Acknowledgements. The authors acknowledge the use of data from the SuperDARN project. SuperDARN is a collection of radars funded by national scientific funding agencies of Australia, Canada, China, France, Italy, Japan, Norway, South Africa, United Kingdom and the United States of America. The work presented in this paper was supported by a number of grants. ML is supported by STFC consolidated grant number ST/M000885/1 and by the SWIGS NERC Directed Highlight Topic Grant number NE/P016928/1/. Funding for KAM at University of Saskatchewan was provided by the Canadian Foundation for Innovation (CFI), the Province of Saskatchewan, and a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada. Initial work by KAM for this paper was carried out at University of Reading on sabbatical leave from University of Saskatchewan. We thank Evan Thomas, Kevin Sterne, Simon Shepherd, Keith Kotyk, Marina Schmidt, Pasha Ponomarenko, Emma Bland, Maria-Theresia Walach, Ashton Reimer, Angeline Burrell, and Daniel Billett for the SuperDARN radar processing toolkit used to analyze the radar data. The authors are also grateful to the staff of: the Space Physics Data Facility, NASA/Goddard Space Flight Center, who prepared and made available the OMNI2 dataset used: these interplanetary data were downloaded from **http://omniweb.gsfc.nasa.gov/ow.html**; the World Data Center for Geomagnetism, Kyoto who generate and make available the *AL* index from **http://wdc.kugi.kyoto-u.ac.jp/aeasy/index.html** and the staff of L'École et Observatoire des Sciences de la Terre (EOST), a joint of the University of Strasbourg and the French National Center for Scientific Research (CNRS) and the International Service of Geomagnetic Indices (ISGI) for making the am index data available from **http://isgi.unistra.fr/data_download.php**

References

Adebesin, B. O. (2016). Investigation into the linear relationship between the AE, Dst and ap indices during different magnetic and solar activity conditions. *Acta Geodaetica et Geophysica*, **51** (2), 315–331. doi: 10.1007/s40328-015-0128-2

Alexeev, I.I., E.S. Belenkaya, V.V. Kalegaev, Y.I. Feldstein, and A. Grafe (1996), Magnetic storms and magnetotail currents, *J. Geophys. Res*., **101**, 7737–7747, doi: 10.1029/95JA03509

Arnoldy, R. (1971), Signature in the interplanetary medium for substorms, *J. Geophys. Res*., **76** (22), 5189–5201, doi:10.1029/JA076i022p05189.

Arnoldy, R. L. Bargatze, L. F., R. L. McPherron, D. N. Baker, and J. E. W. Hones (1984), The application of dimensional analysis to solar wind-magnetosphere energy coupling, *in Proceedings of Conference on Achievements of the IMS*, 26–28 June 1984, Graz, Austria, ESA SP-217, pp. 157–160, Eur. Space Agency, Paris.

Bargatze, L. F., D. N. Baker, R. L. McPherron, and E. W. Hones Jr. (1985), Magnetospheric impulse response for many levels of geomagnetic activity, *J. Geophys. Res*., **90** (A7), 6387–6394, doi:10.1029/JA090iA07p06387

Bargatze, L. F., D. N. Baker, and R. L. McPherron (1986), Solar wind-magnetosphere energy input functions, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, pp. 101–109, Terra Sci. Co., Tokyo.

Borovsky, J. E. (2008), The rudiments of a theory of solar wind/magnetosphere coupling derived from first principles, *J. Geophys. Res*., **113**, A08228, doi:10.1029/2007JA012646.

Borovsky, J. E. (2013), Physical improvements to the solar wind reconnection control function for the Earth's magnetosphere, *J. Geophys. Res. Space Physics*, **118**, 2113– 2121, doi:10.1002/jgra.50110.

Borovsky, J. E., and Birn, J. (2014) The solar wind electric field does not control the dayside reconnection rate, *J. Geophys. Res. Space Physic*s, **119**, 751– 760, doi: 10.1002/2013JA019193.

Borovsky, J. E., and M. H. Denton (2006), Differences between CME-driven storms and CIR-driven storms, *J. Geophys. Res*., **111**, A07S08, doi:10.1029/2005JA011447.

Borovsky, J. E., and M. H. Denton (2010), Magnetic field at geosynchronous orbit during high-speed stream-driven storms: Connections to the solar wind, the plasma sheet, and the outer electron radiation belt, *J. Geophys. Res*., **115**, A08217, doi:10.1029/2009JA015116.

Boyle, C. B., Reiff, P. H., and Hairston, M. R. (1997), Empirical polar cap potentials, *J. Geophys. Res.,* **102** (A1), 111– 125, doi:10.1029/96JA01742.

Burton, R. K., R. L. McPherron, and C. T. Russell (1975), An empirical relationship between interplanetary conditions and Dst, *J. Geophys. Res*., **80** (31), 4204-4214 , doi: 10.1029/ja080i031p04204

Camporeale, E. (2019). The challenge of machine learning in Space Weather: Nowcasting and forecasting. *Space Weather*, **17**, 1166– 1207. doi: 10.1029/2018SW002061

Chicco, D. (2017) Ten quick tips for machine learning in computational biology, *BioData Mining*, **10** (1), Article # 35, doi: 10.1186/s13040-017-0155-3

Cowley, S. W. H. (1991). Acceleration and heating of space plasmas—Basic concepts. *Annales Geophysicae*, **9**, 176–187.

Cowley, S.W.H. (1984) Solar wind control of magnetospheric convection, in *Achievements of the International Magnetspheric Study (IMS*), Proceedings of an International Symposium held 26-28 June, 1984 in Graz, Austria. Eds. B. Battrick and E.J. Rolfe, pp.483-494. ESA Special Publication ESA-SP-217. European Space Agency, Paris.

Cowley, S. W. H., and Lockwood, M. (1992), Excitation and decay of solar-wind driven flows in the magnetosphere-ionosphere system, *Annales Geophys*., **10**, 103-115.

Crooker N.U., and Siscoe G.L. (1986) On the limits of energy transfer through dayside merging. *J. Geophys. Res*., **91**: 13393–13397. doi: 10.1029/JA091iA12p13393.

Danilov A.A., G.F. Krymskii, and G.A. Makarov (2013) Geomagnetic Activity as a Reflection of Processes in the Magnetospheric Tail: 1. The Source of Diurnal and Semiannual Variations in Geomagnetic Activity, *Geomag. and Aeron*., **53**, (4), 441–447, doi:10.1134/S0016793213040051

Ebihara, Y., Tanaka, T., & Kamiyoshikawa, N. (2019). New diagnosis for energy flow from solar wind to ionosphere during substorm: Global MHD simulation. *Journal of Geophysical Research: Space Physics*, **124**, 360 –378. doi: 10.1029/2018JA026177

Farrugia, C.J., M.P. Freeman, S.W.H. Cowley, D.J. Southwood, M. Lockwood and A. Etemadi (1989) Pressure-driven magnetopause motions and attendant response on the ground, *Planet. Space Sci*., **37**, 589-608, doi: 10.1016/0032-0633(89)90099-8

Fedder, J. A., C. M. Mobarry, and J. G. Lyon (1991) Reconnection voltage as a function of IMF clock angle., *Geophysical Research Letters*, **18**, (6) 1047-1050. doi: 10.1029/90GL02722

Feynman, J. & Crooker, N. U., (1978) The solar wind at the turn of the century, *Nature,* **275** (5681), 626–627, doi: 10.1038/275626a0

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

Finch, I.D., M. Lockwood, A. P. Rouillard (2008) The effects of solar wind magnetosphere coupling recorded at different geomagnetic latitudes: separation of directly-driven and storage/release systems, *Geophys. Res. Lett*., 3**5**, L21105, doi:10.1029/2008GL035399, 2008

Fischer H. (2010) A History of the Central Limit Theorem. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, New York, NY. ISBN 978-0-387-87856-0. doi: 10.1007/978-0-387-87857-7

Freeman, M. P., and Morley, S. K. (2004), A minimal substorm model that explains the observed statistical distribution of times between substorms, *Geophys. Res. Lett*., **31**, L12807, doi:10.1029/2004GL019989.

Fuselier, S. A., Dayeh, M. A., Livadiotis, G., McComas, D. J., Ogasawara, K., Valek, P., Funsten, H. O., and Petrinec, S. M. (2015), Imaging the development of the cold dense plasma sheet, *Geophys. Res. Lett*., **42**, 7867– 7873, doi:10.1002/2015GL065716.

Gleisner, H., and H. Lundstedt (1999), Ring current influence on auroral electrojet predictions, *Annales Geophys*., **17** (10), 1268–1275. doi: 10.1007/s00585-999-1268-x

Grocott, A., J. A. Wild, S. E. Milan, T. K. Yeoman (2009) Superposed epoch analysis of the ionospheric convection evolution during substorms: onset latitude dependence, *Annales Geophys.*, **27** (2) 591 to 600, doi: 10.5194/angeo-27-591-2009

Hu, Y. Q., Z. Peng, C. Wang, and J. R. Kan (2009), Magnetic merging line and reconnection voltage versus IMF clock angle: Results from global MHD simulations, *J. Geophys. Res.*, **114**, A08220, doi:10.1029/2009JA014118.

Juusola, L., N. Partamies, and E. Tanskanen (2013),Effect of the ring current on preconditioning the magnetosphere for steady magnetospheric convection, *Geophys. Res. Lett.*, **40**, 1917–1921, doi:10.1002/grl.50405

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

Kissinger, J., McPherron, R. L., Hsu, T.-S., and Angelopoulos, V. (2012), Diversion of plasma due to high pressure in the inner magnetosphere during steady magnetospheric convection, *J. Geophys. Res*., **117**, A05206, doi:10.1029/2012JA017579.

Kivelson, M. G., and W. J. Hughes (1990), On the threshold for triggering substorms, *Planet. Space Sci.*, **38**, 211-220, doi: 10.1016/0032-0633(90)90085-5

Knutti, R., Meehl, G.A., Allen, M.R. & Stainforth, D.A. (2006). Constraining climate sensitivity from the seasonal cycle in surface temperature. *Journal of Climate*, **19** (17), 4224 4233. doi : 10.1175/JCLI3865.1

Knape, J., & de Valpine, P. (2011) Effects of weather and climate on the dynamics of animal population time series. *Proc. Royal Society B: Biological Sciences*, **278** (1708), 985-992, doi: 10.1098/rspb.2010.1333

Kubyshkina, M., N. Tsyganenko, V. Semenov, D. Kubyshkina, N. Partamies and E. Gordeev (2015) Further evidence for the role of magnetotail current shape in substorm initiation, *Earth, Planets and Space*, **67**, 139, doi: 10.1186/s40623-015-0304-1

Lagarias, J. C., J. A. Reeds, M. H. Wright, and P. E. Wright. (1998) Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM Journal of Optimization*. **9**, (1) 112–147, doi: 10.1137/s1052623496303470

Laitinen, T. V., M. Palmroth, T. I. Pulkkinen, P. Janhunen, and H. E. J. Koskinen (2007), Continuous reconnection line and pressure-dependent energy conversion on the magnetopause in a global MHD model, *J. Geophys. Res*., **112**, A11201, doi:10.1029/2007JA012352.

Lavraud, B., M. F. Thomsen, J. E. Borovsky, M. H. Denton, and T. I. Pulkkinen (2006), Magnetosphere preconditioning under northward IMF: Evidence from the study of coronal mass ejection and corotating interaction region geoeffectiveness, *J. Geophys. Res.,* **111**, A09208, doi:10.1029/2005JA011566.

Liemohn, M. W., McCollough, J. P., Jordanova, V. K., Ngwira, C. M., Morley, S. K., Cid, C., et al. (2018). Model evaluation guidelines for geomagnetic index predictions. *Space Weather*, **16**, 2079–2102. doi: 10.1029/2018SW002067

Li, X. L., K. S. Oh, and M. Temerin (2007), Prediction of the AL index using solar wind parameters, *J. Geophys. Res*., **112**, A06224, doi: 10.1029/2006JA011918.

Li, H., Wang, C., and Peng, Z. (2013), Solar wind impacts on growth phase duration and substorm intensity: A statistical approach, J. *Geophys. Res. Space Physics*, **118**, 4270– 4278, doi: 10.1002/jgra.50399.

Lockwood, M. (1991) On flow reversal boundaries and cross-cap potential in average models of high latitude convection, *Planet. Space Sci.,* **39**, 397-409, doi: 10.1016/0032-0633(91)90002-R

Lockwood, M. (2004) Solar Outputs, their variations and their effects of Earth in *"The Sun, Solar Analogs and the Climate", Proc. Saas-Fee Advanced Course*, 34 by J.D. Haigh, M. Lockwood and M.S. Giampapa, eds. I. Rüedi, M. Güdel, and W. Schmutz, pp107-304, Springer, ISBN: 3-540-23856-5

Lockwood, M. (2013) Reconstruction and Prediction of Variations in the Open Solar Magnetic Flux and Interplanetary Conditions, *Living Reviews in Solar Physics*, **10**, 4, 2013. doi: 10.12942/lrsp-2013-4

Lockwood, M. (2019) Does adding solar wind Poynting flux improve the optimum solar wind-magnetosphere coupling function? *J. Geophys. Res. (Space Physics),* **124** (7), 5498-5515 doi: 10.1029/2019JA026639

Lockwood, M. & K.A. McWilliams (2021) A survey of 25 years' transpolar voltage data from the SuperDARN radar network and the Expanding-Contracting Polar Cap model, *J. Geophys. Res*., accepted and in press, doi: 10.1029/2021JA029554 (2021)

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

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

Lockwood, H. Nevanlinna, L. Barnard, M.J. Owens, R.G. Harrison, A.P. Rouillard, and C.J. Scott (2014) Reconstruction of Geomagnetic Activity and Near-Earth Interplanetary Conditions over the Past 167 Years: 4. Near-Earth Solar Wind Speed, IMF, and Open Solar Flux, *Annales. Geophys*., **32**, 383-399, doi:10.5194/angeo-32-383-2014

Lockwood, M., M.J. Owens, L.A. Barnard S. Bentley, C.J. Scott, and C.E. Watt (2016) On the Origins and Timescales of Geoeffective IMF, *Space Weather*, **14**, 406 432, doi: 10.1002/2016SW001375

Lockwood, M., M.J. Owens, L.A. Barnard, C.J. Scott, and C.E. Watt (2017) Space Climate and Space Weather over the past 400 years: 1. The Power input to the Magnetosphere, *J. Space Weather Space Clim*., **7**, A25, doi: 10.1051/swsc/2017019 (2017)

Lockwood, M. S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt, and O. Allanson (2019a) The development of a space climatology: 1. Solar-wind magnetosphere coupling as a function of timescale and the effect of data gaps, *Space Weather*, **17**, 133-156. doi: 10.1029/2018SW001856

Lockwood, M., S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt, O. Allanson and M.P. Freeman (2019b) The development of a space climatology: 2. The distribution of power input into the magnetosphere on a 3-hourly timescale, *Space Weather*, **17**, 157-179. doi: 10.1029/2018SW002016

Lockwood, M., Chambodut , A., Finch, I. D., Barnard, L. A., Owens, M.J. and Haines, C. (2019c) Time-of-day / time-of-year response functions of planetary geomagnetic indices, *J. Space Weather Space Clim.*, **9**, A20, doi: 10.1051/swsc/2019017

Lockwood, M., M.J. Owens, L.A. Barnard, C. Haines, C.JV Scott, K.A. McWilliams, and J.C. Coxon (2020a) Semi-annual, annual and Universal Time variations in the magnetosphere and in geomagnetic activity: 1. Geomagnetic data, *J. Space Weather Space Clim*., **10**, 23, doi: 10.1051/swsc/2020023

Lockwood, M., K.A. McWilliams, M.J. Owens, L.A. Barnard, C.E. Watt, C.J. Scott, A. McNeill and J.C. Coxon (2020b) Semi-annual, annual and Universal Time variations in the magnetosphere and in geomagnetic activity: 2. Response to solar wind power input and relationships with solar wind dynamic pressure and magnetospheric flux transport, *J. Space Weather Space Clim*., **10**, 30, doi: 10.1051/swsc/2020033

Lockwood, M.J. Owens, L.A. Barnard, C.E. Watt, C.J. Scott, J.C. Coxon and K.A. McWilliams (2020c) Semi-annual, annual and Universal Time variations in the magnetosphere and in geomagnetic activity: 3. Modelling, *J. Space Weather and Space Climate*, **10**, 61 doi: 10.1051/swsc/2020062

Lockwood, C.A. Haines, L.A. Barnard, J. Owens, C.J. Scott, A. Chambodut, and K.A. McWilliams (2021) Semi-annual, annual and Universal Time variations in the magnetosphere and in geomagnetic activity: 4. Polar Cap motions and origins of the Universal Time effect, *J. Space Weather and Space Climate*, **11**, 15, doi: 10.1051/swsc/2020077

Luo, B. X., X. L. Li, M. Temerin, and S. Q. Liu (2013), Prediction of the AU, AL, and AE indices using solar wind parameters, *J. Geophys. Res. Space Physics*, **118**, 7683–7694, doi: 10.1002/2013JA019188.

Lyatsky, W., P. T. Newell, and A. Hamza (2001), Solar illumination as the cause of the equinoctial preference for geomagnetic activity, *Geophys. Res. Lett*., **28** (12), 2353-2356, doi: 10.1029/2000GL012803

Lyatsky, W., Lyatskaya, S., and Tan, A. (2007), A coupling function for solar wind effect on geomagnetic activity, *Geophys. Res. Lett*., **34**, L02107, doi:10.1029/2006GL027666.

Lyon, J.G., R.E. Lopez, C.C. Goodrich, M. Wiltberger, K. Papadopoulos (1998) Simulation of the March 9, 1995, substorm: auroral brightening and the onset of lobe reconnection, *Geophysical Research Letters*, **25**, 3039 -3042 doi: 10.1029/98GL00662

Mayaud, P.-N. (1980), Derivation, Meaning and Use of Geomagnetic Indices, *Geophysical Monograph*, *22*, American Geophysical Union, Washington, DC. doi: 10.1029/GM022.

McPherron, R. L., D. N. Baker, T. I. Pulkkinen, T. S. Hsu, J. Kissinger, and X. Chu (2013), Changes in solar wind-magnetosphere coupling with solar cycle, season, and time relative to stream interfaces, *J. Atmos. Sol. Terr. Phys.*, **99**, 1–13, doi:10.1016/j.jastp.2012.09.003

McPherron, R. L., T.-S. Hsu, and X. Chu (2015), An optimum solar wind coupling function for the AL index, *J. Geophys. Res. Space Physics*, **120**, 2494–2515, doi: 10.1002/2014JA020619.

McPherron, R. L., T. P. O'Brien, and S. M. Thompson (2005), Solar wind drivers for steady magnetospheric convection, in *Multiscale Coupling of the Sun-Earth Processes*, edited by A. T. Y. Liu, Y. Kamide, and G. Consolini, pp. 113–124, Elsevier, Amsterdam, doi:10.1016/B978-044451881-1/50009-5

Meng, X.-I., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients, *Psychological Bulletin*, **111** (1), 172–175. doi: 10.1037//0033-2909.111.1.172

Milan, S. E., P. D. Boakes, and B. Hubert (2008), Response of the expanding/contracting polar cap to weak and strong solar wind driving: Implications for substorm onset, *J. Geophys. Res*., **113**, A09215, doi:10.1029/2008JA013340

Milan, S. E. (2009), Both solar wind-magnetosphere coupling and ring current intensity control of the size of the auroral oval, *Geophys. Res. Lett*., **36**, L18101, doi:10.1029/2009GL039997

Milan, S.E., J. Hutchinson, P. D. Boakes, and B. Hubert (2009) Influences on the radius of the auroral oval, *Annales Geophys*., **27** (7), 2913 – 2924 doi: 10.5194/angeo-27-2913-2009

Milan, S. E., J.S. Gosling, & B. Hubert (2012) Relationship between interplanetary parameters and the magnetopause reconnection rate quantified from observations of the expanding polar cap. *Journal of Geophysical Research*, **117**, A03226, doi: 10.1029/2011JA017082

Milan, S. E., Carter, J. A., Sangha, H., Bower, G. E., & Anderson, B. J. (2021). Magnetospheric flux throughput in the Dungey cycle: Identification of convection state during 2010. *Journal of Geophysical Research: Space Physics*, **126**, e2020JA028437. doi: 10.1029/2020JA028437

Mori, D. & A.V. Koustov (2013) SuperDARN cross polar cap potential dependence on the solar wind conditions and comparisons with models, *Adv. in Space Res*., **52** (6), 1155-1167, doi: 10.1016/j.asr.2013.06.019

Murayama, T. (1982), Coupling function between solar wind parameters and geomagnetic indices, *Rev. Geophys*., **20** (3), 623– 629, doi:10.1029/RG020i003p00623.

Murayama, T. (1986), Coupling function between the solar wind and the Dst index, *in Solar Wind Magnetosphere Coupling*, pp. 119–126, Terra Sci. Co. and D. Reidel Co., Tokyo, and Dordrecht, Netherlands.

Murayama, T., T. Aoki, H. Nakai, and N. Hakamada (1980), Empirical formula to relate the auroral electrojet intensity with interplanetary parameters, *Planet. Space Sci.*, **28**, 803–813, doi: 10.1016/0032-0633(80)90078-1

Nelder, JA & R. Mead (1965) A Simplex Method for Function Minimization, *The Computer Journal*, **7**, (4), 308–313. doi: 10.1093/comjnl/7.4.308 and Errata, *The Computer Journal*, 8 (1) Issue 1, 27, doi: org/10.1093/comjnl/8.1.27

Newell, P.T., T. Sotirelis, J.P. Skura, C.-I. Meng and W. Lyatsky (2002) Ultraviolet insolation drives seasonal and diurnal space weather variations, *J. Geophys. Res*., **107** (A10), 1305, doi:10.1029/2001JA000296.

Newell, P. T., Sotirelis, T., Liou, K., Meng, C.-I., and Rich, F. J. (2007), A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables, *J. Geophys. Res*., **112**, A01206, doi:10.1029/2006JA012015.

Newell, P. T., Sotirelis, T., Liou, K., and Rich, F. J. (2008), Pairs of solar wind-magnetosphere coupling functions: Combining a merging term with a viscous term works best, *J. Geophys. Res*., **113**, A04218, doi:10.1029/2007JA012825.

O'Brien, T. P., S. M. Thompson, and R. L. McPherron (2002), Steady magnetospheric convection: Statistical signatures in the solar wind and AE, *Geophys. Res. Lett.*, **29** (7), 1130-1133, doi:10.1029/2001GL014641.

Owens, M. J. (2018). Time-window approaches to space-weather forecast metrics: A solar wind case study. *Space Weather*, **16**, 1847– 1861. https://doi.org/10.1029/2018SW002059

Perreault, P.D. (1974). On the relationship between interplanetary magnetic fields and magnetospheric storms and substorms, PhD thesis, Univ. of Alaska Fairbanks, Ann Arbor. (Order No. 7910269). Available from ProQuest One Academic. (302740827). Retrieved from https://www.proquest.com/dissertations-theses/on-relationship-between-interplanetary-magnetic/docview/302740827/se-2?accountid=13460

Perreault, P.D., & Akasofu, S. I. (1978). A study of geomagnetic storms. *Geophysical Journal International*, **54**, (3), 547-573. doi: 10.1111/j.1365-246X.1978.tb05494.x

Pulkkinen, T.I. and M. Wiltberger (2000) Thin current sheet evolution as seen in observations, empirical models and MHD simulations, *Geophysical Research Letters*, **27**, 1363-1366 doi: 10.1029/1999GL003726

Reiff, P. H., and J. G. Luhmann (1986), Solar wind control of the polar-cap voltage, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, pp. 453–476, Terra Sci., Tokyo.

Rouillard, A. P., Lockwood, M., and Finch, I. D. (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.T., Wang Y.L., and Raeder J. (2003) Possible dipole tilt dependence of dayside magnetopause reconnection. Geophys. Res. Lett., 30 (18), 1937-1940, doi: 10.1029/2003GL017725.

Scurry, L., and C. T. Russell (1991), Proxy studies of energy-transfer to the magnetosphere, *J. Geophys. Res*., **96**, 9541–9548, doi:10.1029/91JA00569

Siscoe, G. L., G. M. Erickson, B. U. O. Sonnerup, N. C. Maynard, J. A., Schoendorf, K. D. Siebert, D. R. Weimer, W. W. White, and G. R. Wilson (2002), Hill model of transpolar potential saturation: Comparisons with MHD simulations, *J. Geophys. Res*., **107** (A6), 1075, doi:10.1029/2001JA000109.

Spencer, E., A. Rao, W. Horton, and M. L. Mays (2009), Evaluation of solar wind-magnetosphere coupling functions during geomagnetic storms with the WINDMI model, *J. Geophys. Res*., **114**, A02206, doi: 10.1029/2008JA013530.

Stephens, G. K., Bingham, S. T., Sitnov, M. I., Gkioulidou, M., Merkin, V. G., Korth, H., et al. (2020). Storm time plasma pressure inferred from multi-mission measurements and its validation using Van Allen Probes particle data. *Space Weather,* **18**, e2020SW002583. doi: 10.1029/2020SW002583

Svalgaard, L. (1977), Geomagnetic activity: Dependence on solar wind parameters, in *A Monograph from Skylab Solar Workshop 1, Coronal Holes and High Speed Wind Streams*, edited by J. B. Zirker, pp. 371–441, Colo. Assoc. Univ. Press, Boulder.

Svalgaard, L., Cliver, E. W., and Le Sager, P. (2003) Determination of interplanetary magnetic field strength, solar wind speed and EUV irradiance, 1890–2003, in: *Solar Variability as an Input to the Earth's Environment*, edited by: Wilson, A., ESA Special Publication, 535, 15–23, European Space Agency, Nordwijk.

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

Temerin, M., and X. Li (2006), Dst model for 1995– 2002, *J. Geophys. Res*., **111**, A04221, doi: 10.1029/2005JA011257

Vassiliadis, D. (2006), Systems theory for geospace plasma dynamics, *Rev. Geophys*., **44**, RG2002, doi:10.1029/2004RG000161.

Vassiliadis, D., A. Klimas, D. Baker, and D. Roberts (1995), A description of the solar wind-magnetosphere coupling based on nonlinear filters, *J. Geophys. Res*., **100** (A3), 3495–3512, doi: 10.1029/94JA02725.

Vasyliunas, V. M., Kan, J. R., Siscoe, G. L., & Akasofu, S.-I. (1982). Scaling relations governing magnetospheric energy transfer, *Planetary and Space Science*, **30** (4), 359–365. doi: 10.1016/0032-0633(82)90041-1

Walsh, B. M., Komar, C. M., and Pfau-Kempf, Y. (2017), Spacecraft measurements constraining the spatial extent of a magnetopause reconnection X line, *Geophys. Res. Lett*., **44**, 3038– 3046, doi: 10.1002/2017GL073379.

Wang, C., Han, J. P., Li, H., Peng, Z., and Richardson, J. D. (2014), Solar wind-magnetosphere energy coupling function fitting: Results from a global MHD simulation, *J. Geophys. Res. Space Physics*, **119**, 6199– 6212, doi:10.**1002**/2014JA019834.

Wiltberger, M., Pulkkinen, T. I., Lyon, J. G., and Goodrich, C. C. (2000), MHD simulation of the magnetotail during the December 10, 1996, substorm, *J. Geophys. Res*., **105** (A12), 27649– 27663, doi:10.1029/1999JA000251.

Wygant, J. R., Torbert, R. B., and Mozer, F. S. (1983), Comparison of S3-3 polar cap potential drops with the interplanetary magnetic field and models of magnetopause reconnection, *J. Geophys. Res*., **88** (A7), 5727– 5735, doi:10.1029/JA088iA07p05727.

Xie, H., N. Gopalswamy, O. C. St. Cyr, and S. Yashiro (2008), Effects of solar wind dynamic pressure and preconditioning on large geomagnetic storms, *Geophys. Res. Lett*., 35, L06S08, doi:10.1029/2007GL032298.

Yu, Y., and Ridley, A. J. (2013), Exploring the influence of ionospheric O+ outflow on magnetospheric dynamics: dependence on the source location, *J. Geophys. Res. Space Physics*, **118**, 1711– 1722, doi:10.1029/2012JA018411

Basis	**coupling function** $B^a \rho_{SW}{}^b V_{SW}{}^c F(\theta_{GSM})^d$	**a**	**b**	**c**	**d**	$F(\theta)$	**τ**	**Reference**
IMF (empirical fit to inter-diurnal geomagnetic data)	B	1	0	0	0	-	1 yr	*Svalgaard & Cliver* (2005)
solar wind speed	V_{SW}	0	0	1	0	-	1 yr	*Feynmann & Crooker* (1978)
(benchmark test)	$V_{SW}{}^2$	0	0	2	0	-	1day-1yr	*Finch & Lockwood* (2007)
empirical fit to inter-diurnal geomagnetic data	$BV_{SW}{}^{-0.1}$	1	0	−0.1	0	-	1 yr	*Lockwood et al.* (2014)
empirical fit to range geomagnetic data	$BV_{SW}{}^{1.7}$	1	0	1.7	0	-	1 yr	*Lockwood et al.* (2014)
southward IMF in GSM (benchmark test)	$[B_S]_{GSM}$	1	0	0	1	h.w.r.	1day-1yr	*Finch & Lockwood* (2007)
h.w.r. interplanetary electric field applied to *Dst*	$E_{SW}= [B_S]_{GSM}V_{SW}$	1	0	0	1	h.w.r.	2.5 min	*Burton et al.* (1975)
h.w.r. interplanetary electric field applied to Φ_{PC}	$E_{SW}= [B_S]_{GSM}V_{SW}$	1	0	1	1	h.w.r.	~ 10 min	*Cowley* (1984)
dawn-dusk interplanetary electric field applied to Φ_{PC}	$BV_{SW} sin^4(\theta_{GSM}/2)$	1	0	1	4	$sin^d(\theta/2)$	1 hr	*Wygant et al.* (1983)
(benchmark test)	$[B_S]_{GSM}V_{SW}{}^2$	1	0	2	1	h.w.r.	1day-1yr	*Finch & Lockwood* (2007)
solar wind Poynting flux (basis of ε)	$B_\perp{}^2V_{SW}$	2	0	1	0	-	-	-
solar wind kinetic energy flux (basis of P_α)	$\rho_{SW}V_{SW}{}^3$	0	1	3	0	-	-	-
solar wind Poynting flux with θ_{GSM} control	$B_\perp{}^2V_{SW} sin^4(\theta_{GSM}/2)$	2	0	1	4	$sin^d(\theta/2)$	-	-
epsilon factor	$\varepsilon = B^2V_{SW} sin^4(\theta_{GSM}/2)$	2	0	1	4	$sin^d(\theta/2)$	-	*Perreault & Akasofu* (1978)
solar wind dynamic pressure (benchmark test)	$p_{SW} = \rho_{SW}V_{SW}{}^2$	0	1	2	0	-	1day-1yr	*Finch & Lockwood* (2007)
empirical fit to *am*	$B_\perp\rho_{SW}{}^{1/2}V_{SW}{}^2 sin^4(\theta_{GSM}/2)$	1	0.5	2	4	$sin^d(\theta/2)$	3 hr	*Scurry and Russell* (1991)
empirical fit to Φ_D	$B_\perp V_{SW}{}^{4/3} sin^{9/2}(\theta_{GSM}/2)$	1	0	1.33	4.5	$sin^d(\theta/2)$	5 min	*Milan et al* (2012)
empirical fit to *Dst*	$BV_{SW}{}^2N_{SW}{}^{1/2} sin^6(\theta_{GSM}/2)$	1	0.5	2	6	$sin^d(\theta/2)$	1hr	*Temerin & Lee* (2006)
near-universal coupling function 1: based on Φ_D	$B^{2/3}V_{SW}{}^{4/3} sin^{8/3}(\theta_{GSM}/2)$	0.67	0	1.33	2.67	$sin^d(\theta/2)$	1 hr	*Newell et al.* (2007)
near-universal coupling function 2: fit to *Dst*	$B^{2/3}\rho_{SW}{}^{1/2}V_s{}^{7/3} sin^{8/3}(\theta_{GSM}/2)$	0.67	0.5	2.33	2.67	$sin^d(\theta/2)$	1hr	*Newell et al.* (2007)
theory of Φ_{PC}	$B_s\rho_{SW}{}^{-1/6}V_{SW}{}^{2/3}$	1	−0.17	0.67	4	h.w.r.	-	*Siscoe et al (2002)*
empirical fit to *Dst*	$B_\perp\rho_{SW}{}^{1/3}V_{SW}{}^{5/3} sin^4(\theta_{GSM}/2)$	1	0.33	1.67	4	$sin^d(\theta/2)$	1 hr	*Murayama* (1986)
empirical fit to *Dst*	$B_\perp\rho_{SW}{}^{1/2}V_s{}^{7/3} sin^6(\theta_{GSM}/2)$	1	0.5	2.33	6	$sin^d(\theta/2)$	1 hr	*Balikhin et al.* (2010)
theoretical estimate of Φ_D	$B_\perp V_{SW}\ sin^2(\theta_{GSM}/2)$	1	0	1	2	$sin^d(\theta/2)$	-	*Kan and Lee* (1979)
power input to the magnetosphere	$P_\alpha= B_\perp{}^{2\alpha}V_{SW}{}^{(7/3-2\alpha)}\rho_{SW}{}^{(2/3-\alpha)}\ sin^2(\theta_{GSM}/2)$	2α	2/3−α	7/3−2α	2	$sin^d(\theta/2)$	All	*Vasyliunas et al* (1982)
P_α fitted to *AL*	P_α for α= 0.50	1	0.27	1.33	4	$sin^d(\theta/2)$	1 min	*Bargatze et al* (1986)
P_α fitted to *AL* data, allow for data gaps	P_α for α= 0.42	0.84	0.25	1.49	4	$sin^d(\theta/2)$	1 hr	*Lockwood et al* (2019a)
P_α fitted to *AL* data allow for data gaps	P_α for α= 0.44	0.88	0.23	1.45	4	$sin^d(\theta/2)$	1 yr	*Lockwood et al* (2019a)
P_α fitted to range geomagnetic data	P_α for α= 0.36	0.72	0.31	1.61	4	$sin^d(\theta/2)$	1 day	*Lockwood* (2019)
Theory and fits to various geomagnetic data	$\approx B^{0.93}N_{SW}{}^{0.04}V_{SW}{}^{1.07} sin^2(\theta_{GSM}/2)$	0.93	0.04	1.07	2	$sin^d(\theta/2)$	1 hr	*Borovsky (2013)*
Theory and fits to various geomagnetic data	$\approx B^{1.26}N_{SW}{}^{-0.13}V_{SW}{}^{0.74} sin^2(\theta_{GSM}/2)$	1.26	−0.13	0.74	2	$sin^d(\theta/2)$	1 hr	*Borovsky (2013)*
empirical fit to *AL*	$B_\perp{}^{0.7}V_{SW}{}^{1.92}N_{SW}{}^{0.1}\sin^{3.67}(\theta_{GSM}/2)$	0.9	0.05	2.14	4.85	$sin^d(\theta/2)$	1 min	*Luo et al.* (2013)
numerical simulation	$B_\perp{}^{0.86}V_{SW}{}^{1.47}N_{SW}{}^{0.24}\{\sin^{2.70}(\theta_{GSM}/2)+0.25\}$	0.86	0.24	1.47	2.70	$sin^d(\theta/2)$	-	*Wang et al.* (2014)
empirical fit to *AL*	$B_\perp{}^{0.7}V_{SW}{}^{1.92}N_{SW}{}^{0.1}\sin^{3.67}(\theta_{GSM}/2)$	0.70 ±0.01	0.096 ±0.009	1.92 ±0.04	3.67 ±0.04	$sin^d(\theta/2)$	1 hr	*McPherron et al.* (2015)
empirical fit to *am*	$B_\perp{}^{0.955}\rho_{SW}{}^{0.355}V_{SW}{}^{2.434} sin^{3.87}(\theta_{GSM}/2)$	0.81 ±0.02	0.36 ±0.02	2.58 ±0.05	3.00 ±0.22	$sin^d(\theta/2)$	1 hr	*this paper*
empirical fit to Φ_{PC}	$B_\perp{}^{0.792}\rho_{SW}{}^{0.031}V_{SW}{}^{0.457} sin^{3.84}(\theta_{GSM}/2)$	0.64 ±0.05	0.02 ±0.01	0.55 ±0.03	2.45 ±0.05	$sin^d(\theta/2)$	1 hr	*this paper*

Table 1. A list of proposed coupling functions that share the general functional form $B^{a}\,\rho_{SW}^{b}\,V_{SW}^{c}\,F(\theta)^{d}$ used here. The first column gives the basis of the formulation in each case, which is given in the second column. Columns 3-6 give the exponents *a*, *b*, *c* and *d* and column 7 the $F(\theta)$ function used (h.w.r. stands for half-wave rectified). Column 8 gives the time resolution of the data on which the function was mainly developed and used. The last column is a reference to a paper using or proposing the formulation. Note that in some cases the formulation is not proposed as a viable coupling function and has only used to make comparisons with proposed coupling functions, some are physical properties of the interplanetary medium and given here only to record the exponents *a*, *b* and *c* that they yield.

<table>
<tr><th rowspan="2">T</th><th rowspan="2">Dates</th><th rowspan="2">lag, δt (min)</th><th colspan="6">optimum values</th></tr>
<tr><th>d</th><th>r_p</th><th>r_p^2</th><th>a</th><th>b</th><th>c</th></tr>
<tr><td>Φ_{PC}</td><td>1995-2020</td><td rowspan="3">18.5 ±1.3</td><td>2.50 ±0.07</td><td>0.865</td><td>0.748</td><td>0.642±0.019</td><td>0.018±0.008</td><td>0.550±0.047</td></tr>
<tr><td>Φ_{PC}</td><td>2012-2019</td><td rowspan="2">2.45 ±0.10</td><td>0.856</td><td>0.733</td><td rowspan="2">0.654±0.022</td><td rowspan="2">0.056±0.013</td><td rowspan="2">0.660±0.068</td></tr>
<tr><td>Φ_{PC}</td><td>1995-2011</td><td>0.887</td><td>0.787</td></tr>
<tr><td>*am*</td><td>1995-2020</td><td rowspan="3">30.5 ±4.0</td><td>3.00 ±0.22</td><td>0.814</td><td>0.663</td><td>0.802±0.022</td><td>0.360±0.012</td><td>2.560±0.072</td></tr>
<tr><td>*am*</td><td>2012-2019</td><td rowspan="2">3.05 (−0.20 to +0.40)</td><td>0.812</td><td>0.659</td><td rowspan="2">0.848±0.034</td><td rowspan="2">0.304±0.020</td><td rowspan="2">2.410±0.102</td></tr>
<tr><td>*am*</td><td>1995-2011</td><td>0.822</td><td>0.676</td></tr>
</table>

Table 2. The best fit exponents *a*, *b*, *c* and *d* and the resulting peak correlation coefficient r_p for the terrestrial parameters Φ_{PC} and *am*, from fits using the data from the range of dates given. Errors in the exponents *a*, *b* and *c* are generally dominated by the uncertainties in the fits at the optimum *d* and errors due to the uncertainty in the estimated optimum *d* are generally small. The results for 1995-2020 and 2012-2020 are for fitting and testing to the same set. The results for 1995-2011 are for fit exponents derived from the 2012-2020 subset. Uncertainties in *a*, *b* and *c* allow for both the fit uncertainties at a given *d* and the uncertainty caused by the uncertainty in *d*.

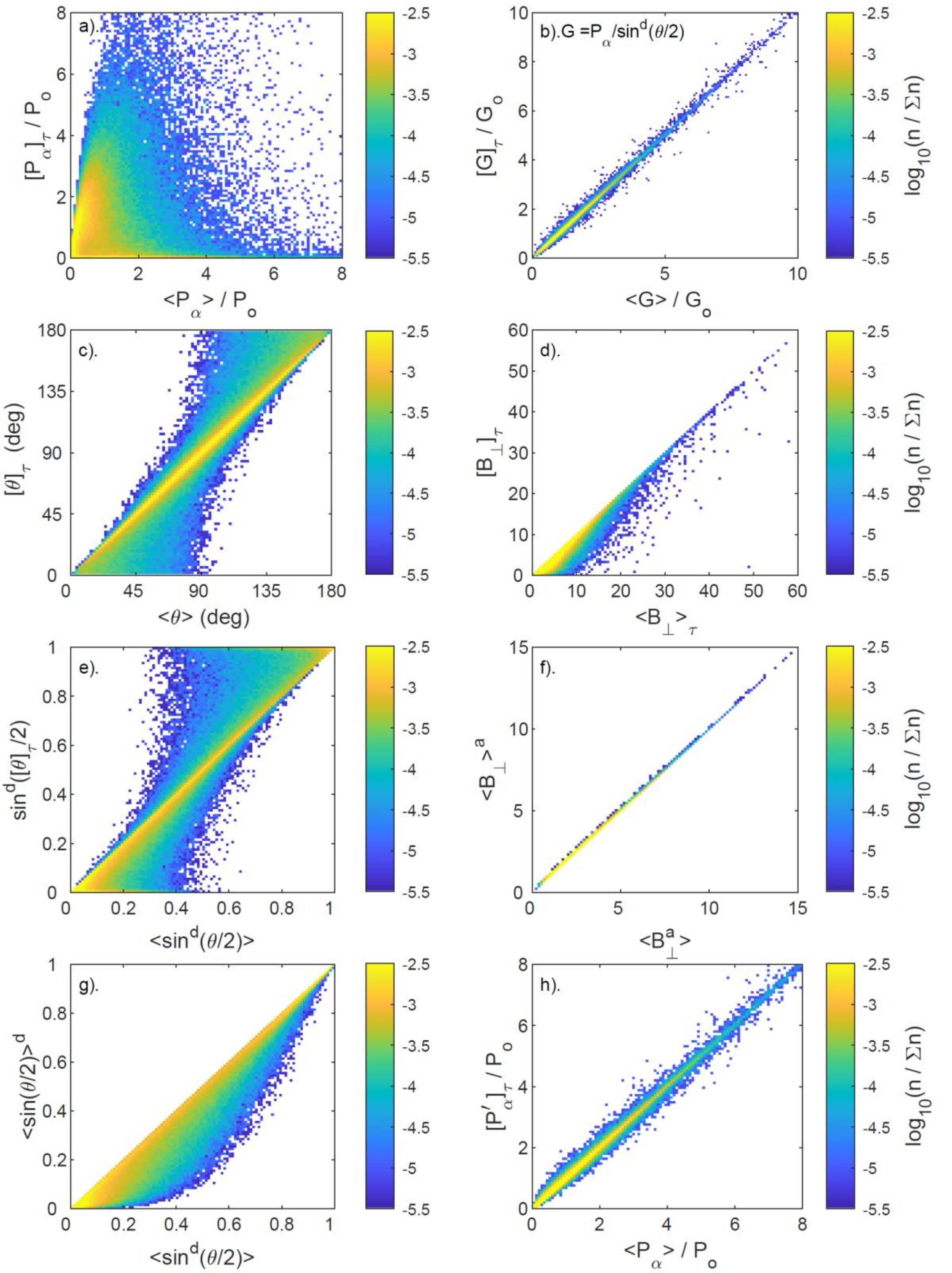

Figure 1. Comparison of combine-then-average, average-then-combine and our compromise hybrid procedure for averaging 1-minute data into 1-hour data (τ = 1hr). In all panels, the horizontal axis gives the result of the combine-then-average approach which is what we ideally would wish to use to mimic solar wind forcing of the magnetosphere. The vertical

axes in (a)-(e) give the result of a full average-then-combine procedure. In each case the fraction of samples $n/\Sigma n$ is color-coded, where n is the number of samples small bins. The raw data used are 9,930,183 valid 1-minute integrations of estimated power input to the magnetosphere, P_α , and 11,646,678 valid 1-minute values of the IMF clock angle θ and tangential component $B_\perp$ observed between 1995-2020 (inclusive). (a) is for the coupling function P_α for $\alpha = 1/3$ and $d = 4$ (the normalizing factor P_o is the arithmetic mean of P_α for all datapoints) in bins of P_α/P_o of size 0.08; (b) is the corresponding plot for G, which is P_α without the IMF orientation factor; (c) is for the IMF clock angle (in the GSM frame of reference) θ in bins that are $2° \times 2°$; (d) is for the tangential IMF component $B_\perp = (B_y^2+B_x^2)^{1/2}$ in bins of $0.5\text{nT} \times 0.5\text{nT}$ and (e) is for $sin^d(\theta/2)$ in bins 0.01×0.01. Part (f) compares $<B_\perp>^a$ with $<B_\perp^a>$ (where $a = 2\alpha$ for the P_α coupling function) and part (g) compares $<sin(\theta/2)>^d$ with $<sin^d(\theta/2)>$. In part (h) the y-axis is the result of our hybrid averaging procedure for P_α, $[P'_\alpha]_{1hr}$, defined by Equation (15).

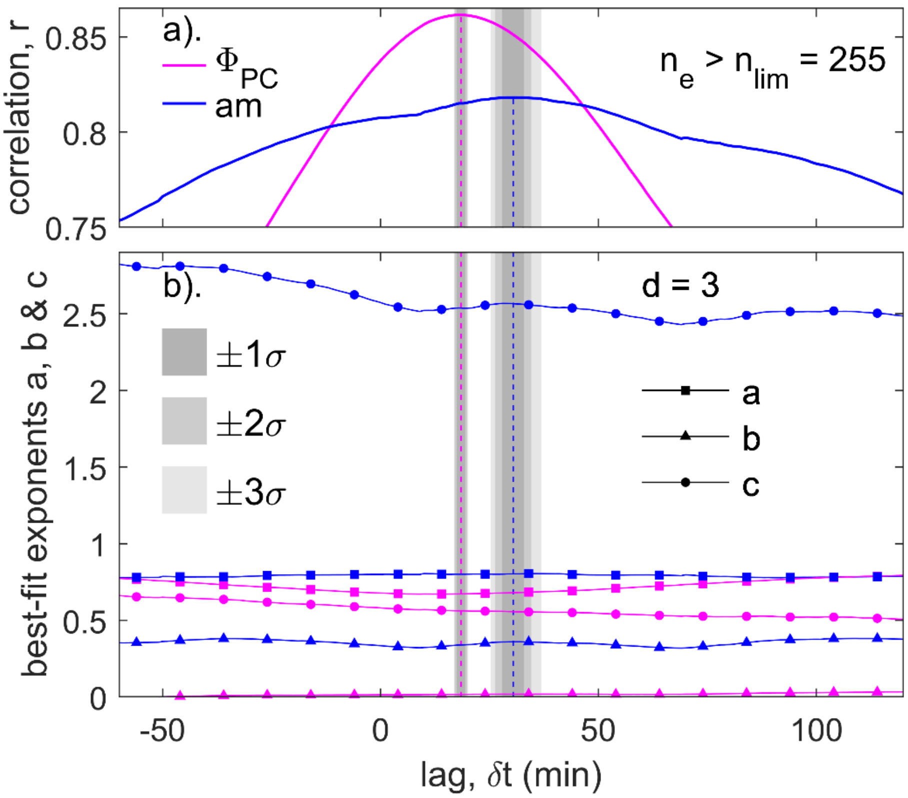

Figure 2. (Top) Lag correlograms (linear correlation coefficient, *r*, as a function lag, δt) of predicted variations using 61-point boxcar (running) means of the coupling function C_f from 1-minute interplanetary parameters with hourly observations of the transpolar voltage Φ_{PC} (in mauve) and the *am* geomagnetic index (in blue). Note that unless otherwise stated, C_f in this and later figures refers to hourly means $[C'_f]_{1hr}$, derived from our hybrid formulation, Equation (15). Both the Φ_{PC} and the *am* data are for the full 25-year dataset, but for times when the number of SuperDARN radar echoes n_e exceeds the threshold n_{min}. This yields N = 65,133 data points. The lag δt = 0 means that the radar data and the Omni interplanetary data are averaged over the same one-hour interval and positive δt corresponds to the interplanetary data leading the terrestrial data. The exponent *d* is assumed to be 3 but tests of values between 1 and 6 made negligible differences to the optimum δt , δt_p, derived. The hourly *am* data are derived from the observed 3-hourly *am* values using PCHIP interpolation to the mid-points of the hourly integration periods for the radar data. The dark gray, lighter gray, and lightest gray areas define, respectively, the 1-σ, 2-σ and 3-σ uncertainty bands in the lag δt_p and are defined using the Meng-Z test (see text for details). The vertical dashed lines give the lag δt_p that yields the peak *r*, r_p, which is 0.862 at δt_p = 18.5 ± 1.3 min for Φ_{PC} and is 0.818 at δt_p = 30.5 ± 4.0 min for *am*, the uncertainties being at the 2-σ level. (Bottom) The best-fit exponents *a*, *b* and *c* as a function of δt (lines marked by squares, triangles and circles, respectively), derived using the Nelder-Mead search algorithm to maximise *r*.

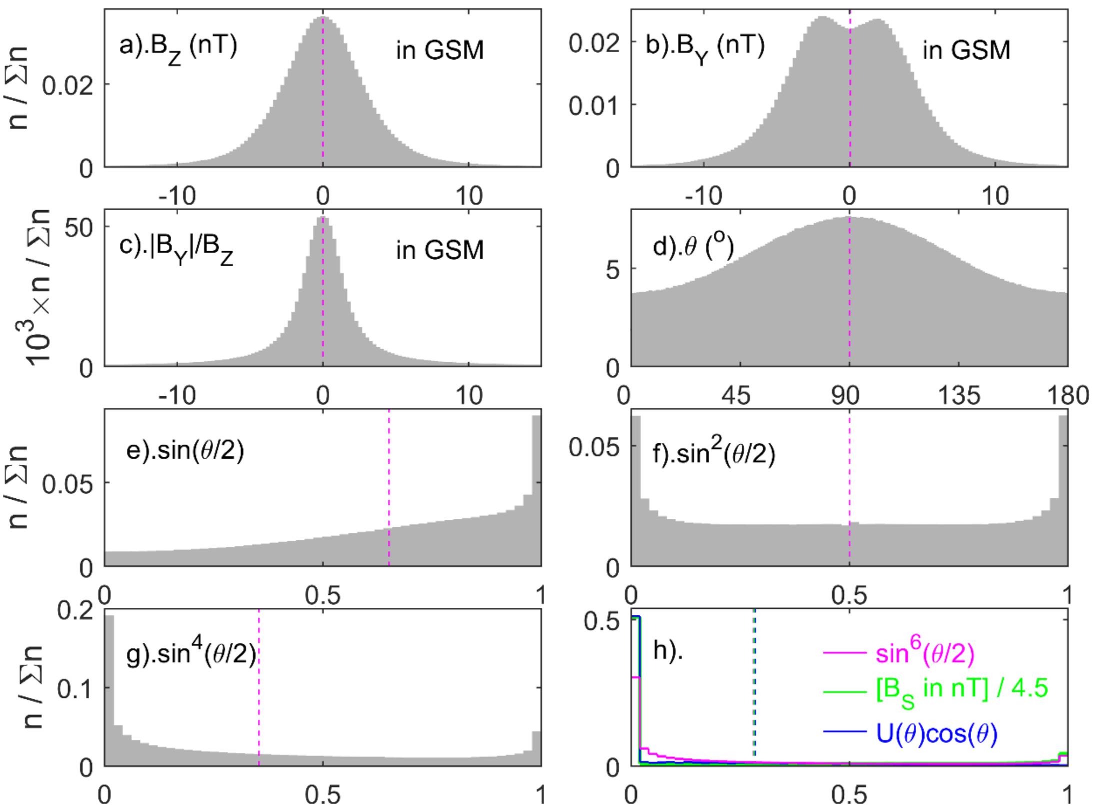

Figure 3. Distributions of 1-minute interplanetary parameters relating to IMF orientation in the GSM frame of reference: (a) the IMF B_Z component; (b) the IMF B_Y component; (c). the ratio $|B_Y|/B_Z$; (d). the clock angle $\theta = tan^{-1}(|B_Y|/B_Z)$; (e). $sin(\theta/2)$; (f). $sin^2(\theta/2)$; (g). $sin^4(\theta/2)$; and (h) $sin^6(\theta/2)$ in mauve, $U(\theta)cos(\theta)$ in blue (where $U(\theta) = 0$ for θ<90° and $U(\theta) = -1$ for $\theta \geq 90°$) and $B_S/4.5$ in green (where B_S is the half-wave rectified southward component of the IMF, $B_S = -B_Z$ for $B_Z < 0$ and $B_S = 0$ for $B_Z \geq 0$: the factor 4.5 is used because it makes the mean value on the axis used the same as for $sin^6(\theta/2)$ and $U(\theta)cos(\theta)$ for the scale used). The data are 116,466,78 1-minute samples from the Omni database for 1995-2020 (inclusive), and the vertical axis is the fraction of samples in each bin, $n/\Sigma n$, where n is the number of samples in bins that are 1% in width of the range shown on the horizontal axis in each case. Vertical dashed lines give the mean value for the whole interval.

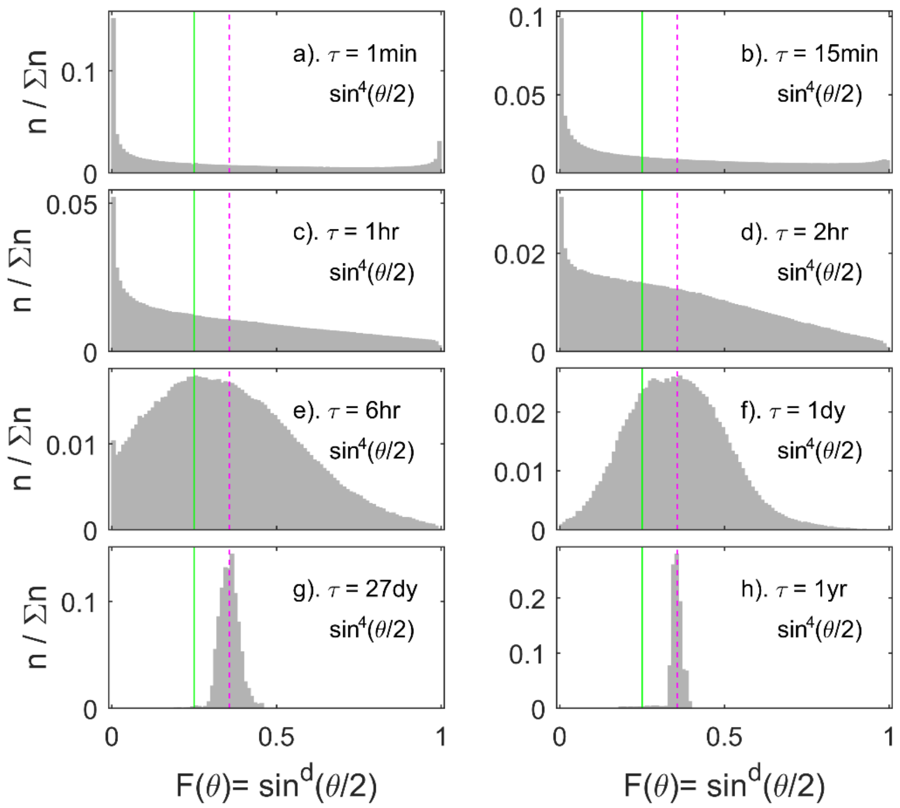

Figure 4. Distributions of the IMF orientation factor $F(\theta) = sin^d(\theta/2)$ for $d = 4$, where θ is the IMF clock angle in GSM coordinates, for data averaging timescales τ of: (a) 1 minute; (b) 15 minutes; (c) 1 hour (used in this paper); (d) 2 hours; (e) 6 hours; (f) 1 day; (g) a solar rotation period of 27 days and (h). one year. The numbers of samples, n, as a fraction of the total number Σn, in bins 0.01 wide are shown in each case and the dataset used is the same as in Figure 3. The vertical mauve dashed lines are for the overall average of all samples. The vertical green line is at $\theta = 90°$ for which the IMF lies the GSM equatorial plane. Note that the lowest bin in $sin^4(\theta/2)$, which is 0-0.01, corresponds to a range in θ of 0-36.9° whereas the highest bin (0.99-1) corresponds to 171.9-180°.

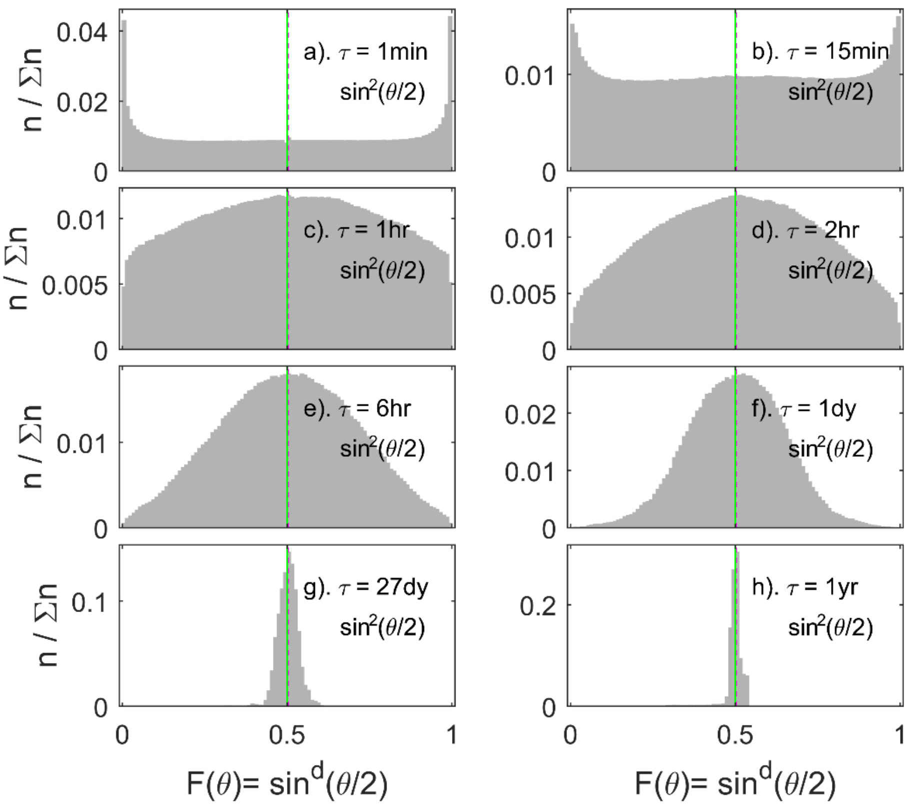

Figure 5. Distributions of the IMF orientation factor $F(\theta) = sin^d(\theta/2)$ for $d = 2$, in the same format as Figure 4 and for the same dataset. Here the lowest bin in $sin^2(\theta/2)$, which is 0-0.01, corresponds to a range in θ of 0-11.5°, whereas the highest bin (0.99-1) corresponds to 168.5-180°.

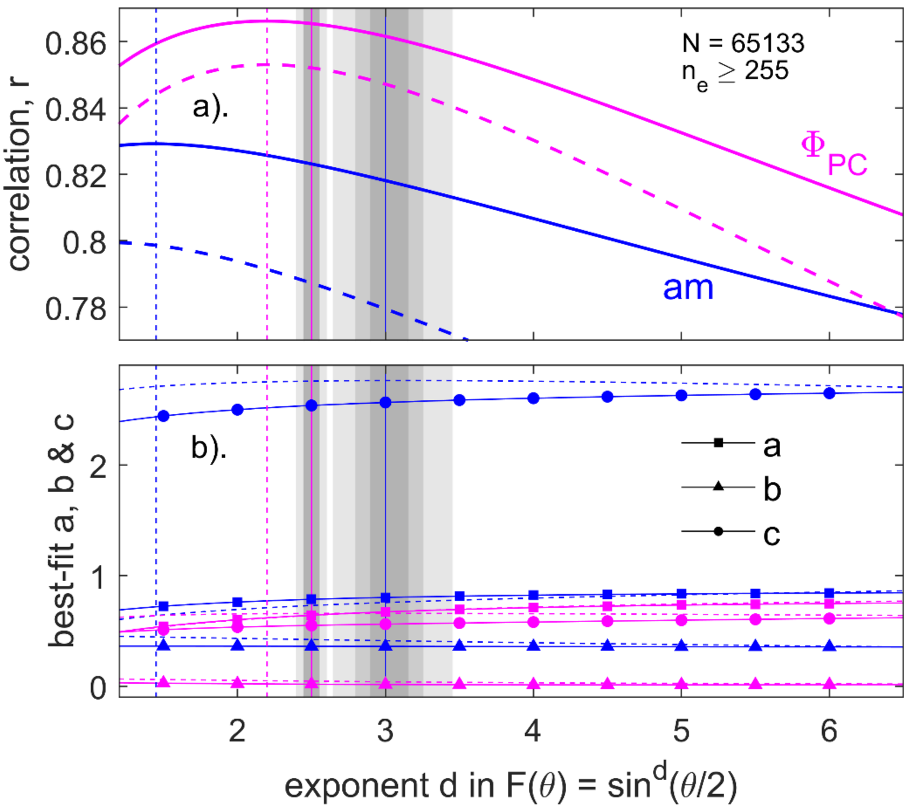

Figure 6. Analysis of the exponent of the d of the $F(\theta) = sin^d(\theta/2)$ IMF orientation factor for all N = 65133 samples which meet the criterion of the hourly mean number of radar echoes $n_e > n_{min}$ = 255. For each value of d, the value of the other three exponents a, b, and c are derived by the Nelder-Mead simplex search method to maximise the correlation coefficient r between the hourly lagged coupling function. The solid lines are for the hybrid hourly mean averaging coupling function $[C'_f]_{1hr} = <B>^a . <\rho_{SW}>^b . <V_{SW}>^c . <sin^d(\theta/2)>$ and the dashed lines are for the average-then-combine coupling function $[C_f]_{1hr} = <B>^a . <\rho_{SW}>^b . <V_{SW}>^c . sin^d([\theta]_{1hr}/2)$. The results for observed Φ_{PC} are in mauve and interpolated hourly values of *am* are in blue. The vertical dashed lines mark the peak correlation in each case, the vertical solid lines the optimum d (determined from Figures 7 and 8) and the gray areas the 2σ uncertainty bands of the optimum d. (Top) The correlation coefficients, r, as a function of d. (Bottom) The best fit values of the exponents a (identified by squares), b (triangles) and c (circles) as a function of d for $[C'_f]_{1hr}$ as solid lines and solid symbols and $[C_f]_{1hr}$ as dashed lines and open symbols.

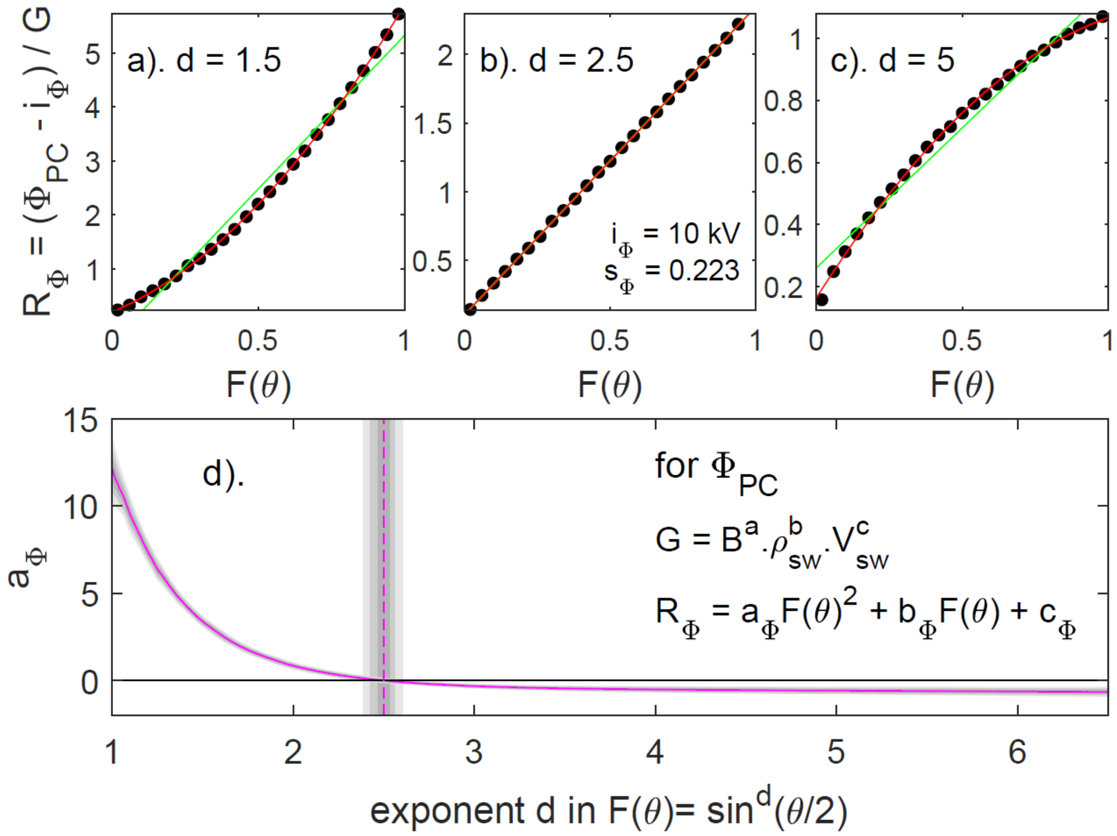

Figure 7. Tests of the IMF orientation term, $F(\theta) = sin^d(\theta/2)$ for the transpolar voltage Φ_{PC}. Parts (a), (b) and (c) show plots of the means of $R_\Phi = (\Phi_{PC} - i_\Phi)/G$ as a function of mean $F(\theta)$, both averaged for 25 bins of $F(\theta)$ that are 0.04 wide. G is given by Equation (14), where C_f is the optimum coupling function for the optimum exponents a, b and c for the d value in question, as shown in Figure 6. (a) is for d = 1.5, (b) for the derived best d of 2.60 and (c) is for d = 5. The green and red lines are linear and quadratic fits, respectively, to the mean values. The values of the linear regression coefficients s_Φ and i_Φ (see equations 16 and 17) are given in (b), where the s_Φ values are for $B_\perp$ in nT, ρ_{SW} in kg m^{-3} and V_{SW} in km s^{-1}. (d). The mauve line is coefficient of the quadratic term of the second-order polynomial fit to the means, a_Φ, as a function of d: the optimum d gives a proportional relationship between $<R_\Phi>$ and $<F(\theta)>$, i.e., when $a_\Phi = 0$, marked by the vertical dashed line. Under the mauve line in three shades of gray area are the 1-σ, 2-σ and 3-σ uncertainty band in a_Φ, the limits to which define the corresponding uncertainty bands in the optimum d, giving a 2-σ uncertainty in the optimum d of ±0.07. Note that in this case for Φ_{PC} the differences between the uncertainty bands are often so small that they cannot be discerned; they are more clearly seen in Figure 8 for *am*. Part (b) confirms this proportional relation at this optimum d = 2.50 for which the exponents are given in Table 2. The uncertainty in these values is evaluated using Figures 13-15.

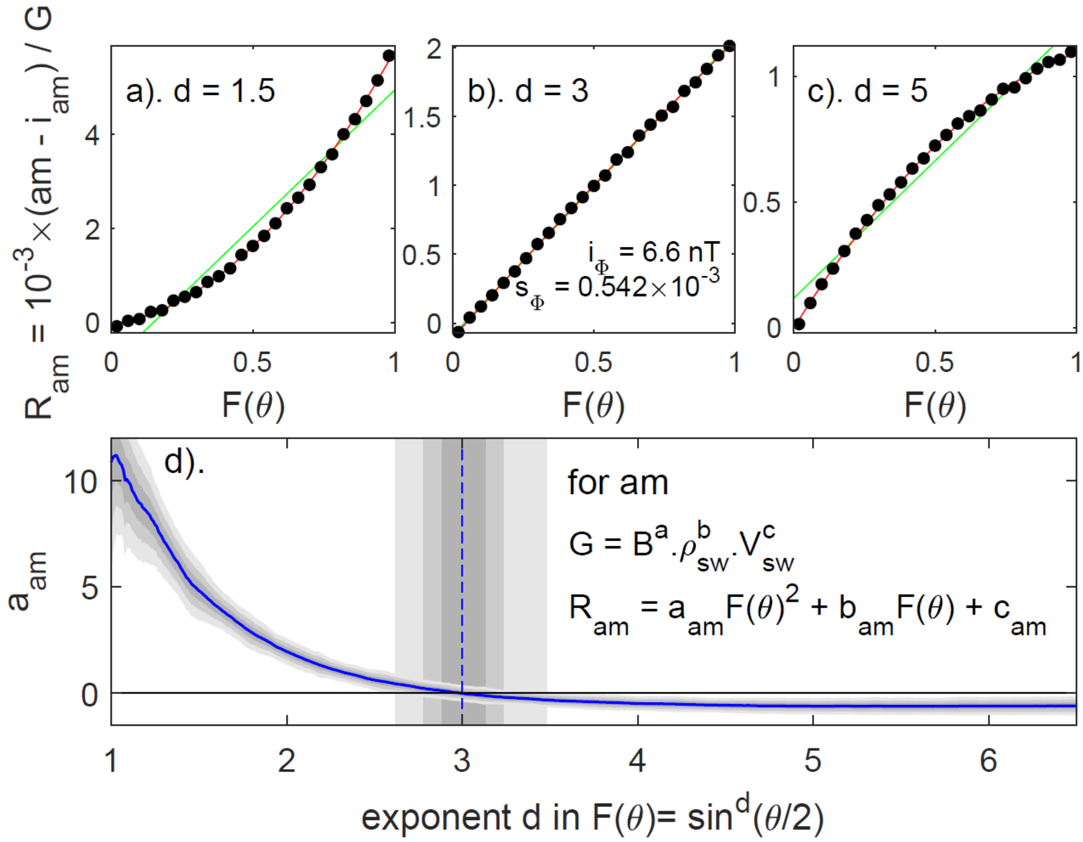

Figure 8. The same as Figure 7 for the *am* index. The blue line in part (d) is the best-fit a_Φ under which the three gray areas define the 1-σ, 2-σ and 3-σ uncertainty bands in a_Φ , the limits to which define the vertical uncertainty bands in the optimum *d* shown. The optimum *d* giving the proportional relationship is $d = 3.00\pm0.22$ for which the exponents *a*, *b* and *c* are given in Table 2. The uncertainty in these values is evaluated in Figures 13-15.

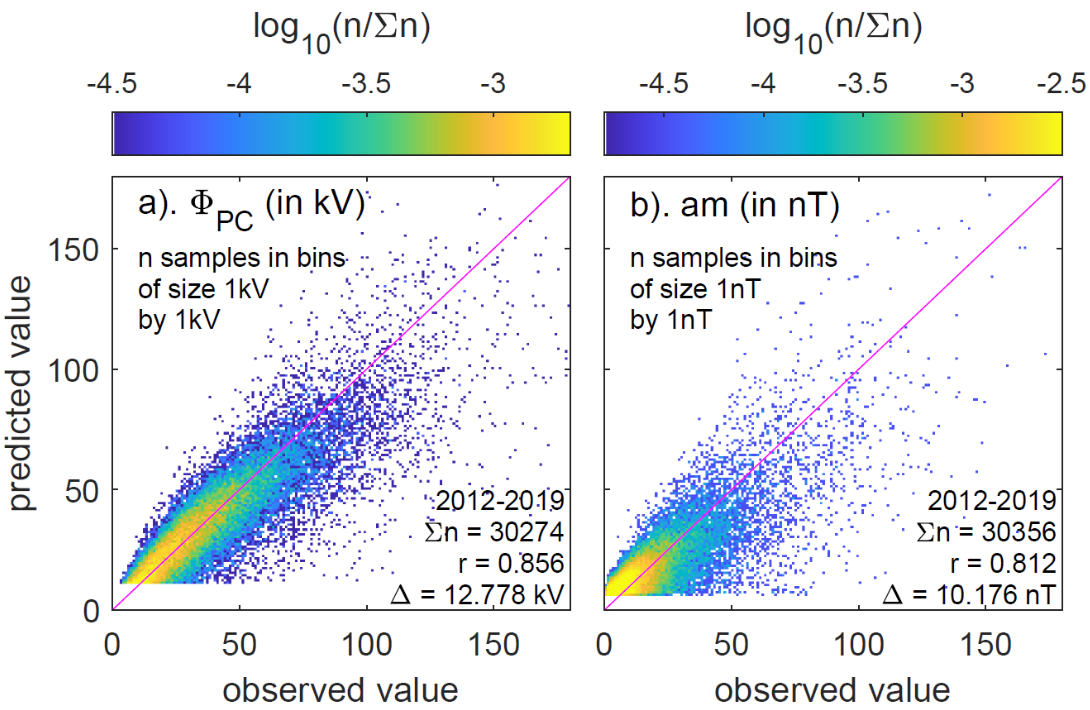

Figure 9. Datapoint density plots of predicted against observed values of (a) the transpolar voltage Φ_{PC} and (b) the *am* geomagnetic index for the optimum *d* values defined in section 3. These data are for the fit dataset which is for 2012-2020. The best fit exponents and their uncertainties are given in Table 2. In both cases, the optimum fit of C_f has been scaled to the data by ordinary least-squares linear regression. The numbers samples *n* (as a faction of the total number Σn) in bins, which are 1kV × 1kV wide in (a) and 1nT × 1nT wide in (b), are colour-coded on the logarithmic scales given. The diagonal mauve lines mark perfect agreement of observed and predicted values. The correlation coefficient *r* and the root mean square deviation Δ of observed and predicted values are given in each panel, along with the total number of valid data-point pairs, *N* .

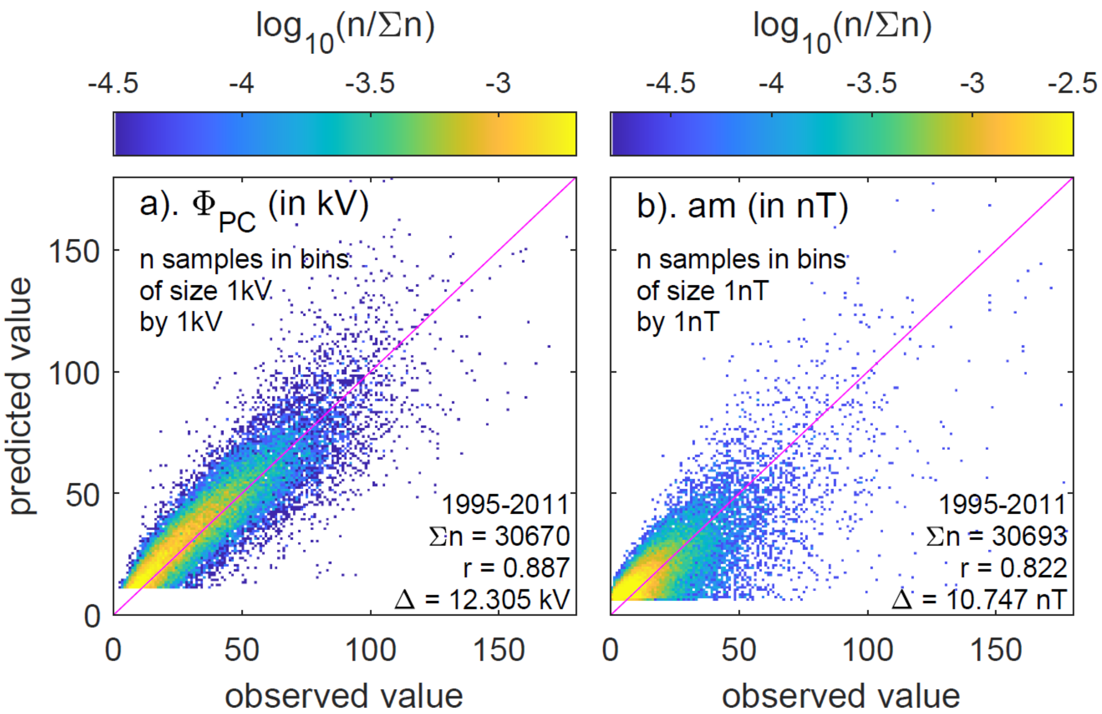

Figure 10. Same as Figure 9 but for the independent test dataset from 1995-2011 but using the best-fit exponents at the optimum lag derived for the fit dataset (2012-2020). The correlation coefficients *r* and the root mean square deviations *Δ* are very similar to the corresponding values for the fit dataset shown in Figure 9.

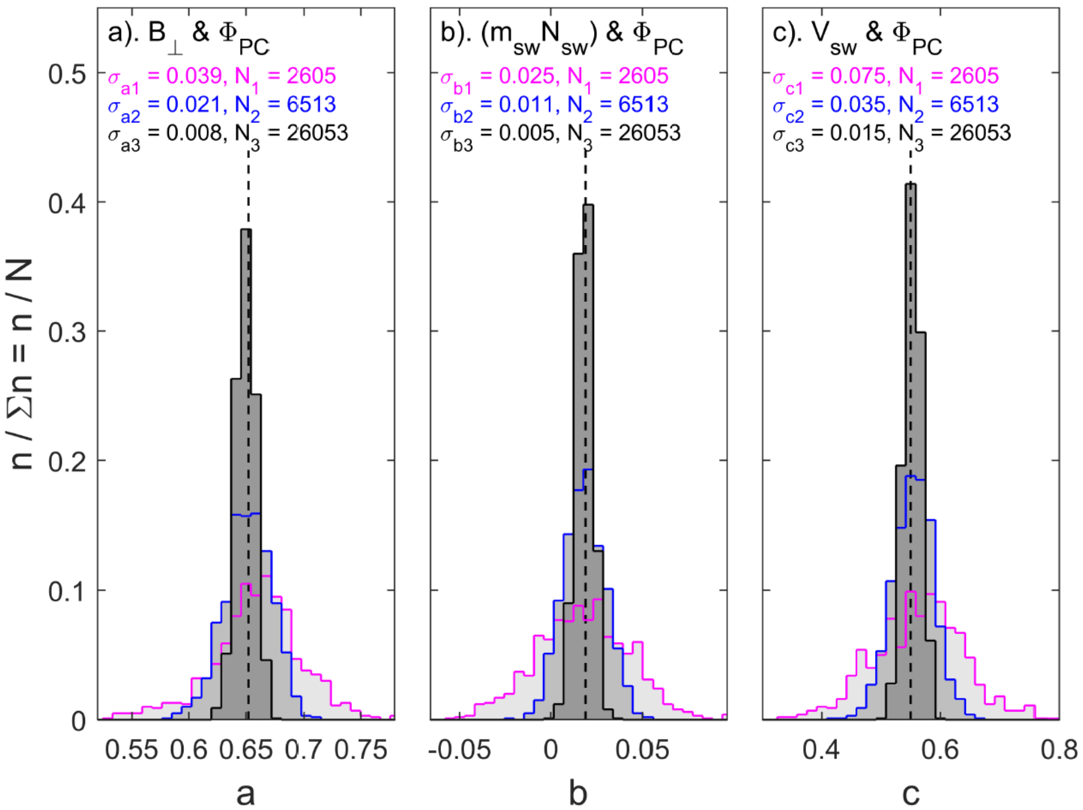

Figure 11. Distributions of fitted values of exponents a (left panel), b (middle panel) and c (right panel) for fits to the transpolar voltage, Φ_{PC}, drawn from the entire 25-year dataset of 65133 values with $n_e > n_{min} = 255$. The fraction of samples $n/\Sigma n$ in bins of width (1/30) of the maximum range of each exponent are plotted. In each case, three histograms are shown: (1) the light grey histogram bounded by the mauve line is for (1/25) of the whole dataset (N = 2606 samples, on average corresponding to 1 yr of data); (2) the darker grey bounded by the blue line is for (1/10) of the whole dataset (N = 6513 samples, on average corresponding to 2.5 yr of data); the darkest grey bounded by the black line is for (1/2.5) of the whole dataset (N = 26503 samples, on average corresponding to 10 yr of data). The standard deviation of the distribution is given in each case with the generic name σ_{xi} where x is the exponent in question and i is the number of the dataset number. The distributions are generated by taking 1000 random selections of N samples from the total of 65130 samples with $n_e > n_{min} = 255$ available. The vertical dashed lines give the values for the full set of 65130 samples.

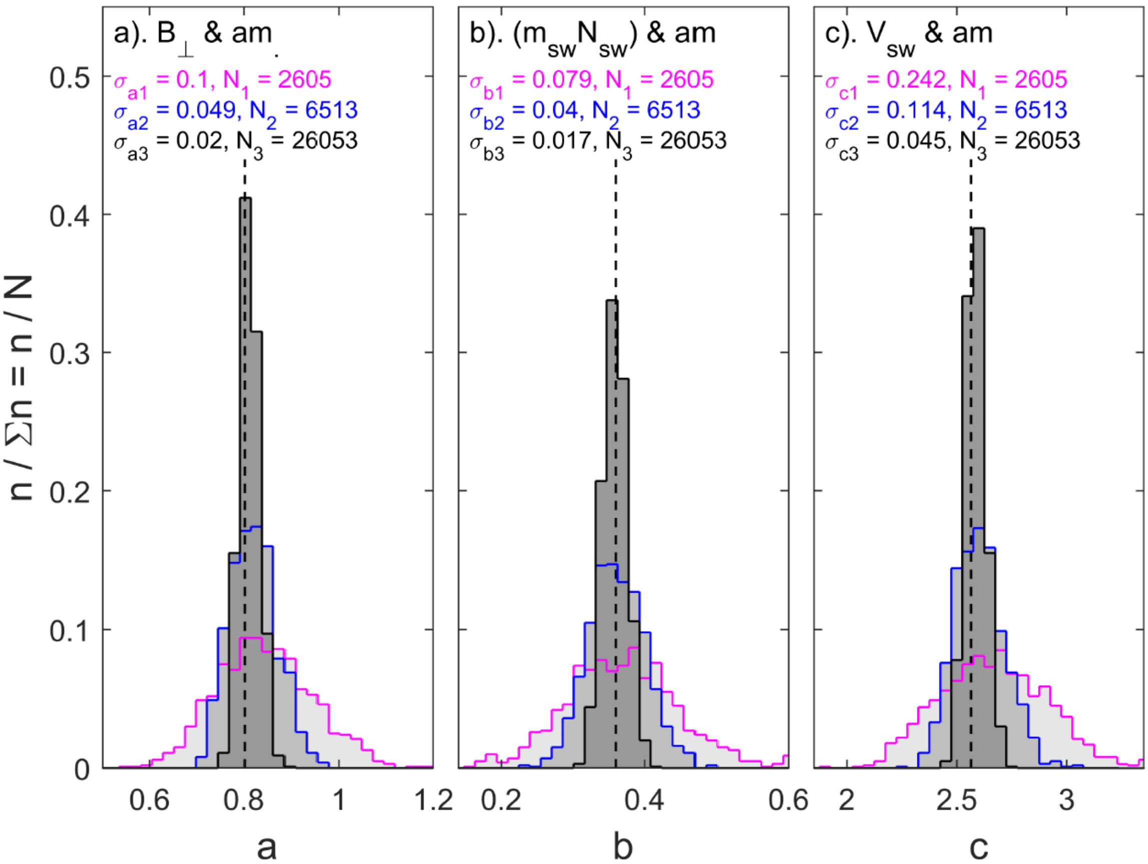

Figure 12. Same as Figure 11 for the *am* geomagnetic index data.

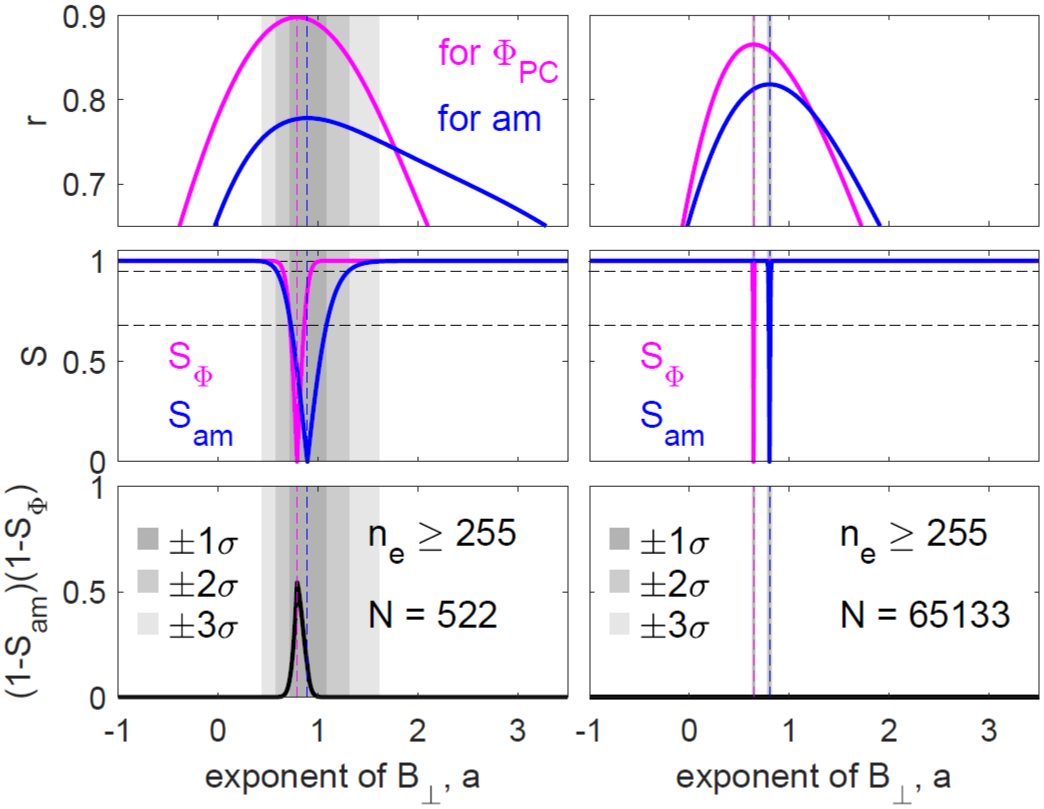

Figure 13. Analysis of the effect of the number of samples, N on the optimum exponents a of the interplanetary magnetic field component, $B_{\perp}$. The left-hand column is for a random selection of N = 522 samples from the full dataset with $n_e > n_{min} = 255$ (which is 0.4% for the full dataset); the right-hand column is a for the full dataset of 61533 samples. In all panels mauve is for Φ_{PC}, blue for *am*. (Top) the correlation coefficient as a function of a. The vertical dashed lines mark the peak correlation in each case, r_p. Note that the exponents b and c are found for each imposed a value using the Nelder-Mead simplex search as was used to find all three exponents when none of the three were fixed. (Middle). The significance S = (1-p) of the difference in a general correlation coefficient $r(a)$ and the optimum value r_p as a function of a, where p is the p-value for the null hypothesis that $r(a)$ and r_p are the same. S_{Φ} is S for Φ_{PC}, S_{am} is for *am*. (bottom) The probability that the optimum d is the same for Φ_{PC} and *am*, $(1-S_{\Phi})$ $(1- S_{am})$. For the full dataset the optimum d values give $a = 0.642\pm0.015$ for Φ_{PC} and $a = 0.802\pm0.010$ for *am*, where uncertainties are at the 2-σ level. If we allow for the uncertainty in d these uncertainties increase to ±0.019 and ±0.022. Optimum values and their 2-σ uncertainty for both the full 1995-2020 dataset and for the 2012-2020 subset are given in Table 2.

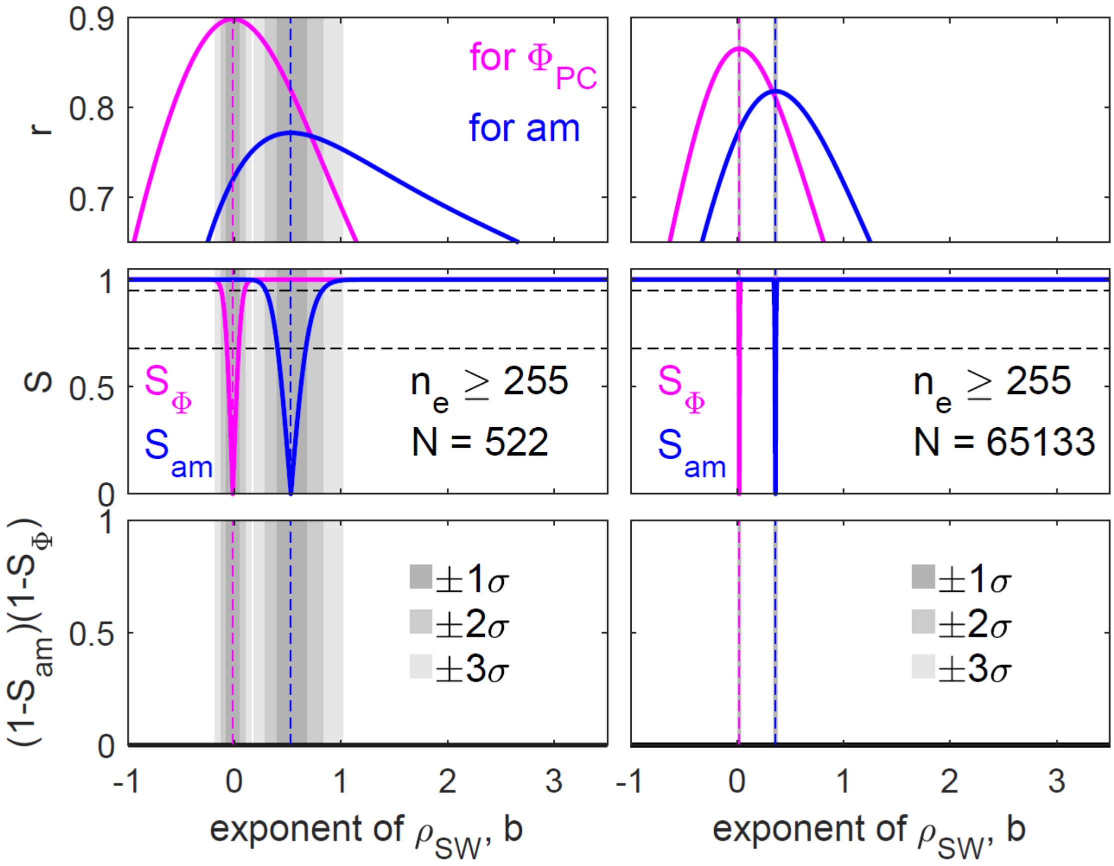

Figure 14. Analysis of the effect of the number of samples, N on the optimum exponents b of the solar wind number density, ρ_{SW}. The format, axes and sample numbers are the same as Figure 13. Best fit values of c and their uncertainty for the full 1995-2020 dataset and for the 2012-2020 subset are given in Table 2.

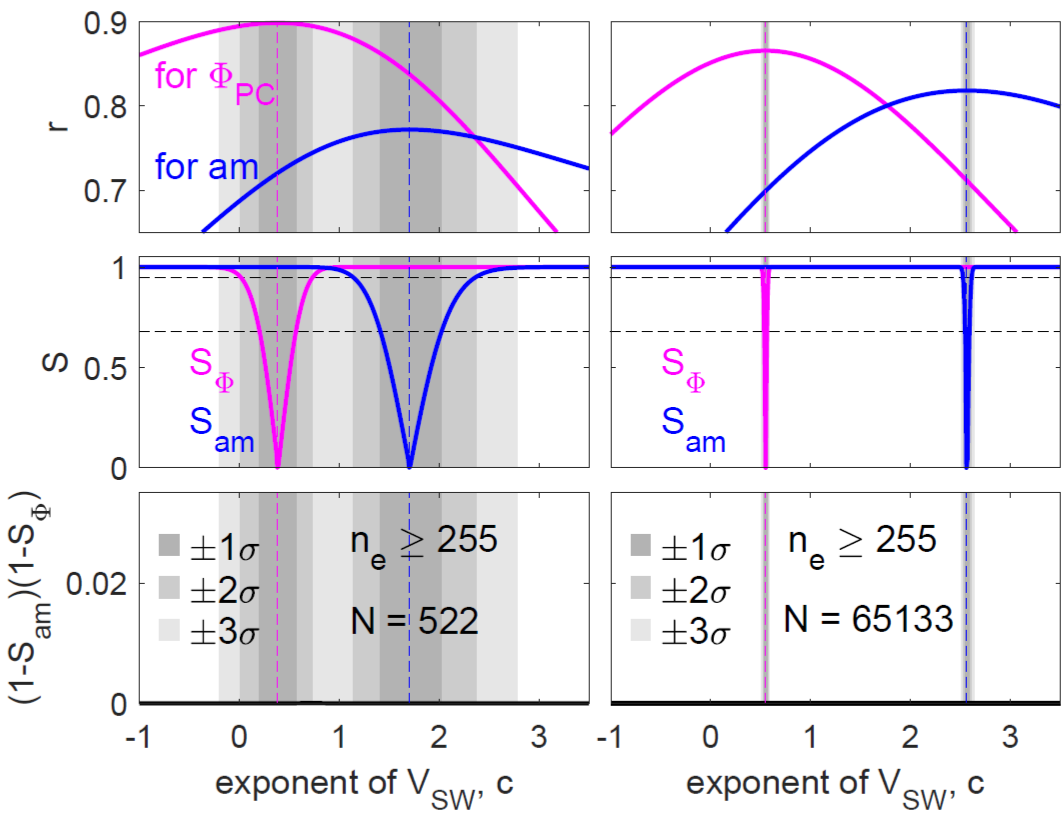

Figure 15. Analysis of the effect of the number of samples, *N* on the optimum exponents *c* of the solar wind speed, V_{SW}. The format, axes and sample numbers are the same as Figure 13. Best fit values of *c* and their uncertainty for the full 1995-2020 dataset and for the 2012-2020 subset are given in Table 2.