University of Reading

The development of a space climatology: 1. solar-wind magnetosphere coupling as a function of timescale and the effect of data gaps

Article

Accepted Version

Lockwood, M. ORCID: https://orcid.org/0000-0002-7397-2172, Bentley, S. N., Owens, M. J. ORCID: https://orcid.org/0000-0003-2061-2453, Barnard, L. A. ORCID: https://orcid.org/0000-0001-9876-4612, Scott, C. J. ORCID: https://orcid.org/0000-0001-6411-5649, Watt, C. E. and Allanson, O. (2019) 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 (1). pp. 133-156. ISSN 1542-7390 doi: 10.1029/2018SW001856 Available at https://centaur.reading.ac.uk/78209/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1029/2018SW001856

Publisher: American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Paper accepted by Space Weather, July 2018. doi: 10.1029/2018SW001856

The development of a space climatology: 1. Solar-wind magnetosphere coupling as a function of timescale and the effect of data gaps

Mike Lockwood[1], Sarah N. Bentley[1], Mathew J. Owens[1], Luke A. Barnard[1],
Chris J. Scott[1], Clare E. Watt[1] and Oliver Allanson[1]

[1]Department of Meteorology, University of Reading, Earley Gate, P.O. Box 243, Reading, Berkshire, RG6 6BB, UK.

Corresponding author: M. Lockwood (m.lockwood@reading.a.uk)

Key Points:

- Gaps in interplanetary data from minutes to years generate significant errors in empirically-derived solar wind coupling functions
- All coupling functions derived using data from before 1995 need to be critically re-evaluated and checked for over-fitting
- The optimum coupling function quantifying power input to the magnetosphere has no detectable variation with averaging timescale

Abstract

Different terrestrial space weather indicators (such as geomagnetic indices, transpolar voltage, and ring current particle content) depend on different "coupling functions" (combinations of near-Earth solar wind parameters) and previous studies also reported a dependence on the averaging timescale, τ. We study the relationships of the *am* and SME geomagnetic indices to the power input into the magnetosphere P_α, estimated using the optimum coupling exponent α, for a range of τ between 1 min and 1 year. The effect of missing data is investigated by introducing synthetic gaps into near-continuous data and the best method for dealing with them when deriving the coupling function, is formally defined. Using P_α, we show that gaps in data recorded before 1995 have introduced considerable errors into coupling functions. From the near-continuous solar wind data for 1996-2016, we find $\alpha = 0.44 \pm 0.02$ and no significant evidence that α depends on τ, yielding $P_\alpha \propto B^{0.88} V_{sw}^{1.90} (m_{sw} N_{sw})^{0.23} \sin^4(\theta/2)$, where B is the Interplanetary Magnetic Field (IMF), N_{sw} the solar wind number density, m_{sw} its mean ion mass, V_{sw} its velocity and θ the IMF clock angle in the Geocentric Solar Magnetospheric reference frame. Values of P_α that are accurate to within ±5% for 1996-2016 have an availability of 83.8% and the correlation between P_α and *am* for these data is shown to be 0.990 (between 0.972 and 0.997 at the 2σ uncertainty level), 0.897±0.004, and 0.790±0.03, for τ of 1 year, 1 day and 3 hours, respectively, and that between P_α and *SME* at τ of 1 min. is 0.7046±0.0004.

1. Introduction

1.1 Coupling functions

On short timescales, the coupled magnetosphere-ionosphere-thermosphere system responds to the magnitude of the southward component of the Interplanetary Magnetic Field (IMF, in a suitable frame oriented with respect to the geomagnetic field axis, such as Geocentric Solar Magnetospheric, GSM). There are two time constants of response, the "directly-driven" system responds on a timescale of order a few minutes [*Nishida*, 1968; *Lockwood et al.*, 1986; *Etemadi et al.*, 1988, *Todd et al.*, 1988]. The directly-driven flows store magnetic energy in the magnetospheric tail and subsequently that energy is released and deposited in the inner magnetosphere and nightside auroral ionosphere and thermosphere via the substorm current wedge. This is the "storage-unloading" system, which generates a response after a delay of typically between about 30 and 60 min. [*Schatten and Wilcox*, 1967; *Arnoldy*, 1971; *Baker et al.*, 1981, 1983; *Lockwood et al*., 1990]. *Finch et al.* [2007] used data from the global network of geomagnetic observatories to show how stations responded differently to the directly-driven and storage/unloading systems depending on their position (in geomagnetic latitude and magnetic local time (MLT) coordinates). Similar conclusions was reached by *Dods et al*. [2015; 2017], using network analysis.

A great many combinations of near-Earth interplanetary parameters (so-called "coupling functions") have been proposed over many years to describe the transfer of energy, and/or mass, and/or momentum, and/or electric field from the solar wind into the Earth's magnetosphere-ionosphere-thermosphere system [e.g., *Burton et al.*,1975; *Crooker et al.*, 1977; *Feynman and Crooker*, 1978; *Perreault and Akasofu*, 1978; *Kan and Lee*, 1979; *Reiff et al.*, 1981; *Murayama*, 1982; *Wygant et al.*, 1983; *Bargatze et al.*,1985; 1986; *Vasyliunas et al.*, 1982; *Scurry and Russell*, 1991; *Wing and Sibeck*,1997; *Wu and Lundstedt* , 1997; *Papitashvili et al*. 2000; *Temerin and Li*, 2002; 2006; *Newell et al.*, 2007; *Borovsky*, 2008; *Balikhin et al.*, 2010; *Spencer et al.*, 2011; *Tenfjord and Ostgaard*, 2013, *McPherron et al*., 2015]. These are derived and tested by comparison with terrestrial space weather disturbance indices (sometimes in combination) which respond to the energy, mass, electric field and/or momentum input to the magnetosphere from the solar wind (in a combination thereof that depends on which terrestrial index is used). *Gonzalez* [1990] examined the empirical coupling functions in use at the time and concluded that almost all could be derived either from electric field or power transfer associated with the reconnection process. Although they are related in a general sense [e.g. *Saba et al.*, 1997], different terrestrial indices respond differently to different combinations of energy, mass, momentum and/or electric field transfer into the magnetosphere (and on different timescales): hence it is not surprising that there are a wide variety of derived coupling functions. For example, *Svalgaard and Cliver* [2005] noted that different geomagnetic indices responded differently to different coupling functions. This is also unsurprising as each is influenced by different combinations of the currents in the terrestrial system: the ionospheric DP1 (auroral electrojet and substorm current wedge); the ionospheric DP2 (growth phase convection) currents; the magnetospheric ring current; the magnetopause (Chapman-Ferraro) currents; the cross-tail current sheet; and (unless ionospheric conductivities are spatially uniform) the field-aligned (Birkeland) currents. *Svalgaard and Cliver* [2005] also noted that this is a very useful feature because it allows reconstruction of multiple solar wind parameters from combinations of indices for times before interplanetary spacecraft were available, as exploited by *Lockwood et al.* [1999], *Lockwood* [2003], *Rouillard et al*. [2009] and *Lockwood et al.* [2014a]. A comparison of the optimum exponents of interplanetary parameters for various commonly-used geomagnetic indices has been presented for annual timescales by *Lockwood* [2013].

On long timescales (e.g., τ = 1 year) the substorm cycles and geomagnetic storm responses are averaged out and also the IMF orientation factor converges to almost a constant factor. On such timescales, the coupling functions that depend only on the IMF magnitude, *B*, work well [*Stamper et al*., 1999; *Lockwood*, 2003; 2013; *Lockwood et al.*, 2017]. A coupling function derived for small averaging timescale τ should, if all the physical mechanisms have been properly accounted for, integrate over time and so also work at large τ. On the other hand, a coupling function that works well at high τ may not apply on shorter timescales (the IMF *B*, as discussed above, being a good example of this). Naturally, cancellation of noise and other factors (i.e., the central limit theorem) means that correlations are higher for larger τ.

Many studies of coupling functions have concentrated on the exponents needed for each of the solar wind parameters that have been shown statistically to play a role in controlling the variation of the terrestrial response, and have used multivariate analysis to adjust the exact functional form of the combination of parameters used in the coupling function. This raises a problem. Including all of the factors with their own weighting factors and/or exponents could result in extremely good fits that are, nevertheless, statistically meaningless as each additional fit parameter reduces the statistical significance of the correlation. Such fits can have limited, and in extreme cases, no predictive capability. This pitfall is called “over-fitting” and it is a serious, but often under-appreciated, problem in multiple regression analysis of geophysical time series that have internal “geophysical” noise. Overfitting refers to the situation when a fit has too many degrees of freedom and starts to approximate the noise in the training subset, which is not robust throughout the whole dataset. This is a recognized pitfall in areas where quasi-chaotic behaviors give large internal noise such as climate science and population growth (see, for example, *Knutti et al.* [2006] and *Knape and de Valpine* [2011]) but is not recognized as widely in space physics, possibly because systems may have been viewed as being somewhat more deterministic and with lower internal variability. To help guard against over-fitting, we favor the physics-based approach to coupling function derivation by *Vasyliunas et al.* [1982] which is described in the next section. This approach yields a single fit parameter, the coupling exponent α, which arises from an unknown dependence on the solar wind Alfvén Mach number but which influences the dependence of the coupling function on solar wind speed, V_{SW}, number density N_{SW}, mean ion mass, m_{SW} and the interplanetary field B – all in a consistent way. Despite having just the one free fit parameter, extremely high correlations with geomagnetic indices (often exceeding 0.99 for annual timescales) can be derived for some geomagnetic indices.

We here concentrate on the optimum coupling function to predict the *am* geomagnetic index [*Mayaud*, 1980]. This is a "range" index, being based on the range of variation of the more variable horizontal component of the field in 3-hour windows. This range is used to give a *K*-value, as introduced by *Bartels et al.* [1939]. The stations used to compile the *am* index are situated at sub-auroral latitudes close to corrected geomagnetic latitude $\Lambda_{CG} = 50°$. They are grouped into longitude sectors, with 5 such groups in the Northern hemisphere, and 4 in the Southern. The *K* indices for stations in a longitude sector are averaged together and the result is converted into a sector a_K value using the standard "K2aK" scale. Weighted averages of these sector a_K values are then generated in each hemisphere giving *an* and *as*, the weighting factors accounting for the differences in the longitude extents of the sectors. The index *am* is the average of the hemispheric indices *an* and *as*.

The study presented here as also been repeated using the *ap* index, generated from *K* values from a network of 11–13 stations at $|\Lambda_{CG}|$ between 44° and 60°. To generate this index, the station *K* indices are first converted into standardized K_S values, using conversion tables for each observatory that were defined by *Bartels* [1949, 1957]. The *K*p index is the arithmetic mean of the 3-hourly K_S values for the observatories employed. The *K*p values are converted into *ap* values using the standard K2aK scale that is constructed such that *ap* may be regarded as the range of the most-disturbed of the two horizontal field components, expressed in the unit of 2nT, at a station at dipole latitude of 50°. More details of the compilation of the *a*p and *am* indices are provided by *Mayaud* [1980], *Menvielle and Berthelier* [1991] and *Dieminger et al.* [1996].

We here use the *am* index in preference to *ap* because *am* is constructed using only measured *K*-values from a relatively uniform global network of stations, whereas *ap* employs a less uniform distribution of stations and relies on an empirical model (via Bartels' *K* to *K*s conversion tables) to become global in nature. Our research (reported elsewhere) has shown that although this gives more accurate annual means for *ap*, it also generates some spurious time-of-year and time-of-day variations in the response of *ap* (of amplitude about 20%) whereas the time-of-year/time-of-day response of *am* is constant to within 2.5%. All the results presented here are also found for *ap*, but uncertainties are slightly lower for *am* except on annual timescales. Any differences between results for *ap* and *am* will be noted and the results for *ap* are given in the supplementary materials.

The *am* index and its daily means, *A*m, responds strongly to substorm currents [*Saba et al.*, 1997], such that the linear correlation coefficient between all available coincident 56 annual means of the auroral electrojet AE(12) index (from 1959-2017, inclusive) and *Am* is 0.92±0.04 (significant at >99.99% level), and the correlation between the 20301 coincident daily means of AE(12) and *Am* is 0.855±0.004 (significant to at least the same level). One reason for the lower correlation for daily means is evident from the scatter plot which is linear at low disturbance levels but shows a marked non-linearity in large disturbances, with large *Am* values consistently exceeding the corresponding linearly-regressed values from AE(12) (see, for example, Figure 4 of *Adebesin* [2016] and the scatter plots presented in the supplementary material for the present paper). This is consistent with the effect on AE(12) of an extremely expanded auroral oval migrating increasingly equatorward of the fixed latitude of the ring of 12 stations that it is compiled from.

The mid-latitude planetary range indices such as *am* are particularly sensitive to the ionospheric currents that flow in the substorm current wedge. These currents are strong during substorm expansion phases, in which the energy that was extracted from the solar wind and stored as magnetic energy in the geomagnetic tail is released and deposited in the upper atmosphere. (See description of energy flow as a function of substorm cycle phase using Poynting's theorem by

Lockwood [2004]). Hence *am* is an index that we expect to depend on the energy input to the magnetosphere and so correlate well with P_α.

The substorm current wedge (that *am* responds to) forms when the near-Earth edge of the cross-tail current is disrupted at, or soon after, substorm expansion phase onset and, due to the shape of the magnetosphere, the solar wind dynamic pressure constrains the cross-sectional area of the geomagnetic tail lobe at such locations. This means that solar wind dynamic pressure plays a role in the growth-phase rise in magnetic field intensity at and around the distances down the tail (the X-coordinate in the GSM frame) of the inner edge of the cross tail current. Hence the dynamic pressure (as well as the open magnetospheric flux in tail) controls the magnetic sheer and hence the cross-tail current at these locations [*Lockwood*, 2013]. (Note that this is not the case in the far tail (greater $-X$) where the tail lobe area expands as the open flux increases and the cross-tail current is constant, being set by the static pressure of the solar wind). This role of solar wind dynamic pressure in setting the cross-tail current (the current that becomes deflected into the substorm current wedge at substorm onset) means that the *am* index should show a strong variation on V_{SW}^{n}, where the exponent n is close to 2 [*Lockwood*, 2013]. The *K*p (and hence *ap*) and *am* indices are also known to be good indicators of the strength of convection in the magnetosphere-ionosphere system [*Thomsen*, 2004]. That *am* is good indicator of both substorm currents and convection is also expected from the Expanding-Contracting Polar Cap model in which convection is understood as the net effect of the opening of field lines by magnetopause reconnection (associated with the directly-driven current system, and the storage (growth) phase of the substorm cycle) and their re-closing by reconnection in the cross-tail current sheet in expansion and recovery phases (associated with the unloading phase of the substorm cycle) [*Lockwood et al.*, 1990; *Lockwood and Cowley*, 1992; *Cowley and Lockwood*, 1992; *Milan et al*., 2003; 2008].

1.2. Power Input into the magnetosphere

Vasyluinas et al. [1982] estimated the power input from the solar wind into the magnetosphere, P_α. As discussed below, a key point about the coupling function that this theory yields is that it has just one free fit parameter, the various solar wind parameters being linked by the theory used. To demonstrate this, the formula used by *Vasyluinas et al*. [1982] is:

$$P_\alpha = (\pi L_o^2) \times (m_{sw} N_{sw} V_{sw}^3/2) \times (t_r) \qquad (1)$$

where L_o is the mean cross-sectional radius of the magnetosphere, such that the geomagnetic field presents an area πL_o^2 to the solar wind flow and $(m_{sw} N_{sw} V_{sw}^3/2)$ is the kinetic energy flux of the particles, which is the dominant energy flux in the solar wind (where m_{sw} is the mean ion mass, N_{sw} the number density and V_{sw} the speed). The term t_r is a dimensionless transfer function, the fraction of the incident energy flux that is transferred to inside the magnetosphere.

Assuming a hemispheric shape to the dayside magnetosphere, pressure balance at the nose of the magnetosphere gives [e.g. *Farrugia et al.,* 1989]:

$$L_o = k_1 (M_E^2/P_{sw} \mu_o)^{1/6} \qquad (2)$$

where k_1 is a geometric factor for a blunt-nosed object, M_E is Earth's magnetic dipole moment, μ_o is the magnetic constant, and P_{sw} is the solar wind dynamic pressure, given by

$$P_{sw} = m_{sw} N_{sw} V_{sw}^2 \qquad (3)$$

Vasyluinas et al. [1982] adopted the dimensionless transfer function:

$$t_r = k_2 M_A^{-2\alpha} \sin^4(\theta /2) \qquad (4)$$

where α is the "coupling exponent" (the one free fit parameter used), θ is the clock angle that the IMF makes with the north in the GSM frame of reference (so $\theta = \tan^{-1}(B_{yM}/B_{zM})$, where B_{yM} and B_{zM} are the Y and Z components of the IMF in the GSM frame) k_2 is a constant (which below is combined with other constant factors) and M_A is the Alfvén Mach number of the solar wind flow given by

$$M_A = V_{SW} (\mu_o \, m_{sw} \, N_{sw})^{1/2} / B \qquad (5)$$

Substituting equations (2) - (5) into (1) yields

$$P_\alpha = (k_1 k_2 . \pi/2\mu_o^{(1/3-\alpha)}) M_E^{2/3} m_{sw}^{(2/3-\alpha)} \times B^{2\alpha} N_{sw}^{(2/3-\alpha)} V_{sw}^{(7/3-\alpha)} \sin^4(\theta/2) \qquad (6)$$

The predictions from equation compare well with the global MHD simulations of the energy transfer into the magnetosphere by *Wang et al.* [2014]: their results agreed best with P_α for a value of α of 0.34.

The interplanetary parameters B, N_{sw} and V_{sw} have been routinely measured by near-Earth, interplanetary craft. We here use 1-minute samples of these parameters from the Omni-2 dataset [*Couzens and King*, 1986; *Hapgood et al.*, 1991; *King and Papitashvili* , 1994; 2005]. The error introduced into P_α / P_o (where P_o is the mean of P_α over a long reference period) by using a constant value of the mean ion mass, m_{sw} , is less than 5% [*Lockwood et al.*, 2017a]. However, we here reduce that uncertainty by employing all the available information and compute the mean ion mass. The full mass spectrum of the solar wind is not routinely available and we here use the ratio of the number densities, of the two dominant components, protons and Helium ions (He^+ and He^{++}), and neglect the trace higher-mass ions [*Kaspar et al.*, 2007]. Using the typical heavier ion abundances given by *Bochsler* [1987], the effect of neglecting ions heavier than Helium on m_{sw} introduces an uncertainty into the correction of 1.8%. However, we here reduce this uncertainty further by using P_α as a ratio of the mean for the whole period (1996-2016, inclusive), P_o.

The helium abundance ratio, N_{He}/N_p, is relatively constant near 0.04 in high-speed streams, but varies in phase with the sunspot number in slow wind, between about 0.01 and 0.04 [*Wang*, 2016]. Hourly means of N_{He}/N_p are available in the Omni2 dataset from 1972 and are here interpolated to the center times of the 1-minute samples of B, N_{sw} and V_{sw} using Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation [*Fritsch and Carlson*, 1980]. As a check, linear interpolation was also used and the deviation of the results from the two methods quantified to check that it remained small (<10%) and hence that the interpolations were reasonable. The mean ion mass is then given by

$$m_{sw} = m_p \{1 + (4N_{He}/N_p)\} / \{1 + (N_{He}/N_p)\} \qquad (7)$$

where m_p is the mass of the proton.

There is a clear but weak solar cycle variation in this m_{sw} estimate, with the largest annual mean being 1.139 a.m.u. (a deviation from the mean of +3.4%) in the year 2000 (sunspot maximum) and the lowest value being 1.050 a.m.u. (a deviation of −4.7%) in the low sunspot minimum year of 2009. This is allowed for by our procedure but the effects of any corresponding variation in the heavier ion fraction (a.m.u. > 4) is not: from the range in the relative abundances given by *Bochsler* [1987] we estimate this introduces an uncertainty into P_α / P_o of $\leq$0.4%.

By taking the results for P_α as a ratio of the mean for the whole period, P_o, we also remove the dependence on the constant k_3 with the assumption that it remains constant. Earth's magnetic dipole moment M_E varies on long timescales and we use the dipole moment variation provided by the IGRF-11 model [*Thébault et al.*, 2015] to account for any drift.

The coupling exponent α influences almost all the factors in equation (6) but is unknown and has to be derived empirically. *Wang et al.* [2014] present a review of previously derived values which, as noted by *Finch and Lockwood* [2007], appear to also depend on averaging timescale τ (they found an optimum α of 0.4 at τ = 3hrs, falling to α = 0.3 at τ = 1 yr). The results also appear to depend on which indictor of solar wind energy deposition is used: *Finch and Lockwood* were using the *am* geomagnetic index. *Murayama* [1982] also found α = 0.4 for the *am* and the AL indices and τ near 1 day whereas *Stamper et al.* [1999] found α = 0.38 using the *aa* index and τ =1 yr. On the other hand, *Bargatze et al.* [1986] found α = 0.5 for the AL index and τ < 1 hr. *Tenfjord and Ostgaard* [2013] employed combinations of SuperMAG magnetometer data designed to quantify energy sinks in the terrestrial system: using 5-minute data their results were close to α = 0.5.

1.3. The effect of data gaps on studies of coupling functions

Since 1995, we have available a much more continuous data series on the near-Earth interplanetary medium from the WIND, ACE and DSCOVR satellites. However, before this date the data series contained many data gaps with a broad (and bimodal) spectrum of durations [*Finch and Lockwood*, 2007]. An analysis of interplanetary data availability is presented in section 2 of this paper.

Data gaps can influence correlation studies [*George et al.*, 2015]. For many applications such as spectral analysis, it is desirable or necessary to fill the data gaps and a number of methods for doing this are available, but one has to remain aware of the implications of the method used for application in question [*Sturges*, 1983; *Wynn and Wickwar*, 2007; *Henn et al.*, 2013, *Munteanu et al.*, 2016]. Such gap filling techniques have been applied to solar wind data by, for example, *Kondrashov et al.* [2010; 2014] but many require a proxy dataset for either interpolation or testing purposes. This usually involves using a terrestrial space weather disturbance index as the proxy, which means the coupling function is assumed and used to help fill the data gap and so such techniques must not be used in the context of deriving a coupling function. Correlation studies, such as between a solar wind coupling function and a terrestrial space weather response, are not influenced by data gaps at basic time resolution, other than through the loss of correlation significance through the reduced number data point pairs available. The question then arises as to how best to deal with data gaps when the basic time resolution data are averaged into longer intervals τ. Some of the more recent studies of solar wind-magnetosphere coupling have been able to reduce the problem of data gaps by employing only data from after 1995 when data became much more continuous [e.g., *Luo et al.*, 2013; *Temerin and Li*, 2006], despite the drawback that this reduces the number of years covered by the study. Some studies have placed tight restrictions on data gap occurrence [e.g., *Teodorescu et al*., 2015] but these tend to be severe and also greatly reduce the available data such that only some spectral studies are possible and coupling functions are studied through of a series of events [e.g. *Bargatze et al.,* 1985; 1986]. This limits the potential for noise reduction by averaging (which makes correlation coefficients increase with averaging timescale τ) and reduction of the influence of others factors such as ionospheric conductivity [e.g., *Nagatsuma*, 2006; *Luo et al.*, 2013] . Many studies continue to use the pre-1995 data, despite its severe problem of many and long data gaps (as shown in Figures 2 and 3 and discussed in section 2). In a great many of these studies, the averages of all available data have been taken on the (usually tacit) assumption that the effect of data gaps will average out. This philosophy is often introduced by the use of the Omni or Omni2 interplanetary datasets, which supply averages for a requested τ without a limit on the data availability (so in extreme cases a mean is given even if only one data sample is available in the interval). *Lockwood and Finch* [2007] made allowance for data gaps by introducing the solar wind data gaps into the geomagnetic data sequence that they correlated with and showed that the assumption that data gaps can be neglected is often invalid: hence they argued for piecewise removal of the geomagnetic data at times corresponding to the interplanetary data gaps (i.e. allowing for an appropriate lag) before the averaging of both into the intervals of duration τ. For P_α values this means that if any of N_{SW}, V_{SW}, m_{SW}, B or θ were missing, the corresponding geomagnetic data were piecewise masked out. This ensures that the averages contain only the corresponding data in the two data sets. Although they argued that this was the best approach, [*Finch and Lockwood*, 2007] also cautioned that data gaps could still have an influence on the optimum correlations.

On the other hand, *Svalgaard and Cliver* [2005] deployed a linear interpolation scheme on solar rotation timescales to fill the data gaps in the interplanetary data. Because it can give a small number of data points a greatly inflated weighting, *Lockwood et al.* [2006] argued this procedure was unsound.

Other interpolation procedures have been employed to fill data gaps, such as in *Temerin and Li* [2002] in their analysis of a coupling function to predict the Dst index. However, the same authors soon after moved to using only data from after 1995 and assuming that the effects of gaps averaged out, with considerable improvement of the correlation between their coupling function and the Dst geomagnetic index [*Temerin and Li*, 2006]. Others have used interpolation with restrictions, for example *Wu and Lundstedt* [1997] deployed autocorrelation-based interpolation for data gaps of up to 5 hours, and intervals with longer data gaps excluded from the study.

Note that there is also a related issue in that there are two separate ways in which mean coupling function values over an interval τ can be generated. Because, in general, there are non-linear dependencies on interplanetary parameters, the most representative of the two is to evaluate values at high time resolution and then average them (the "combine-then-average" approach). This is the most desirable approach, but this is not possible in some circumstances and the coupling function instead is computed from averages of the required parameters (the "average-then-combine" approach). For example, *Lockwood et al*. [2017a] have shown that, although compiling annual P_α values from annual means of N_{SW},V_{SW}, B and $\sin^4(\theta/2)$ is a less satisfactory approach and does slightly alter the optimum coupling exponent, α, it does not lower the correlation with *ap* or *am* on annual timescales. An example of when the "average-then-combine" method is necessary is when working with the reconstructed parameters which can only be generated as annual means [*Lockwood*, 2013; *Lockwood et al*., 2017a]: reconstructed annual means of N_{SW},V_{SW} and B generated by *Owens et al.* [2017] (with constant mean $\sin^4(\theta/2)$ and m_{SW} which only introduce small uncertainties) were employed by *Lockwood et al.* [2017a] to reconstruct the variation of P_α over the last 400 years.

From the above discussions, the variation of α with τ found by *Finch and Lockwood* [2007] is a concern because, if the physics of the relevant solar wind-magnetosphere coupling is properly described by P_α, it should simply integrate over longer τ and α should not vary. *Finch and Lockwood* did consider it a somewhat unsatisfactory result and cautioned that it may have arisen from the presence of large data gaps in the interplanetary data. In the current paper we investigate the origins of the drift in α with τ. In particular, we study the effects of data gaps by restricting our attention to the interval 1996-2016 during which the interplanetary data come from the WIND, ACE and DISCVR satellites: data gaps are present but relatively few and short in duration compared to in the interval 1974-2003 that was used by *Finch and Lockwood* [2007]. We introduce synthetic data gaps to enable us to study the effects on the derived coupling functions (and on the optimum coupling exponent α_p) of methods for dealing with gaps in the interplanetary data.

2. Datasets

Finch and Lockwood [2007] used Omni2 hourly interplanetary data for 1974-2005 (inclusive), for which the average availability of P_α was just 30%. Since data became available from ACE, Wind and DSCOVR (from 1995 onwards) this data availability has risen to 92% for 1995-2016. We here use a basis dataset of 1-minute resolution from 1996-2016 taken from the Omni2 dataset maintained by the Space Physics Data Facility, NASA/Goddard Space Flight Centre [*Couzens and King*, 1986; *King and Papitashvili*, 1994; 2005]. (We keep 1995 aside to generate independent data gap masks, as described in section 4.2).

A great many coupling function studies have employed the Omni2 hourly means, which are generated for the Omni database even if only one sample is available within the hour (but note the number of samples used can always also be downloaded). In evaluating the full effect of data gaps, the question arises as to how much data within each hour is required to generate a valid hourly mean (i.e., one which is close to the value that would be obtained with continuous sampling). The answer to this depends strongly on the autocorrelation function of the parameter in question. Figure 1a shows the autocorrelation functions (a.c.f.s) of B, N_{sw}, V_{sw}, $\sin^4(\theta/2)$ and P_α (computed here for the optimum α found in section 3, α_p= 0.5) for lags up to 3 hours. The vertical gray line is at 1 hour. It can be seen that over 1 hour, the a.c.f. of $\sin^4(\theta/2)$ has fallen considerably whereas that of V_{sw} has hardly fallen at all. It follows that the number of samples required to make a valid mean to a given uncertainty is much greater for $\sin^4(\theta/2)$ than it is for V_{sw}.

To study the effect of data gaps on 1-hour averages, we here took 1-minute Omni2 data for 1996-2016 (11046240 samples, falling in 184104 hourly intervals). We note that for some of the solar wind measurements, the basic time resolution is greater than 1-minute: for example the ACE satellite spins with 64 s. period setting the time resolution of the solar wind data. In such cases, the long autocorrelation time constant of the solar wind parameters allows interpolation into 1-minute values. However, the same is not true for the IMF data. For each parameter we searched for hourly intervals in which all 60 minutes in the hour gave a 1-minute sample made up from at least one observation such that it was not classed as "missing data" in the Omni dataset. For B, and $\sin^4(\theta/2)$ and such hours numbered 112855; for N_{sw} and V_{sw} the number was 121685 and for P_α it was 1083797 . For these hours we removed N_g 1-minute sub-samples at random (this was done 10 times for each of the hours) where N_g was increased from 1 up to 59 and the percentage error in the mean of the reduced data for that parameter and hour was evaluated by comparing it to the mean for the full data (i.e., the effect of the synthesized data gaps was determined). The distributions of these fractional errors were Gaussian and in Figure 1b the standard deviations (i.e., the one sigma error in the hourly means), ε_{hr}, are shown for each parameter as a function of data availability within the hour, $f_{hr} = 1-(N_g/60)$. As expected, the high a.c.f. of V_{sw} results in the average error introduced into hourly means, ε_{hr}, being very small and it is only 0.2% even if just one sample is available in the hour. For N_{sw} and B, this figure is larger but is still, respectively, only just above and just below 2%. Much larger errors arise from the IMF orientation factor $\sin^4(\theta/2)$, as expected for its greater variability at sub-hour timescales, shown by rapid decline in its a.c.f. with time in Figure 1a. Figure 1b shows that it is the variability in $\sin^4(\theta/2)$ that generates matching behavior for P_α, and hence sets a requirement for a sub-hour $\sin^4(\theta/2)$ availability threshold for coupling studies. From this plot we can find the fraction of samples f_{hr} that is needed in the hour to give a specified 1-σ uncertainty in the hourly means for that parameter, ε_{hr}. We here study the 2% and 5% uncertainty levels. Figure 1b shows that $\varepsilon_{hr} < 2\%$ is achieved for P_α if $f_{hr} \geq 0.96$ for $\sin^4(\theta/2)$ observations and $f_{hr} \geq 0.15$ for N_{sw} data. On the other hand, $\varepsilon_{hr} < 5\%$ achieved for P_α if $f_{hr} \geq 0.82$ for $\sin^4(\theta/2)$ observations and for all other parameters just one sample is adequate.

Figures 2 and 3 present an analysis of the data gaps in the interplanetary data. To avoid being classed as a data gap requires that all the parameters required to compute P_α are available, meaning we require a valid mean of B, B_{yM}, B_{zM}, n_{He}/n_p, N_{sw}. The definitions used of "valid" are discussed below. For Figure 2 we make an exception to this because data on the helium ion fraction n_{He}/n_p are not routinely available before 1972 and many coupling functions do not make use of this parameter, and even in many implementations of P_α it is assumed to be constant because such an assumption introduces only a small uncertainty (one that is comparable with other observational uncertainties [*Lockwood et al.*, 2017a]). Hence for the purpose of Figures 2 and 3, missing data on n_{He}/n_p does not generate a data gap before 1972.

In both Figures 2 and 3, and all subsequent places in the text, we employ hourly means that we have made from the downloaded 1-minute Omni data. In Figure 3 we require enough 1-minute samples in the hour to reduce the statistical uncertainty (as shown by Figure 1) to a given level whereas in Figure 2 we use the criterion of requiring just a single 1-minute sample in the hour to generate an hourly mean in order to investigate the effects of that criterion. To distinguish the two later in this paper, the latter cases are always identified using the phrase "using the Omni criterion" (because the Omni data give an hourly mean value even if only one sample in the hour is available). All other hourly means refer to the former definition. From the analysis presented in Figure 1, the Omni criterion is certainly adequate for V_{sw} and generates acceptable average uncertainties of $\varepsilon_{th} \approx \pm 2\%$ for B and N_{sw}. However, for the IMF orientation, and hence coupling function, the errors in using a single sample are $\varepsilon_{th} \approx \pm 15\%$. Comparison of Figures 2 and 3 investigates the effect of criteria on the occurrence of data gaps because, whereas Figure 2 is for the Omni criterion, Figure 3 uses the criterion that $\varepsilon_{th} \leq 5\%$ (requiring $f_{hr} \geq 0.82$ for the IMF orientation data) and $\varepsilon_{th} \leq 2\%$ (requiring $f_{hr} \geq 0.96$ for the IMF orientation data and $f_{hr} \geq 0.15$ for the solar wind data).

The availability of hourly samples in calendar years, f, is shown for the Omni criterion in Figure 2a. The large rise in f caused by the advent of data from ACE and WIND in 1995 is marked by the vertical green line. Figure 2b shows the spectrum of data-gap durations by giving the probability $p(L)$ of being in a data gap of duration L, as a function of $\log_{10}(L)$, evaluated in bins 1 hour wide. Note that at large L before 1995 (in blue) the spectrum become discrete as there were often just single occurrences of a data gap at that L, but $p(L)$ increases linearly with L for such cases. Figure 2b shows that data gaps at almost all L (but particularly at large L) were considerably rarer after 1995 than before. Figure 2c shows the information as the cumulative probability distribution, which shows clearly the durations of gaps that contribute most to the loss of data. Using the Omni criterion, the total probability for being in a data gap for 1966-1994 is 68% (so the average availability is $<f>$ = 32%) and for 1995-2016 it is 8% ($<f>$ = 92%). Remember from Figure 1b, this gives average errors in the hourly energy input estimates, P_α, of $\varepsilon_{th} \approx \pm 15\%$.

Figure 3 is the same as Figure 2, but for two tighter criteria on what constitutes a valid mean, namely $\varepsilon_{hr} < 5\%$ and $\varepsilon_{hr} < 2\%$. The red and blue lines require IMF data to have an availability of 82% within in each hour (giving $\varepsilon_{th} \leq 5\%$) and this raises the total probability for being in a data gap for 1966-1994 to 76.1% (so the average availability is $<f>$ = 23.9%) and for 1995-2016 it is 16.2% ($<f>$ = 83.8%). Comparison of the red and blue lines in Figures 2a and 3a shows that the additional requirement reduces the availability in most years before 1995 and makes some years around 1980 unusable. After 1995, it amplifies somewhat the trend for loss of data towards the end of the period when the main source of data was the ACE satellite. Comparing parts (b) and (c) of the two figures show that data gaps of all durations are increased by the extra criterion before 1995 but that after then the main effect is the loss of individual hours with data gaps of duration L < 1day. The pink and cyan lines in Figure 3 are for $\varepsilon_{th} \leq 2\%$. The cyan line in Figure 3a shows that this relatively small (3%) gain in the accuracy has a relatively small effect on the availability of valid data before 1995 ($<f>$ = 18.5%) but after then it has a somewhat greater effect, lowering the average availability to $<f>$ = 70.0%. Figure 3b and 3c show that the dominant effect of the lower error threshold is to introduce many short-duration data gaps.

Hence a compromise needs to be struck between the error in hourly P_α means, ε_{th}, and the average availability of those hourly means $<f>$. We here adopt and uncertainty limit of $\varepsilon_{th} \leq 5\%$

which gives the annual availability and data gap spectra after 1 January 1995 shown by the red lines in Figure 3.

The *am* data are continuous and are generated by L'École et Observatoire des Sciences de la Terre (EOST), a joint unit of the University of Strasbourg and the French National Center for Scientific Research (CNRS) institute, on behalf of the International Service of Geomagnetic Indices (ISGI). The data available from http://isgi.unistra.fr/data_download.php.

3. Optimum coupling function as a function of timescale

The hourly interplanetary data formed from hours with more than 82% of 1-minute IMF orientation samples (giving $\varepsilon_{th} \leq 5\%$) were averaged into independent intervals of duration τ. The process was repeated 2922 times as τ was varied between 3hrs and 1average year (365.25 days = 8766hrs) in steps of 3hrs (in other words, the values of τ used were 3hrs, 6hrs, 9hrs … up to 8766hrs). Before the *am* data were similarly averaged, values co-incident in time to data gaps in the interplanetary data were masked out using the procedure of *Finch and Lockwood* [2007] and not included in the means. Note that the data in the Omni2 dataset are lagged by the predicted propagation time between the observing satellite and the nose of the magnetosphere (see description of the procedure by King and Papitashvili available at https://omniweb.gsfc.nasa.gov/html/omni_min_data.html).

The averages of the input power for a given τ, $<P_\alpha>_\tau$ were computed using equation (6) for values of the coupling exponent α between 0 and 1.5 in steps of 0.01. The linear correlation coefficient between $<P_\alpha>_\tau$ and $<am>_\tau$ (with data gaps introduced into *am* to that match those in the P_α data) at each α, $r(\alpha)$, was then determined. Figure 4 presents an example of the results for annual means (i.e. for τ = 1 yr). Figure 4(a) shows the correlogram of $r(\alpha)$ as a function of α. The peak $r(\alpha)$, r_p = 0.990, is at $\alpha = \alpha_p$ = 0.44 (marked by the vertical green line). Figure 4(b) shows the significance $S(\alpha)$ of the difference between $r(\alpha)$ and its peak value, r_p, evaluated using the Meng-Z test of the significance of the difference between two correlations [*Meng et al.*, 1992]. Sometimes also called Steiger's Z test, this is a variant of the Fisher-Z test [e.g. *Asuero et al.*, 2006] that makes allowance for inter-correlation between the comparison time series. We use a one-sided version of the test against the null hypothesis that $r(\alpha)$ is not lower than $r(\alpha_p)$.

By definition, $S(\alpha) = 0$ at $\alpha = \alpha_p$ and increases at α away from the peak as $r(\alpha)$ falls. The vertical red and blue lines are where $S(\alpha)$ rises to 0.68 and so mark the 1-σ uncertainties in the α_p value (giving the optimum α to be $\alpha_p = 0.42$, with 1-σ uncertainties of +0.11 and −0.08 in this case for τ =1yr). Figure 4e shows the time series of $<am>_{\tau=1yr}$ (in black) and the best-fit linear regression of $<P_\alpha>_{\tau=1yr}$ to $<am>_{\tau=1yr}$ (in mauve) for $\alpha = \alpha_p$, the best fit being $s<P_\alpha/P_o>_{\tau=1yr} + c$, for linear coefficients of $s = 13.07 \pm 0.32$ and $c = -2.59 \pm 0.33$, where P_o is the mean of P_α over the whole 1996-2016 dataset and is a convenient normalizing factor. The corresponding scatter plot is shown by the points in Figure 4f and the best-fit linear regression in mauve.

Figures 4(c) and 4(d) show an alternative procedure for determining α_p and its uncertainty. The r.m.s. deviation of *am* from the best-fit linear regression of $<P_\alpha/P_o>_{\tau=1yr}$ (illustrated in Figures 4e and 4f for $\alpha = \alpha_p$), ΔAm_{rms}, is plotted as a function of α in Part (c). This r.m.s. fit residual is a different metric of the same thing as the correlation coefficient (i.e., the level of agreement between Am and $<P_\alpha/P_o>_{\tau=1yr}$), but can be used with different statistical tests. As expected, the two procedures give the same α_p to within ±0.05 (the resolution of the tests deployed as α was incremented in steps of 0.1). Figure 4d looks at the significance S of the difference between ΔAm_{rms} at $\alpha = \alpha_p$ and the values at all other α. This is evaluated using a 2-sample F test of the variances of the distributions of the fit residuals [*Snedecor and Cochran*, 1989]: this allows us to estimate the α values at which ΔAm_{rms} is larger than the minimum by an amount that is significant at a specified level. We use the one-tailed version of the test against the null hypothesis that ΔAm_{rms} at general α is not greater than that $\alpha = \alpha_p$. The F–test is a parametric test (i.e., it assumes Gaussian distributions) and is particularly sensitive to the effect of distortions from that form. Furthermore, a non-normal distribution of fit residuals indicates that the linear regression may also be invalid as it is a violation of one of the assumptions of the regression. Hence, although this is a more direct test of the fit, it is particularly sensitive to the normality of the distributions and it is important to test if the fit residuals are normally distributed. This is done in the quantile-quantile (q-q) plot against a normal distribution in Figure 4g. For a normal distribution, the points would lie along the diagonal line and Figure 4g shows the residuals' distribution is quite close to Gaussian in this case. (However, there is a slight "S-shape" which is discussed below in relation to Figure 5 in which it is more pronounced). The pink and cyan vertical lines in Figure 4c and 4d are at S = 0.68 (the 1-sigma level) for this test and comparison with Figures 4a and 4b shows that the F-test gives slightly larger uncertainties in the best value of α than does the Meng-Z test. Note that the Meng-Z test is also parametric, but is not as sensitive to the assumption of Gaussian distributions and gives valid results even if the distributions tested are only approximately normal.

It is important to stress that we are testing the significance of the difference between the agreement between the optimum fit of $<P_\alpha/P_o>_{\tau=1yr}$ to Am and the fits for other, less optimum, $<P_\alpha/P_o>_{\tau=1yr}$ data series. This is different to testing the significance of the less-good fits. Because the optimum fits are so good (correlation coefficient very close to unity and r.m.s. fit residual close to zero), the agreements for less-good fits are still highly significant in themselves - even when the difference to the optimum fit is found to be significant. Notice also that we here use the 1-sigma level to quantify the uncertainty in α and the estimate would naturally be larger for the 2-sigma or 3-sigma level. This is because 1-sigma is a conservative (pessimistic) estimate for our application of the uncertainties because they are here used to gauge if α is constant with τ and using the smaller uncertainties is a stricter test of the constancy of α.

Figure 5 shows the corresponding plots for τ = 1day. As expected, the peak correlation is lower and the scatter is greater. The optimum α is again α_p = 0.42 and the best-fit linear regression coefficients are $s = 12.46 \pm 0.69$ and $c = -2.06 \pm 0.40$. Thus not only is the optimum α the same for the two timescales illustrated by Figures 4 and 5, but the regression coefficients s and c agree closely, and are the same to within the statistical fit uncertainties. In this case, the uncertainties in α are much smaller. However, the q-q plot in Figure 5g reveals a significant and systematic departure from a normal distribution of residuals, the "S-shape" pattern shape revealing a tail-heavy distribution with lower kurtosis than a Gaussian. Hence the F–test uncertainties are probably less reliable in this case. Note, however, that the minimum r.m.s. fit deviation again gives the same best-fit α and similar uncertainties as the peak correlation method and, as the sensitivity to non-normal distributions is different for the two tests, we can infer the effect of this uncertainty on the uncertainties is small.

Figure 6a studies the full evolution of the $r(\alpha)$ correlograms with averaging timescale τ between 3 hours and 1 year, (i.e. averaging intervals of duration τ = 3, 6, 9, …. 8766 hours, a total of 2922 different durations). The correlation coefficient is color-coded as a function of τ along the horizontal axis and α along the vertical axis. The middle white line gives the α_p value which yields the peak correlation r_p (given by the vertical green lines in the examples presented in Figures 4 and 5) and the two black lines give the $\pm 1\sigma$ uncertainty points from the Meng-Z test (given by the vertical red and blue lines in Figures 4 and 5). Figure 6b compares the uncertainty bands from the two methods or computing the uncertainty, as a function of averaging timescale, τ, and using the same line colors as Figures 4 and 5. It can be seen that they give very similar results, except at lower τ when there are concerns about the normality of the distributions, particularly for the F test. We here use the results from the Meng-Z test and the general similarity of the F test results gives us confidence in these uncertainty estimates. Figure 6c plots the corresponding variation with τ of the peak correlation r_p.

Range indices, such as *am*, cannot easily be generated for $\tau < 3$hrs. In addition, at such timescales the correlations become complicated by lag times with the directly-driven system responding within about 5 min. of changes in P_α arriving at the dayside magnetopause and the storage-release system responding, typically, after between 30 min. and an hour (see section 1). Hence the analysis becomes more complex and depends critically on the combination of current systems to which a given index responds. Nevertheless the near-constant α found for τ between 3 hours and 1 year in figure 6 is found to apply at $\tau < 3$hrs for indices that respond primarily to the auroral electrojet and the substorm current wedge. Figure 8 shows an example of the analysis used to test this at τ = 1 min. We use the 1-min. resolution SME index generated by the SuperMAG project [*Newell and Gjerloev* , 2011]. This index is equivalent to the AE index and responds strongly to the nightside auroral electrojet, but is compiled from the large SuperMAG magnetometer network and so does not suffer the nonlinear effect due to the limited latitudinal coverage of the AE(12) stations (see supplementary materials). Figure 8 shows the correlation between P_α and SME for τ = 1 min. as a function of coupling exponent, α (horizontal axis) and time lag, Δt (vertical axis). Positive Δt is when P_α is lagged, i.e. the SME variation follows P_α. The peak correlation is r_p = 0.705 at α = 0.46 and Δt = 37 min. The plot suggests that α for this

τ is slightly higher than 0.44 (marked by the vertical white dashed line) but inside the white contour the difference between r and r_p is not significant at even the 1-σ level. Hence we cannot regard the difference between r_p for τ = 1min. and 0.44 as significant. However, it is quite possible that α is indeed slightly greater than 0.44 in this case because SME is influenced by both the directly-driven and the storage/release systems and analysis of indices such as AU that respond primarily to the directly-driven system find a dependence on V_{sw}^{n} with a smaller exponent n than for the storage-release system. We do know that the dominant response of SME is due to the storage-release system because the optimum lag is Δt = 37±3 min., which we equate to the average duration of substorm growth phases for strong solar wind forcing [*Li et al.*, 2013]. Note that the correlation is large (>0.65) for a wide range of lags Δt (between about 15 min and 1 hour) and this great variability in the length of the substorm growth phase will have reduced the peak correlation at Δt = 37min whereas it would have had no effect on 3-hourly means. This, and equivalent tests at other τ, show that α = 0.44 is consistent with the best correlation at all τ between 1 min and 1 year.

Figure 6 shows that, at all τ, the optimum α is consistent with a constant value of 0.44. This contrast somewhat with the results of *Finch and Lockwood* [2007], who found an optimum α of α_p = 0.4 at τ = 3 for the *am* index hours which fell to α_p = 0.3 at τ = 1 year. The major difference between the present study and that by *Finch and Lockwood* is that, as shown by Figure 2, much of the interplanetary data used by *Finch and Lockwood* contained many data gaps of a wide range of durations whereas we here employ data from 20 years that has a relatively small number of relatively short data gaps. P_α is designed to be an estimate of the power input that drives the substorm cycle and the geomagnetic am index has been shown to be an excellent indicator of the geomagnetic response to that energy input on timescales longer than the substorm cycle of energy storage and release (typically 1-2 hours) [*Finch et al*., 2007; *Finch*, 2008]. Hence, there is no physical reason why the optimum α should vary with averaging timescale and so *Finch and Lockwood* considered the variation of α with τ a somewhat unsatisfactory result that may have arisen from the presence of large data gaps in interplanetary data. In the next section we introduce synthetic data gaps into the near-continuous data since the start of 1996 to investigate their effects and see if this was indeed the case.

4. The effect of data gaps on the peak correlation and the derived optimum coupling exponent

4.1. Simulation of the effects of the interplanetary data gaps before 1995

To test for the effect of data gaps we here use the interplanetary data for 1996-2016, inclusive (for which over 80% of the data meets our availability criterion of $\varepsilon_{hr} \leq 5\%$) but introduce synthetic data gaps using 500 masks. To keep the same distributions of gap durations and frequency as were observed in the earlier (pre-1995) interplanetary data series (see Figure 2), we here take the sequence data gaps in the Omni hourly dataset (30-D) years earlier where D is varied between 0 and 10 years in steps of 0.02 years (175 hours). Because of the large number of data masks considered and because of the need to repeat for the full range of α values, the tests are here restricted to annual means (i.e. τ = 1 year). For each of the 500 cases, the synthetic data gaps were then dealt with in 4 different ways:

A. The interplanetary data were averaged into a mean over interval of length τ = 3 hours only if there are 3 valid hourly means available (by our criterion of sufficient one-minute samples in the hour to give an error below 5%). The simultaneous *am* data during gaps in these 3-hourly P_α data series were piecewise removed using the procedure of *Finch and Lockwood* [2017] before both *am* and P_α were averaged into annual means. Similar piecewise removal of Dst data during solar wind data gaps was used by *Temerin and Li* [2006].
B. The full *am* data series (with no piecewise data removal) and all the available hourly P_α data were used (i.e. the presence of data gaps in P_α was ignored).
C. The full *am* data series were used and the available 1-hour P_α data were linearly interpolated to fill the data gaps before both *am* and P_α were averaged. This method was used (at 10-minute resolution) in the coupling function study by *Temerin and Li* [2002].
D. The full *am* data series were used and the P_α data interpolated to fill data gaps using the scheme adopted by *Svalgaard and Cliver* [2005]. Specifically, the hourly means were calculated and combined into daily means (in Universal Time), even if only one hourly mean was available; the 27-day Bartels rotation mean was calculated from available daily

means (again even if only one was available); if there were no data for a rotation, its mean was linearly interpolated from surrounding rotations. The average for a year was then calculated from the Bartels rotations with center dates in the year in question.

In all cases both the "combine-then-average" and the "average-then-combine" procedures were studied. This is an additional complication that applies to all 4 methods. We here term annual values of P_α generated by the former $<P_\alpha>_{1yr}$ and by the latter as $<P_\alpha>_{ann}$. Figure 9a shows the time series of $<P_\alpha>_{1yr}$ (in red) and $<P_\alpha>_{ann}$ (in blue) (both normalized to ratio of Po, their average value over the whole period) and Figure 9b the scatter plot of $<P_\alpha>_{ann} / P_o$ as a function of $<P_\alpha>_{1yr} / P_o$. The agreement is very close indeed for after 1995 (the linear correlation coefficient is 0.99, with a 2-σ uncertainty range of 0.97-0.995) but the increased scatter of the green points shows it is not so close for the data before 1996 (linear correlation coefficient 0.93, with a 2-σ uncertainty range of 0.86-0.97). Hence it appears that data gaps also play a role in creating a difference between these two ways of generating annual estimates of this coupling function.

Figure 10 shows the results of using Method B to deal with data gaps. In this method the (synthetic) data gaps are effectively ignored and the annual mean P_α data series constructed from the available data (after removal of the synthetic data gaps), whereas $<am>_{1yr}$ is constructed using all of the (continuous) *am* data. This method assumes that the effects of data gaps average out, and this is the most common way of dealing with the missing data. In Figure 10a, the red and blue lines are the correlograms (correlation coefficient r as a function of assumed coupling exponent, α) for all the 1996-2016 data (i.e. no synthetic data gaps are introduced) showing, in red, $r(\alpha)$ for *am* and $<P_\alpha>_{1yr}$ (the combine-then-average annual estimates) and, in blue, $r(\alpha)$ for *am* and $<P_\alpha>_{ann}$ (the average-then-combine annual estimates). As found by *Lockwood et al.* [2017a] the peak, r_p, is at slightly higher α for $<P_\alpha>_{ann}$ than for $<P_\alpha>_{1yr}$ (α_p = 0.44 and α_p = 0.48, marked by the vertical green and orange lines, respectively). Note that this is despite the fact that, as shown by figure 9, $<P_\alpha>_{ann}$ and $<P_\alpha>_{1yr}$ are very highly correlated (r =0.99) for this interval. The solid points are the means of the values of $r(\alpha)$ for *am* and $<P_\alpha>_{1yr}$ from the runs for the N_S = 500 different data masks, and the error bars are

±1 standard deviation. It can be seen that correlations at a given α are generally reduced compared to that for the full available data series. The mean correlation at the peak (at $\alpha = \alpha_p = 0.44$, marked by the vertical green line) is reduced from 0.99 to $r_{avp} = 0.94$ with a standard deviation for the N_S simulations of $\sigma_{avp} = 0.04$ and the optimum α is increased to 0.48, compared to the 0.44 obtained without data gaps. The gray histogram in Figure 10(b) shows the distribution of the derived α_p values (giving the peak $r = r_p$ between $<P_\alpha>_{1yr}$ and $<am>_{1yr}$) for the $N_S = 500$ different data gap masks. It can be seen that the mode value of this distribution is $\alpha_m = 0.48$ and the median $\alpha_{0.5} = 0.43$. It must be remembered that we have introduced 500 sets of data gaps but only one would have existed in the actual pre-1996 data and so it is really the range of possible individual values in Figure 10b that we need to consider, rather than the mean, median or mode of the distribution. The standard deviation of the distribution of the 500 α_p values is $\sigma = 0.116$, the minimum to maximum range is 0.14 - 1.00 and the lower and upper 2-σ points are $\alpha_{0.05} = 0.24$ and $\alpha_{0.95} = 0.67$. It must be remembered that the only difference between the 500 simulations is when the data gaps happen, by chance, to fall and neglecting the effect of data gaps (the most commonly used procedure) could generate any one of the α values in this distribution.

Figure 10c shows the corresponding distribution of α_p values giving the peak r between $<P_\alpha>_{ann}$ and $<am>_{1yr}$ for the 500 different data gap masks. The effects of data gaps are very similar to those for $<P_\alpha>_{1yr}$, namely a very wide range of α_p values are possible, depending on when the data gaps happen to fall, and the most likely value is larger than the "true" value (by "true" in this context we mean the value obtained without the introduction of synthetic data gaps). We stress here that we have used the sequence of data gaps that actually existed in the pre-1995 interplanetary data series and so these large uncertainties apply to all studies that used such data but made no allowance or data gaps.

Table 1 summarizes the results shown by Figure 10 for Method B and compares them to the results or the other methods. The plots corresponding to Figure 10 for the other methods are included in the supplementary materials to this paper. In this section we highlight the major differences between the results.

The results or the piecewise removal of *am* data using the procedure of *Finch and Lockwood* [2007] (Method A) are surprisingly similar to those shown in Figure 10 for Method B. On average, the correlations are very slightly increased (r_{avp} is 0.95 instead of 0.94) but the distributions α_p are actually slightly broader for Method A for both $<P_\alpha>_{1yr}$ and $<P_\alpha>_{ann}$. However the distribution medians are closer to the "true" values (data gap free, i.e. 0.44 and 0.48) in both cases. Hence we can say that the procedure of *Finch and Lockwood* [2007] is an improvement over neglecting data gaps and averaging over all available data; however, it is only a very small improvement in the context of improving the peak correlation and is slightly more likely to give the correct estimate of the optimum α.

The most unsatisfactory method for dealing with the (synthetic) gaps in 1-hour P_α data is to fill them using interpolation (method C). This is here implemented using linear interpolation but results are even worse if cubic splines or PCIP are used. The peak correlations are lower, the distributions of α_p are very wide and the median and mode values of those distributions shifted away from the "true" (i.e., data gap free) value. We note this method was used (on pre-1995 data) by *Temerin and Li* [2002]; however, we also note that the same authors soon after moved to using a variant of method A and only data from after 1995 with considerable improvement of the correlation between their coupling function and the Dst geomagnetic index [*Temerin and Li*, 2006]. There are many variants of interpolation procedures. Others have used interpolation with restrictions, for example *Wu and Lundstedt* [1997] deployed autocorrelation-based interpolation for data gaps of up to 5 hours, and intervals with longer data gaps excluded from the study (a variant of method A).

Method D was used by *Svalgaard and Cliver* [2005]. This method performs better than Method C but significantly less well than Methods A and B. Hence it is better to simply ignore the effect of data gaps than use Method D. The reason is that Method D enables some datapoints to take on far too great statistical weight. For example, if there were a single 1-minute sample in a whole Bartels rotation, it would be used in method D to give the Bartels rotation mean value, thereby giving that data point a huge weighting. This extreme possibility illustrates why the Method D introduces extra noise by giving some datapoints too much weight.

4.2. For the availability of interplanetary data after 1995

The previous section shows that Method A is the best method for dealing with data gaps; however, it is only a very marginal improvement on making no allowance for data gaps (Method B). These two methods perform significantly better than the others tested, but Table 1 shows both give a considerable spread of α_p values (standard deviations $\sigma > 0.1$ and deviations of 2-σ points > 0.2). Hence data gaps in the pre-1995 will have been a large factor in producing the range of optimum α values in the literature.

In this section, we assess the effects of data gaps when the data availability rises to over 80%, as has been the case since 1995 (with $\varepsilon_{th} \leq 5\%$). To do this we use the same procedure as was used in the last section, but we have fewer data available (for 1995 and after) from which to make the masks to introduce synthetic data gaps. Hence we test the 1996-2016 dataset using 500 masks that are the occurrence of data gaps 1-*D* years earlier where *D* is varied between 0 and 1 yr in steps of 17.5 hours.

Figure 11 shows the results in the same format as Figure 10. It can be seen that the correlations for any of the 500 masks are only slightly reduced compared to the values for no additional data gaps. The optimum α is unchanged and the 500 masks give a standard deviation σ of only 0.014. It is instructive to compare this to the uncertainty in the α of peak correlation (α_p) between *am* and $\langle P_\alpha \rangle_{1yr}$ for this τ (1 year) with the +0.11 to –0.08 1-σ error from the Meng-Z test, as shown in figure 4. Hence data gaps are contributing to the uncertainty in α_p when data availability is 83% (post 1995) but are far from the dominant factor and so uncertainties introduced by the data gaps are smaller than that inherent in the correlations due to instrumental and geophysical noise in both P_α and *am*. In contrast, the test of Method A with data gap masks drawn from before 1995 (when average data availability is 30%) shows that these uncertainties rise to ±0.13 and so errors due to data gaps would be dominant.

5. Discussion and Conclusions

Figure 1 and comparison of the columns in Table 1gives a number of important insights into the effect of data gaps on the tuning of coupling functions to reproduce and predict terrestrial space weather responses.

The first is that the low auto-correlation time constant of the IMF orientation factor means that hourly samples that do not have 82% data availability within the hour cause errors greater than 5% (at the 1-σ level) in hourly values of coupling functions, such as P_α, and should be treated as bad data and removed. Note that this error could be reduced by requiring a higher availability, but this would generate a great many and longer data gaps.

The second insight is that piecewise removal of the terrestrial response index during interplanetary data gaps is the best option for dealing with data gaps although, on annual scales at least, the improvement over simply neglecting them is very small. We recommend Method A on principle and tests at lower τ (not presented here) show it sometimes performs significantly better than Method B. All the interpolation methods to fill in data gaps that we tested introduced errors and performed less well than ignoring data gaps and we recommend that interpolation should always be avoided in this context. If interpolation is to be used we recommend a "reanalysis" approach, using a physical model such as ENLIL (and not a statistical empirical model) and even then it should only be applied over appropriate timescales.

Using the dataset with the requirement that IMF availability in each hour is 82% (which limits the uncertainty in hourly P_α values to ±5% at the 1-σ level) we find that the optimum coupling exponent of $\alpha = \alpha_p = 0.44 \pm 0.02$ at all averaging timescales τ between 1 minute and 1 year. We find no significant variation of this value with τ. From equation (6), this yields dependencies on $B^{0.88}$, $V_{sw}^{1.90}$, and $(m_{sw}N_{sw})^{0.23}$. A dependence on B has been found in a great many proposed coupling functions [e.g., *Kan and Lee*, 1979; *Wygant et al.,* 1983; *Scurry and Russell*, 1991; *Temerin and Li*, 2006]. The 1-σ uncertainty on the optimum α found here is ±0.02 which allows a variation of anywhere between $B^{0.84}$ and $B^{0.92}$; however, this is not enough to explain the discrepancy with a dependence on B, nor with the result of *Newell et al.* [2007] who find a $B^{0.67}$ dependence (probably because these authors were attempting to match a wide range of terrestrial response measures). The dependence found here is close to BV_{sw}^2 and similar to that found in a great many other studies. *Lockwood* [2013] studied the optimum values of n for functions of the form BV_{sw}^{n} and found that $n = 2$ for AL, $n = 1.9$ for *aa*, $n = 1.8$ for *am*, and $n = 1.6$ for *ap*. The optimum value deduced here for *am* is $n = 1.90$ and allowing for the uncertainty in the peak α (± 0.02) the, n derived here is between 1.75 and 1.91. However, bearing in mind that *Lockwood* [2013] used annual means of data from 1966 onwards with the *Finch and Lockwood* [2007] method (A) to deal with data gaps, Table 1 shows a 1-σ uncertainty in α of ± 0.12 would apply to the *Lockwood* [2013] values which means that n could be between about 1.86 and 2.24 for AL and between 1.55 and 2.04 for *am*. Hence results of *Lockwood* [2013] for the mid-latitude station range indices *am*, ap and aa and the auroral electrojet index AL are not inconsistent with the result found here for *am* when we allow for the potential effect of data gaps.

The IMF orientation factor $\sin^4(\theta/2)$ is here assumed rather than fitted, but we note overall correlations are extremely high at all τ, rising to 0.990 for $\tau = 1$ year for *am* (and 0.997 or *ap*). However at $\tau = 1$ year this factor averages out to an almost constant number [*Lockwood*, 2013; *Lockwood et al.*, 2017a] and so the exponent used has no effect. Although $\sin^4(\theta/2)$ is the most used formulation, *Kan and Lee* [1979] used $\sin^2(\theta/2)$, *Temerin and Li* [2006] used $\sin^6(\theta/2)$, and *Newell et al.* [2007] derived $\sin^{8/3}(\theta/2)$. We have repeated our correlation study for $\sin^{i/3}(\theta/2)$ with i of [6:1:18]. Although there were very slight differences (<1%) in peak correlation, r_p, and optimum α, α_p at low τ , they were not statistically significant for any of the i tested. Hence the correlogram with i is exceptionally flat and we find no evidence that the widely-used $\sin^4(\theta/2)$ dependence is not optimum.

The derived dependence on $(m_{sw}N_{sw})^{0.23}$ is interesting. The exponent of this term would be zero if $\alpha = 2/3$, removing any dependence on the solar wind mass or number flux. This is because for $\alpha = 2/3$ the effect of an increase in mass flux ($m_{sw}N_{sw}$) on the particle kinetic energy flux in the solar wind ($m_{sw}N_{sw}V_{sw}^3$) would be cancelled by the compressional effect of the solar wind dynamic pressure ($m_{sw}N_{sw}V_{sw}^2$) and the reduction of the target area presented to the solar wind flow by the geomagnetic field. The lower α of 0.44 means that the first effect is the slightly larger of the two and there is a (weak) dependence on solar wind mass flux ($m_{sw}N_{sw}$), variations in which are dominated by the number flux (N_{sw}). Several studies have found that increased N_{sw}

increases terrestrial space weather response to IMF changes [e.g., *Lopez et al.* 2004; *Weigel,* 2010] but many coupling functions do not include a term in either ($m_{sw}N_{sw}$) or N_{sw}. *Newell et al*. [2007] also found a $(m_{sw}N_{sw})^{1/6}$ dependence, but chose to omit it from their coupling function in the interests of making it match a wide range of terrestrial responses.

Plots corresponding to those presented in this paper, but using the geomagnetic *ap* index rather than *am*, are given in the supplementary materials. These plots are very similar indeed to those for *am* in all cases. Table 2 shows how *ap* correlates slightly better with $P\alpha$ on annual averaging timescales, but *am* performs slightly better at 27 days, 1 day and 3 hours. This is consistent with an analysis of the two indices that have carried out showing that the empirical, tabular K-to-K_S conversions used to make *K*p (and hence *ap*) introduce spurious diurnal and annual variations, but these are averaged out on annual timescales, on which *ap* gains an advantage in noise suppression by averaging data from a more concentrated cluster of stations in Europe. The best fit α values are higher for *ap* than for *am* by between 0.02 and 0.06. Figure 7 investigates the implications of the precise value of α. Parts (c) and (d) of Figure 7 demonstrate that the difference in α_p for the two indices makes almost no difference to the distribution of the normalized power input to the magnetosphere, P_α/P_o, but part (b) shows it does influence the average value over the whole interval 1996-2016, P_o, and hence the absolute magnitude of the power input which is estimated. Because *am* has a more uniform time-of-day-time-of-year response pattern, we prefer the value obtained for *am*, which is $P_o = (0.38\pm0.06)\times10^{19}$ W (at the 2σ uncertainty level). However, the sensitivity of this value to the index used illustrates how difficult it is to get an absolute power input value accurately even though the variation in P_α/P_o is well defined.

Lockwood et al. [2017] have estimated the annual mean power input into the magnetosphere for all years back to 1612 from the reconstructed solar wind and interplanetary field parameters derived by *Owens et al*. [2017], and from this *Lockwood et al.* [2018a] have derived the annual means of *Ap* and AE back to this date. In the two subsequent papers in the present series [*Lockwood et al.,* 2018b; c] we begin to construct a space weather climatology by studying the distributions of space weather parameters about these averages and, in particular, how these distributions evolve with timescale. The present paper is an important first step in this because it shows that the formula for the optimum coupling function does not significantly evolve with timescale. Previous studies that suggested there was such a variation had been influenced by data gaps which have different effects on different timescales. Removing this potential complication (by showing that the optimum coupling exponent α is, in fact, independent of timescale) is a valuable first step in the construction of a useful space weather climatology.

Acknowledgments and Data

The authors are grateful to the staff of Space Physics Data Facility, NASA/Goddard Space Flight Center, who prepared and made available the OMNI2 dataset used. The data were downloaded from http://omniweb.gsfc.nasa.gov/ow.html. They are also grateful to 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. We also thank GeoForschungsZentrum (GFZ) Potsdam, Adolf-Schmidt-Observatorium für Geomagnetismus, Niemegk, Germany who generate the *ap* data. The *ap* data were downloaded from the UK Solar System Data Centre from https://www.ukssdc.ac.uk/ with updating of recent data from BGS Edinburgh http://www.geomag.bgs.ac.uk/data_service/data/magnetic_indices/apindex.html. For the SuperMAG indices data we gratefully acknowledge: Intermagnet; USGS, Jeffrey J. Love; CARISMA, PI Ian Mann; CANMOS; The S-RAMP Database, PI K. Yumoto and Dr. K. Shiokawa; The SPIDR database; AARI, PI Oleg Troshichev; The MACCS program, PI M. Engebretson, Geomagnetism Unit of the Geological Survey of Canada; GIMA; MEASURE, UCLA IGPP and Florida Institute of Technology; SAMBA, PI Eftyhia Zesta; 210 Chain, PI K. Yumoto; SAMNET, PI Farideh Honary; The institutes who maintain the IMAGE magnetometer array, PI Eija Tanskanen; PENGUIN; AUTUMN, PI Martin Connors; DTU Space, PI Dr. Rico Behlke; South Pole and McMurdo Magnetometer, PI's Louis J. Lanzarotti and Alan T. Weatherwax; ICESTAR; RAPIDMAG; PENGUIn; British Artarctic Survey; McMac, PI Dr. Peter Chi; BGS, PI Dr. Susan Macmillan; Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation (IZMIRAN); GFZ, PI Dr. Juergen Matzka; MFGI, PI B. Heilig; IGFPAS, PI J. Reda; University of L'Aquila, PI M. Vellante; BCMT, V. Lesur and A. Chambodut; Data obtained in cooperation with Geoscience Australia, PI Marina Costelloe; SuperMAG, PI Jesper W. Gjerloev. The work presented in this paper is supported by STFC consolidated grant number ST/M000885/1, the work of ML and MJO is also supported by the SWIGS NERC Directed Highlight Topic Grant number NE/P016928/1/ and of OA by NERC grant NE/P017274/1. SB is supported by an NERC PhD studentship.

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 Geod Geophys*, 51 (2), 315–331, doi: 10.1007/s40328-015-0128-2

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

Asuero, A.G., A. Sayago & A. G. González (2006) The Correlation Coefficient: An Overview, *Critical Reviews in Analytical Chemistry*, 36 (1), 41-59, DOI: 10.1080/10408340500526766

Baker, D. N., E. W. Hones, Jr., J. B. Payne, and W. C. Feldman (1981) A high time resolution study of interplanetary parameter correlations with AE, *Geophys. Res. Lett.*, 8, 179-182, doi: 10.1029/gl008i002p00179

Baker, D. N., R. D. Zwickl, S. J. Bame, E. W. Hones Jr., B. T. Tsurutani, E. J. Smith, and S.-I. Akasofu (1983) An ISEE 3 high time resolution study of interplanetary parameter correlations with magnetospheric activity, *J. Geophys. Res.*, 88 (A8), 6230–6242, doi:10.1029/JA088iA08p06230.

Balikhin, M. A., R. J. Boynton, S. A. Billings, M. Gedalin, N. Ganushkina, D. Coca, and H. Wei (2010), Data based quest for solar wind–magnetosphere coupling function, *Geophys. Res. Lett.*, 37, L24107, doi:10.1029/2010GL045733.

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., R. L. McPherron, and D. N. Baker (1986), Solar wind-magnetosphere energy input functions, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, pp. 93–100, Terrapub/Reidel, Tokyo, Japan, doi: 10.1007/978-94-009-4722-1_7

Bartels, J., Heck, N.H. and Johnston, H.F. (1939) The three-hour-range index measuring geomagnetic activity, *Terr. Magn. Atmos. Electr.*, 44, 411–454, doi: 10.1029/TE044i004p00411

Bartels, J. (1949) The standardized index Ks and the planetary index Kp, IATME Bulletin 12b, 97.

Bartels, J. (1957) The geomagnetic measures for the time-variations of solar corpuscular radiation, described for use in correlation studies in other geophysical fields, Ann. Intern. Geophys. Year 4, 227-236

Bochsler, P. (1987) Solar-wind ion composition, *Physica Scripta*, T18, 55-60, doi: 10.1088/0031-8949/1987/T18/007

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.

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

Couzens, D.A. and J.H., King (1986) Interplanetary Medium Data Book – Supplement 3, 1977 – 1985, NSSDC/WDC-A/R&S, 86-04, NASA, Greenbelt, MD, 1986, http://www.archive.org/details/nasa_techdoc_19890001401

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

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

Crooker, N. U., J. Feynman, and J. T. Gosling (1977) On the high correlation between long-term averages of solar wind speed and geomagnetic activity, *J. Geophys. Res*., 82, 1933-1937, doi: 10.1029/ja082i013p01933.

Dieminger, W., G.K. Hartmann, and R. Leitinger (1996) Geomagnetic Activity Indices, in *The Upper Atmosphere* (Eds. W. Dieminger et al.), 887-911, Springer, Berlin Heidelberg, doi: 10.1007/978-3-642-78717-1_26

Dods, J., S. C. Chapman, and J. W. Gjerloev (2015), Network analysis of geomagnetic substorms using the SuperMAG database of ground-based magnetometer stations, *J. Geophys. Res. Space Physics*, 120, 7774–7784, doi: 10.1002/2015JA021456.

Dods, J., S. C. Chapman, and J. W. Gjerloev (2017), Characterizing the ionospheric current pattern response to southward and northward IMF turnings with dynamical SuperMAG correlation networks, *J. Geophys. Res. Space Physics*, 122, 1883–1902, doi:10.1002/2016JA023686.

Etemadi, A., S.W.H. Cowley, M. Lockwood, B.J.I. Bromage, D.M. Willis, and H. Lühr (1988) The dependence of high-latitude dayside ionospheric flows on the north-south component of the IMF: a high time resolution correlation analysis using EISCAT "POLAR" and AMPTE UKS and IRM data, *Planet. Space Sci.*, 36, 471-498, doi: 10.1016/0032-0633(88)90107-9

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

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

Finch, I.D. (2008) The use of geomagnetic activity observations in studies of solar wind-magnetosphere coupling and centennial solar change, *PhD thesis, University of Southampton*, British Library Ethos ID: uk.bl.ethos.485008

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*., 35, L21105, doi: 10.1029/2008GL035399

Fritsch, F.N. and R.E. Carlson (1980) Monotone Piecewise Cubic Interpolation, *SIAM Journal on Numerical Analysis*., 17, .238–246, doi: 10.1137/0717021

George, S.V., G. Ambika, R. Misra (2015) Effect of data gaps on correlation dimension computed from light curves of variable stars *Astrophys. and Space Sci*., 360 (1), 5, doi: 10.1007/s10509-015-2516-z

Gonzalez, W.D. (1990) A unified view of solar wind-magnetosphere coupling functions, *Planet. Space Sci*., 38(5) , 627-632, 10.1016/0032-0633(90)90068-2

Hapgood, M.A., G. Bowe, M. Lockwood, D.M. Willis, and Y. Tulunay (1991) Variability of the interplanetary magnetic field at 1 A.U. over 24 years: 1963 – 1986, *Planet. Space Sci.*, 39, 411-423, doi: 10.1016/0032-0633(91)90003-S

Henn, B., M.S. Raleigh, A. Fisher, and J.D. Lundquist (2013) A Comparison of Methods for Filling Gaps in Hourly Near-Surface Air Temperature Data, *J. Hydromet*., 14(3) on pages 929-945, doi: 10.1175/jhm-d-12-027.1

Kan, J. R., and L. C. Lee (1979), Energy coupling and the solar wind dynamo, *Geophys. Res. Lett*., 6, 577-580, doi: 10.1029/gl006i007p00577

Kasper, J.C., M.L. Stevens, A.J. Lazarus, J.T. Steinberg, K.W. Ogilvie (2007) Solar Wind Helium Abundance as a Function of Speed and Heliographic Latitude: Variation through a Solar Cycle, *Astrophys. J*, 660 (1), 901 – 910, doi: 10.1086/510842

King, J. H., and N. E. Papitashvili (1994) Interplanetary Medium Data Book, Supplement 5, 1988-1993, NSSDC/WDC-A-R&S 94-08, NASA/National Space Science Data Center, GSFC, Greenbelt, Maryland, 1994.

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

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

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

Kondrashov, D., Y. Shprits, and M. Ghil (2010) Gap filling of solar wind data by singular spectrum analysis, *Geophys. Res. Lett*., 37, L15101, doi:10.1029/2010GL044138.

Kondrashov, D., R. Denton, Y. Y. Shprits, and H. J. Singer (2014), Reconstruction of gaps in the past history of solar wind parameters, *Geophys. Res. Lett*., 41, 2702–2707, doi:10.1002/2014GL059741.

Li, H., C. Wang, and Z. Peng (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, N. (2003) Twenty-three cycles of changing open solar flux, *J. Geophys. Res.*, 108 (A3), 1128, doi: 10.1029/2002JA009431

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 Rev. Solar Physics*, 10, 4, 2013. doi: 10.12942/lrsp-2013-4

Lockwood, M. and S.W.H. Cowley (1992) Ionospheric Convection and the substorm cycle, in *Substorms 1, Proceedings of the First International Conference on Substorms, ICS-1,* ed C. Mattock, ESA-SP-335, 99-109, European Space Agency Publications, Nordvijk, The Netherlands

Lockwood, S.W.H. Cowley, and M.P. Freeman (1990) The excitation of plasma convection in the high latitude ionosphere, *J. Geophys Res*., 95, 7961-7971, doi: 10.1029/JA095iA06p07961

Lockwood, M., A.P. van Eyken, B.J.I. Bromage, D.M. Willis, and S.W.H. Cowley (1996) Eastward propagation of a plasma convection enhancement following a southward turning of the interplanetary magnetic field, *Geophys. Res. Lett.*, 13, 72-75, doi: 10.1029/GL013i001p00072

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, 1999

Lockwood, M., A.P. Rouillard, I.D. Finch, and R. Stamper (2006) Comment on "The IDV index: its derivation and use in inferring long-term variations of the interplanetary magnetic field strength" by Svalgaard and Cliver, *J. Geophys. Res.*, 111, A09109, doi:10.1029/2006JA011640

Lockwood, M., H. Nevanlinna, L. Barnard, M.J. Owens, R.G. Harrison, A.P. Rouillard, and C.J. Scot (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

Lockwood, M., M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt and S. Bentley (2018a) Space Climate and Space Weather over the past 400 years: 2. Proxy indicators of geomagnetic storm and substorm occurrence, J. Space Weather Space Clim., in press, doi: 10.1051/swsc/2017048

Lockwood, M., S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt and O. Allanson (2018b) The development of a space climatology: 2. the variation of space weather parameters with timescale, *Space Weather*, submitted with present paper

Lockwood, M., S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt and O. Allanson (2018c) The development of a space climatology: 3. The evolution of distributions of space weather parameters with timescale

Lopez, R. E., M. Wiltberger, S. Hernandez, and J. G. Lyon (2004), Solar wind density control of energy transfer to the magnetosphere, *Geophys. Res. Lett*., 31, L08804, doi:10.1029/2003GL018780.

Luo, B., X. Li, M. Temerin, and S. 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.

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

Meng, X-I., R. Rosenthal and D. B. Rubin (1992), Comparing Correlated Correlation Coefficients, *Psych. Bulletin*, 111 (1), 172-175, doi: 10.1037//0033-2909.111.1.172

Menvielle, M., and A. Berthelier (1991) The K-derived planetary indices: Description and availability, *Rev. Geophys.*, 29 (3), 415–432, doi:10.1029/91RG00994.

Milan S.E., M. Lester, S.W.H. Cowley, K. Oksavik, M. Brittnacher, R.A. Greenwald, G.Sofko, and J.-P. Villain (2003) Variations in polar cap area during two substorm cycles, *Ann. Geophys*., 21, 1121-1140, doi: 10.5194/angeo-21-1121-2003

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.

McPherron R.L., G. Siscoe G and C.N. Arge (2004) Probabilistic forecasting of the 3-h Ap index, IEEE Trans. Plasma Sci. 32 1425, doi: 10.1109/TPS.2004.833387

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.

Munteanu, C., C. Negrea, M. Echim, and K. Mursula (2016) Effect of data gaps: comparison of different spectral analysis methods, *Ann. Geophys*., 34, 437–449, doi: 10.5194/angeo-34-437-2016

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

Nagatsuma, T. (2006) Diurnal, semiannual, and solar cycle variations of solar wind magnetosphere-ionosphere coupling, *J. Geophys. Res.*, 111, A09202, doi: 10.1029/2005JA011122, 2006

Newell, P. T., and J. W. Gjerloev (2011), Evaluation of SuperMAG auroral electrojet indices as indicators of substorms and auroral power, *J. Geophys. Res*., 116, A12211, doi:10.1029/2011JA016779.

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

Nishida, A. (1968), Coherence of geomagnetic DP 2 fluctuations with interplanetary magnetic variations, *J. Geophys. Res.*, 73 (17), 5549–5559, doi:10.1029/JA073i017p05549

Owens, M.J., M. Lockwood, and P. Riley (2017) Global solar wind variations over the last four centuries, *Nature Scientific Reports*, 7, Article number 41548, doi: 10.1038/srep41548

Papitashvili, V.O., N.E. Papitashvili, and J.H. King (2000), Solar cycle effects in planetary geomagnetic activity: Analysis of 36-year long OMNI dataset, *Geophys. Res. Lett.*, 27, 2797-2800, doi: 10.1029/2000gl000064.

Perreault, W. K., and S-I. Akasofu (1978) A study of geomagnetic storms, *Geophys. J. Int.*, 54, 547-573, 10.1111/j.1365-246x.1978.tb05494.x.

Reiff, P.H., R.W. Spiro, and T.W. Hill (1981) Dependence of polar cap potential drop on interplanetary parameters, *J. Geophys. Res.*, 86, 7639-7648, doi: 10.1029/ja086ia09p07639.

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

Saba, M.M.F., W.D. Gonzalez, A.L. Clua de Gonzalez (1997) Relationships between the AE, ap and Dst indices near solar minimum (1974) and at solar maximum (1979), *Annales Geophys*., 15 (10), 1265-1270.

Schatten, K. H., and J. M. Wilcox (1967) Response of the geomagnetic activity index Kp to the interplanetary magnetic field, *J. Geophys. Res*., 72 (21), 5185–5191, doi: 10.1029/JZ072i021p05185.

Scurry, L., and C.T. Russell (1991), Proxy studies of energy transfer to the magnetopause, *J. Geophys. Res*., 96 (A6), 9541–9548, doi:10.1029/91JA00569.

Snedecor, G.W. and W.G. Cochran (1989) Statistical Methods, 8th edition, Ames: Iowa State University Press, ISBN 0 8138 1561 6

Spencer, E., P. Kasturi, S. Patra, W. Horton, and M. L. Mays (2011), Influence of solar wind–magnetosphere coupling functions on the Dst index, *J. Geophys. Res.*, 116, A12235, doi: 10.1029/2011JA016780.

Stamper, R., M. Lockwood, M.N. Wild, and T.D.G. Clark (1999) Solar Causes of the Long Term Increase in Geomagnetic Activity, *J. Geophys Res*., 104, 28325-28342, doi: 10.1029/1999JA900311

Sturges, W. (1983) On interpolating gappy records for time series analysis, *J. Geophys. Res. (Oceans)*, 88 0736-9740, doi: 10.1029/jc088ic14p09736

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, 2005, doi: 10.1029/2005JA011203

Temerin, M., and X. Li, A new model for the prediction of Dst on the basis of the solar wind , *J. Geophys. Res*., 107 (A12), 1472, doi:10.1029/2001JA007532, 2002.

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

Tenfjord, P., and N. Ostgaard (2013), Energy transfer and flow in the solar wind-magnetosphere-ionosphere system: A new coupling function, *J. Geophys. Res. Space Physics*, 118 (9), 5659–5672, doi:10.1002/Jgra.50545.

Teodorescu, E., M. Echim, C. Munteanu, T. Zhang, R. Bruno and P. Kovacs (2015) Inertial range turbulence of fast and slow solar wind at 0.72 AU and solar minimum, *Astrophys. J. Lett.*, 804, doi:10.1088/2041-8205/804/2/L41, 2015.

Thébault, E., C.C. Finlay, C.D. Beggan, P. Alken, J. Aubert, O. Barrois, F. Bertrand, T. Bondar, A. Boness, L. Brocco, E. Canet, A. Chambodut, A. Chulliat, P.Coïsson, F. Civet, A. Du, A. Fournier, I. Fratter, N. Gillet, B. Hamilton, M. Hamoudi, G. Hulot, T. Jager, M. Korte, W. Kuang, X. Lalanne, B. Langlais, J.-M. Léger, V. Lesur, F.J. Lowes, S. Macmillan, M. Mandea, C. Manoj, S. Maus, N. Olsen, V. Petrov, V. Ridley, M. Rother, T.J. Sabaka, D. Saturnino, R. Schachtschneider, O. Sirol, A. Tangborn, A. Thomson, L. Tøffner-Clausen, P. Vigneron, I. Wardinski and T. Zvereva (2015) International Geomagnetic Reference Field: the 12th generation, *Earth, Planets and Space*, 67, 79, doi: 10.1186/s40623-015-0228-9

Thomsen, M. F. (2004), Why Kp is such a good measure of magnetospheric convection?, *Space Weather*, 2, S11004, doi:10.1029/2004SW000089.

Todd, H., S.W.H. Cowley, M. Lockwood, D.M. Willis, and H. Lühr (1988) Response time of the high-latitude dayside ionosphere to sudden changes in the north-south component of the IMF, *Planet. Space Sci*., 36, 1415-1428, doi: 10.1016/0032-0633(88)90008-6

Vassiliadis, D., A. J. Klimas, D. N. Baker, and D. A. 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., J. R. Kan, G. L. Siscoe, and S.-I. Akasofu (1982) Scaling relations governing magnetospheric energy transfer, *Planet. Space Sci*., 30, 359–365, doi: 10.1016/0032-0633(82)90041-1

Wang, C., J. P. Han, H. Li, Z. Peng, and J. D. Richardson (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.

Wang,Y.-M. (2016) Role of the coronal Alfvén speed in modulating the solar-wind helium abundance, *Astrophys. J. Lett.*, 833 (2) L21, doi: 10.3847/2041-8213/833/2/l21

Weigel, R. S. (2010), Solar wind density influence on geomagnetic storm intensity, J. Geophys. Res., 115, A09201, doi:10.1029/2009JA015062

Wing, S., and D. G. Sibeck (1997) Effects of interplanetary magnetic field z component and the solar wind pressure on the geosynchonous magnetic field, *J. Geophys. Res.*, 102, 7207-7216, doi: 10.1029/97ja00150

Wu, J.-G., and H. Lundstedt (1997), Neural network modeling of solar wind-magnetosphere interaction, *J. Geophys. Res.*, 102 (A7), 14457–14466, doi:10.1029/97JA01081.

Wygant, J. R., R. B. Torbert, and F. S. Mozer (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.

Wynn T.A. and V.B. Wickwar (2007) The effects of large data gaps on estimating linear trend in autocorrelated data, in *Annual Fellowship Symposium of the Rocky Mountain NASA Space Grant Consortium*, Salt Lake City, Utah, May 2007. http://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1000&context=atmlidar_conf

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.

Xu, D., T. Chen, X.X. Zhang, and Z. Liu (2009) Statistical relationship between solar wind conditions and geomagnetic storms in 1998–2008, Planet. Space Sci., 57 (12), 1500–1513. doi:10.1016/j.pss.2009.07.015

Table 1. Comparison of the performance of methods to handle data gaps or annual data. The rows, from top to bottom, give: (1) the ranking order of the accuracy the method; (2) the interval used to give data gap masks; (3) the number of masked data simulations, N_s; (4) the peak value of the average correlation coefficient, r_{avp}, from the N_s simulations, plus and minus (when $N_s > 1$) the standard deviation at that peak, σ_{avp} (shown in Figure 10a for the case of Method C by the black dot and error bar aligned with the vertical green line); (5) and (10) the value of α giving that peak, α_{avp} (marked by the vertical green line); (6) and (11) the mode value α_m of the distribution (given in Figure 10b or10c for Method C) of the N_s individual α values giving peak correlation; (7) and (12) the standard deviation of that distribution, σ; (8) and (13) the median value of that distribution, $\alpha_{0.5}$; (9) and (14) the range between the upper and lower 2-σ points of that distribution, ($\alpha_{0.05}$-$\alpha_{0.95}$). Rows (4) to (9) apply to the combine-then-average annual means, $<P_\alpha>_{1yr}$ (the distribution shown in Figure10b for Method C) and rows (10) to (14) apply to the combine-then-average annual means, $<P_\alpha>_{ann}$ (the distribution shown in Figure10c for Method C). Row (15) gives the relevant figure. The columns are for: the single analysis with no simulated data gaps (the ideal case that the methods are trying to reproduce) and Methods A-D. Method A is applied with data masks drawn from both 1995-2015 and from 1966-1994. The other methods are only used with masks from 1966-1994.

Averaging used			No added data gaps	Method A	Method A	Method B	Method D	Method C
	1	Rank	-	-	1	2	3	4
	2	Dates giving gap masks	-	1995-2015	1966-1994	1966-1994	1966-1994	1966-1994
	3	Number of gap simulations, N_s	1	500	500	500	500	500
Combine then average, $<P_\alpha>_{1yr}$	4	peak average correlation $r_{avp} \pm \sigma_{avp}$	0.990	0.989±0.002	0.96±0.04	0.94±0.04	0.92±0.06	0.86±0.09
	5	α giving peak r_{av}, α_{avp}	0.44	0.44	0.47	0.48	0.5	0.52
	6	Mode of α distribution, α_m	0.44	0.44	0.40	0.48	0.52	0.56
	7	Standard deviation, σ	0	0.014	0.132	0.116	0.150	0.239
	8	Median, $\alpha_{0.5}$	0.44	0.435	0.44	0.43	0.46	0.49
	9	5 percentile range ($\alpha_{0.05}$-$\alpha_{0.95}$)	0.44-0.44	0.404-0.445	0.24-0.67	0.27-0.65	0.20-0.71	0.16-0.96
Average then combine, $<P_\alpha>_{ann}$	10	α giving peak r_{av}, α_{avp}	0.48	0.51	0.50	0.54	0.54	0.46
	11	Mode of α distribution, α_m	0.48	0.48	0.50	0.49	0.56	0.56
	12	Standard deviation, σ	0	0.014	0.152	0.105	0.130	0.217
	13	Median, $\alpha_{0.5}$	0.48	0.476	0.49	0.50	0.51	0.58
	14	5 percentile range ($\alpha_{0.05}$-$\alpha_{0.95}$)	0.48-0.48	0.451-0.498	0.31-0.76	0.35-0.68	0.32-0.74	0.29-1.05
	15	Figure(s)	3	11 and S9	S5	10 and S6	S7	S8

Table 2. Correlations between Pα and the *ap* and *am* indices

averaging timescale , τ	correlation coefficient, *r*		2-σ range in *r*		optimum coupling exponent, α_p	
Index	*ap*	*am*	*ap*	*am*	*Ap*	*am*
1 year	0.997	0.990	0.993- 0.999	0.971 - 0.995	0.46	0.44
27 days	0.959	0.968	0.949 - 0.968	0.960 - 0.975	0.50	0.44
1 day	0.897	0.927	0.893 - 0.901	0.923 - 0.930	0.48	0.43
3 hours	0.790	0.855	0.787 - 0.793	0.853 - 0.858	0.48	0.45

Figures

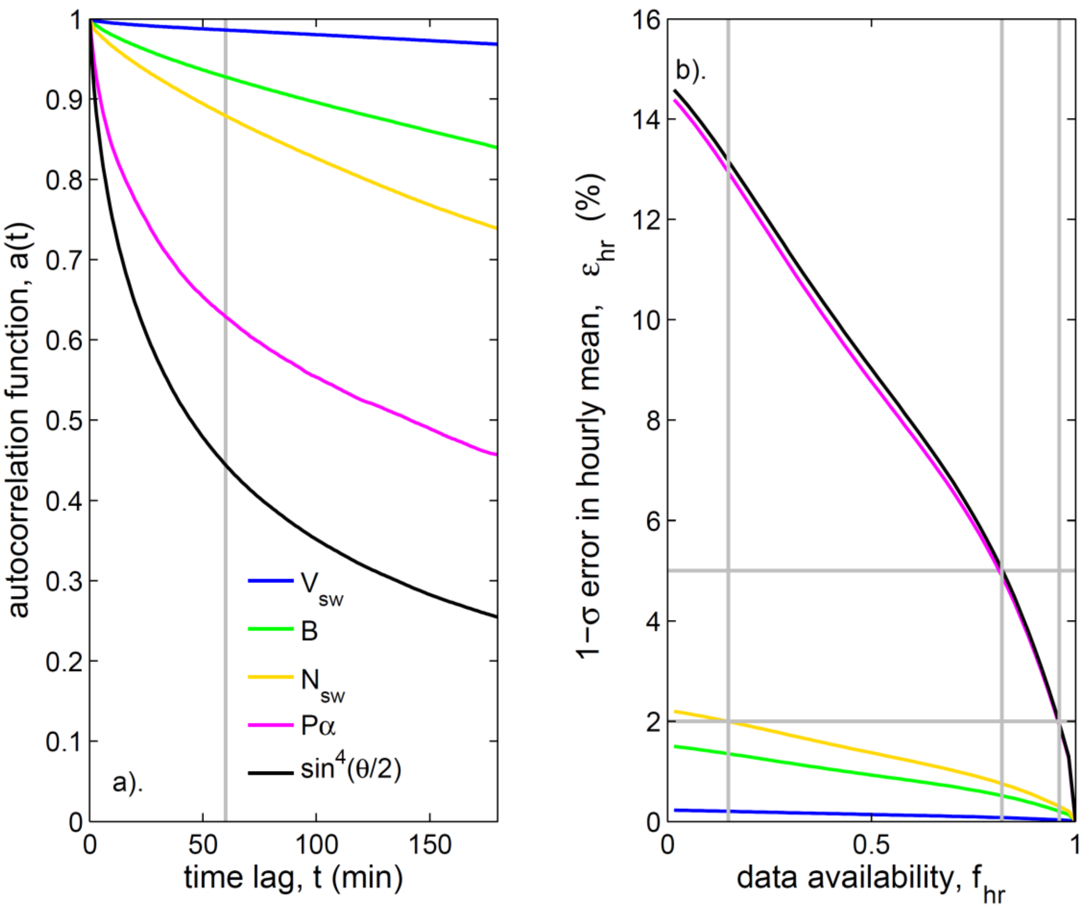


Figure 1. (a) Autocorrelation functions, $a(t)$ as a function of lag time, t, for 1-minute samples of: (blue) solar wind speed, V_{sw}; (orange) solar wind number density, N_{sw}; (green) interplanetary magnetic field (IMF) , B; (mauve) power input to the magnetosphere (for α = 0.5), P_α ; and (black) IMF orientation factor, $\sin^4(\theta/2)$. The vertical gray line marks t = 1 hour. (b) 1-σ percentage errors in hourly means, ε_{hr}, as a function of data availability within the hour, f_{hr} , for the same parameters (modulated by introducing synthetic data gaps, see text), shown using the same color scheme. The horizontal gray lines in (b) are uncertainties ε_{hr} of 5% and 2% which set threshold requirements for IMF orientation data availability of $f_{hr} \geq 0.82$ and $f_{hr} \geq 0.96$, respectively. The $\varepsilon_{hr} \leq 2\%$ condition sets an additional availability requirement on N_{sw} data of f_{hr}

≥ 0.15. These f_{hr} thresholds are shown by the vertical gray lines. Plots are constructed using 1-minute Omni samples for 1996-2016.

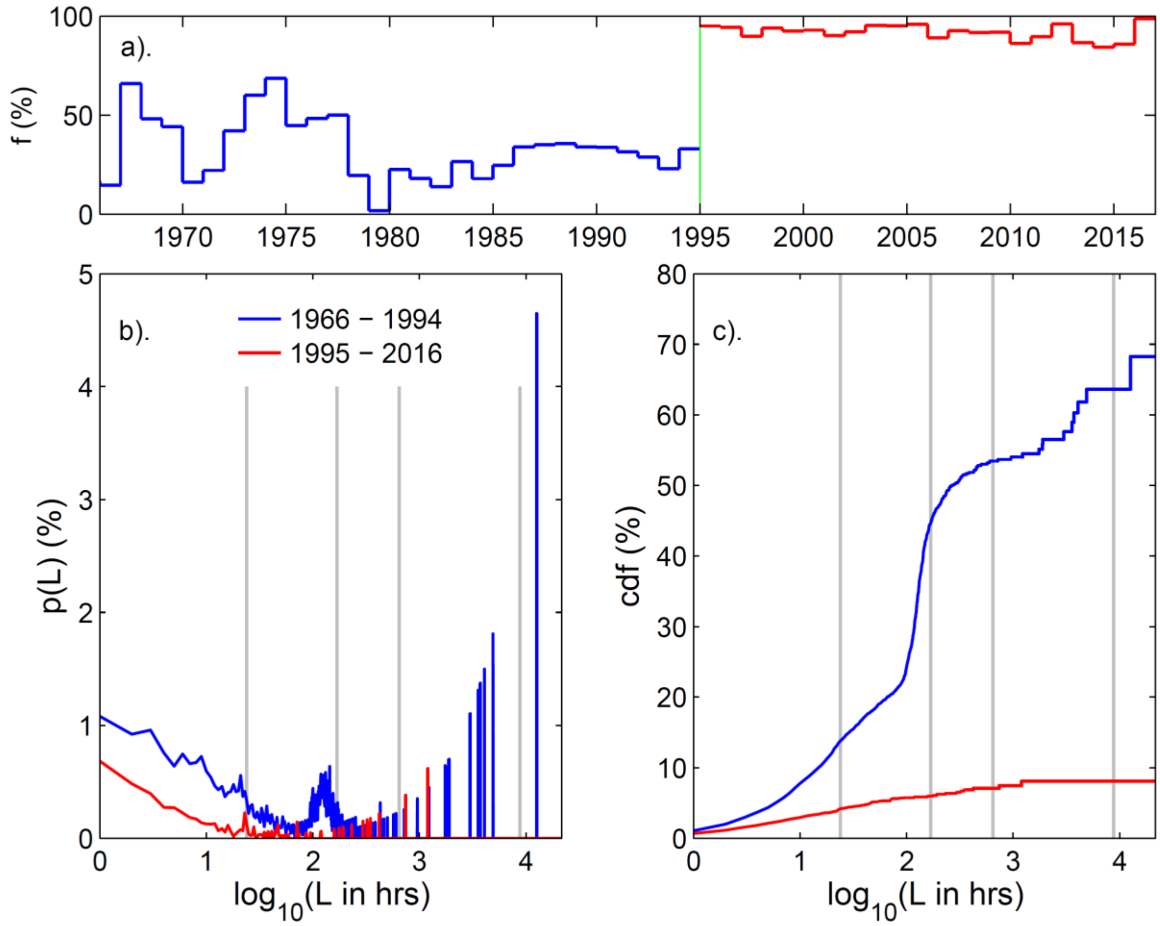


Figure 2. Analysis of data gap occurrence in the P_α interplanetary data. Data in blue are for 1966-1994 (inclusive) and in red are for 1995-2016 (inclusive). Data gaps are here defined by the absence of any sample in one-hour intervals, as used in compiling Omni2 data hourly samples. (a) the availability f of 1-hour means in calendar years. (b) The probability of being in a data gap of duration L, $p(L)$, shown as a function $\log_{10}(L)$ for L between 1 hour and 1 year, for before and after 1 January 1995 (blue and red lines, respectively). (c) The cumulative probability distribution as a function of $\log_{10}(L)$. The total probability for being within a data gap for 1966-1994 is 67% (so the average data availability is $<f>$ = 33%) and for 1995-2016 it is 8% ($<f>$ = 92%). Gray vertical lines in (b) and (c) are at 1 day, 7 days, 27 days and 1 year.

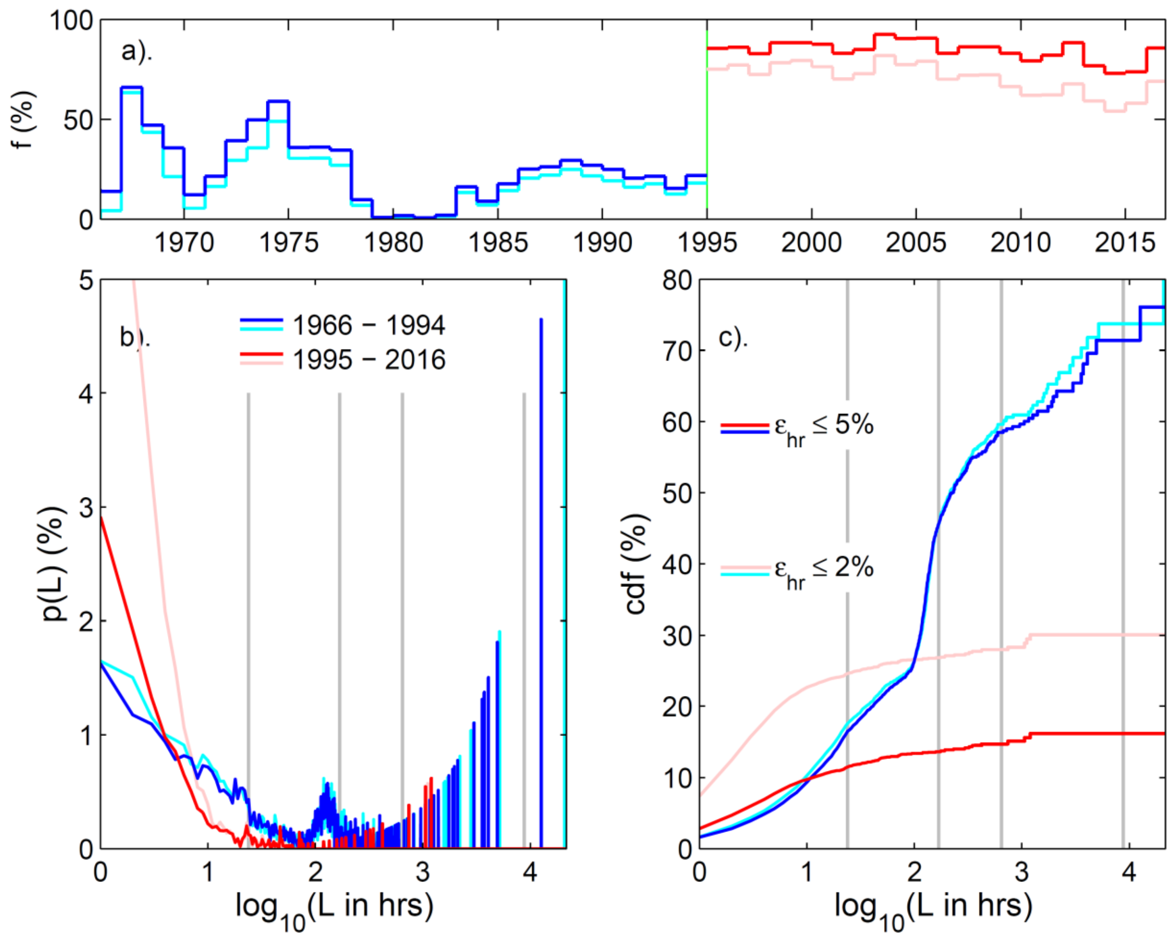
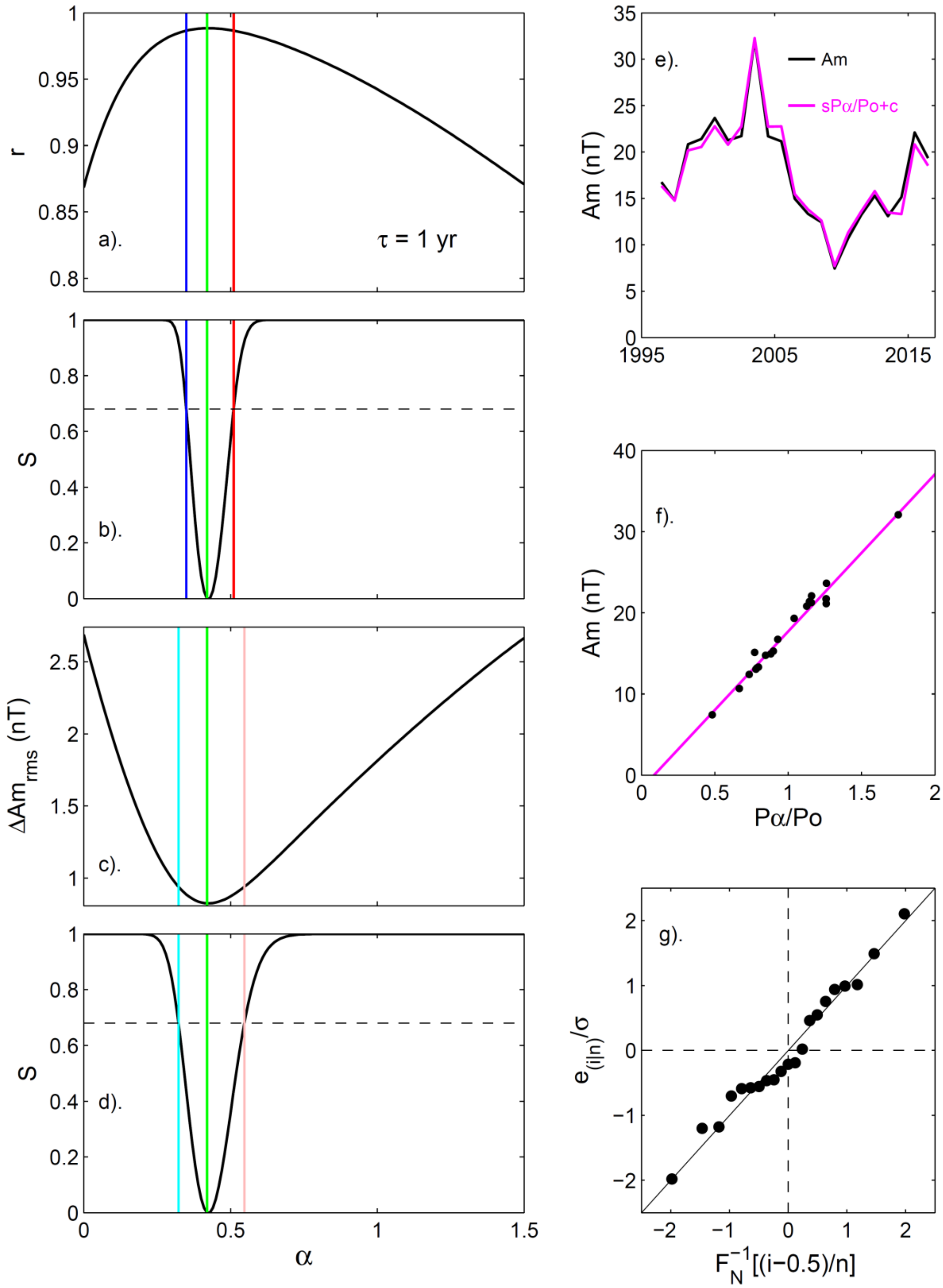
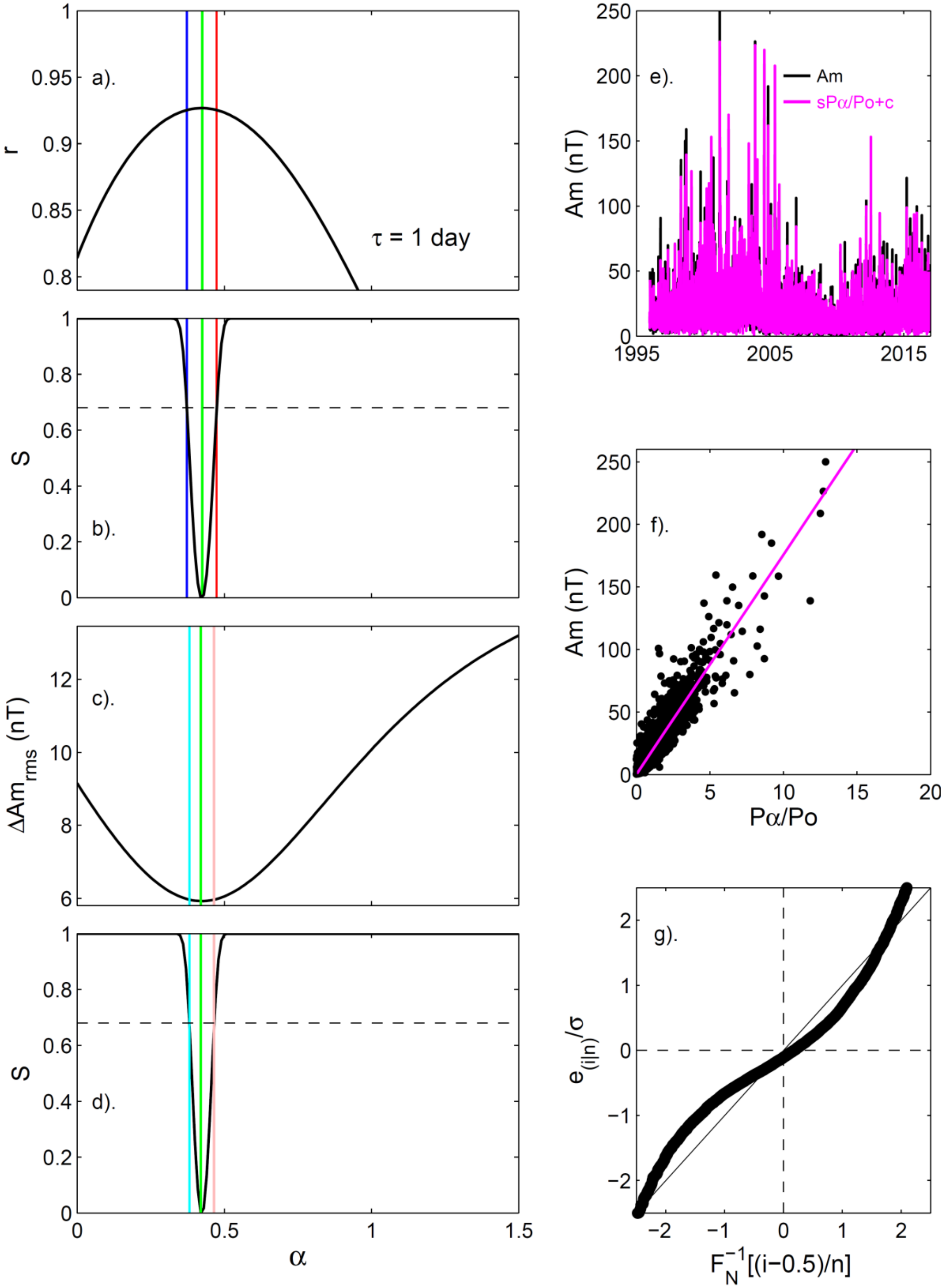


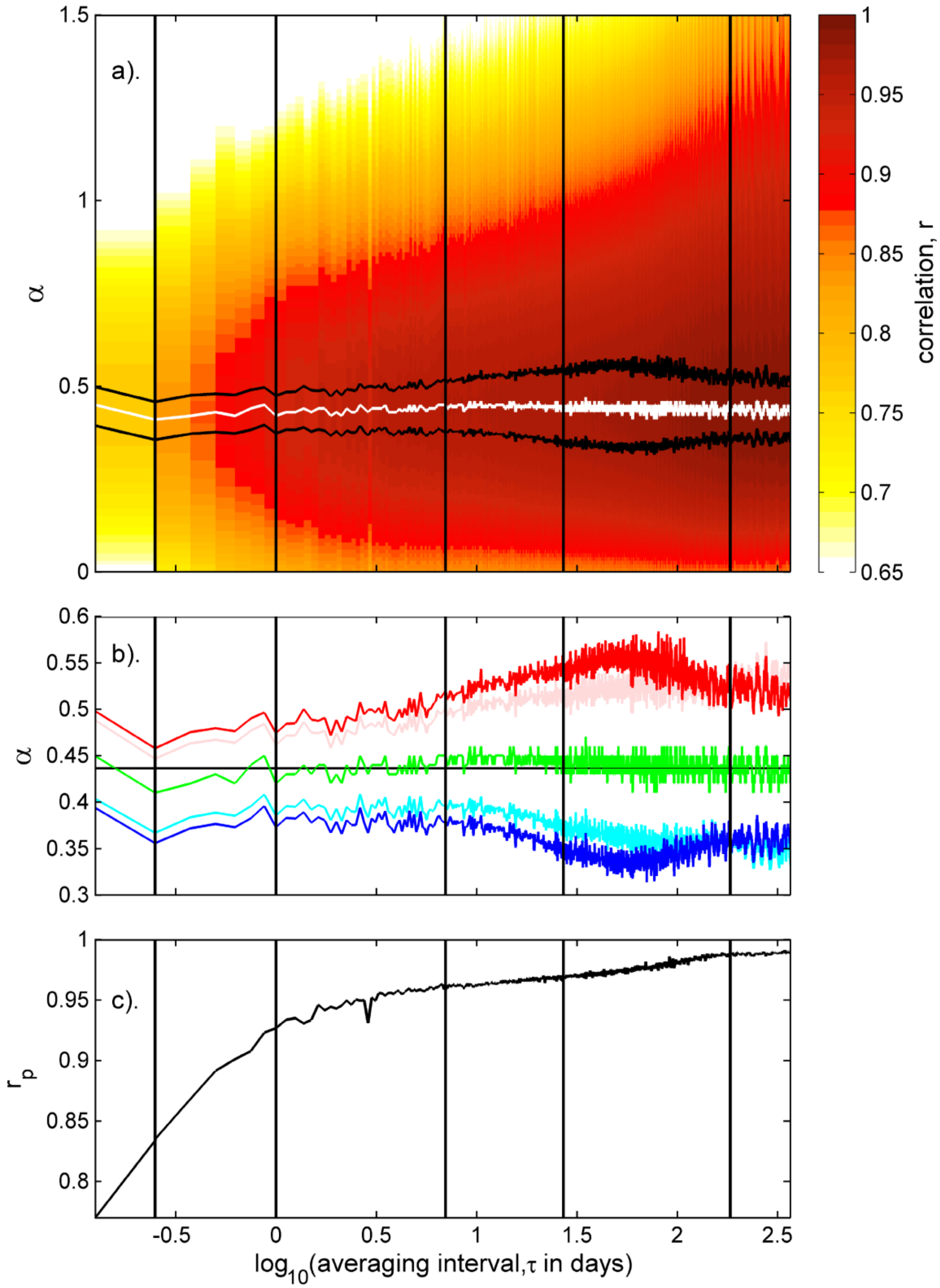
Figure 3. Same as Figure 2 for more stringent criteria as to what constitutes a valid hourly sample. Specifically, for the red and blue lines we require that solar wind data availability gives an average uncertainty in the hourly coupling function P_α values of $\varepsilon_{hr} \leq 5\%$ and for the pink and cyan lines we require $\varepsilon_{hr} \leq 2\%$. Blue and cyan lines are for before 1995: $\varepsilon_{hr} \leq 5\%$ gives the blue lines and 23.9% availability of hourly sample; lowering the uncertainty limit to $\varepsilon_{hr} \leq 2\%$ lowers this availability to 18.9% (cyan lines). Red and pink lines are for 1995 and after: $\varepsilon_{hr} \leq 5\%$ gives the red lines with 83.8% availability of hourly samples; lowering the uncertainty limit to $\varepsilon_{hr} \leq 2\%$ lowers this availability to 70.0% (pink lines).



1186 **Figure 4.** Analysis of correlation between power input to the magnetosphere, P_{α}, and the *am* geomagnetic index for averaging timescale τ = 1 yr over the interval 1996-2016 (inclusive). Panels (a)-(d) show, as a function of the coupling exponent α, (a) the correlation coefficient r between $<Am>_{1yr}$ and $< P_\alpha >_{1yr}$; (b) the significance, S, of the difference between the peak value of r, r_p (marked by the vertical green line at $\alpha = \alpha_p = 0.42$) and the r at any other α, $r(\alpha)$, computed using the Meng-Z text (see text for details): the vertical red and blue lines are where $S = 0.68$ and so mark where r is significantly lower than r_p at the 1-σ level; (c) the r.m.s. difference, ΔAm_{rms} between $<Am>_{1yr}$ and the best fit (for $\alpha = \alpha_p$) of $<P_\alpha/P_o>_{1yr}$, $Am_{fit} = s.<P_\alpha/P_o>_{1yr} + c$, where P_o is the average of P_α for all the data and the linear best-fit regression coefficients $s = 19.37 \pm 0.42$ and $c = -1.64 \pm 0.31$. (d) shows the significance S of the difference between ΔAm_{rms} for $\alpha = \alpha_p$ and that at other α, computed using a 2-sample variance F test (see text for details): pink and cyan vertical lines are at $S = 0.68$. (e) shows the time series of $<Am>_{1yr}$ (in black) and the best-fit Am_{fit} (in mauve);. (f) Scatter plot of $<Am>_{1yr}$ against $<P_\alpha / P_o>_{1yr}$ for $\alpha = \alpha_p$ with the best-fit linear regression line shown in mauve. (g) is a quantile-quantile (q-q) plot of the best-fit fit residuals of Am_{fit} and $<Am>$ used to test for the normality of their distribution: $e_{(i|n)}/\sigma$ are the ordered standardized residuals and $F_N^{-1}[(i-0.5)/n]$ are the quantiles of the standard normal distribution.



1203 **Figure 5.** Same as Figure 4 for an averaging timescale τ = 1 day. In this case, α_p = 0.42 and the best-fit linear regression coefficients are $s = 17.58 \pm 0.89$ and $c = 0.02 \pm 0.40$.



1205 **Figure 6.** (a). Correlation coefficients between $<am>_\tau$ and $<P_\alpha>_\tau$ for the interval 1996-2016

1206 (inclusive), r, color-coded as a function the logarithm of averaging timescale, $\log_{10}(\tau)$, and the coupling exponent, α. The middle white line gives the peak of each vertical slice, i.e. the optimum α, α_p (for which r has a maximum value r_p) (as shown by the vertical green lines in Figures 4a and 5a for $\tau = 1$ yr and $\tau = 1$ day, respectively). The upper and lower black lines give the $\pm 1\sigma$ uncertainty of α_p from the Meng-Z test (as shown by the vertical red and blue lines in Figure 4a for τ of 1 year and in Figure 5a for τ of 1 day, respectively). The left-hand edge of the plot is at $\tau = 3$ hrs, the right hand edge at $\tau = 1$ yr, and the vertical lines show τ of 6 hours, 1 day, 7 days, 27 days and 0.5 year. (b) Variation of best α estimates (α_p) and uncertainties. Line colors are as are used in Figures 4 and 5. The green line give the results for α_p from both the peak correlation ($r = r_p$) and the minimum r.m.s. fit deviation (ΔAm_{rms}), which are always exactly the same. The blue and red lines are the ± 1-σ uncertainties in α_p computed from the $r(\alpha)$ variation at each τ using the Meng-Z test. The cyan and pink are the 1-sigma uncertainties in α_p computed from the fit residuals $\Delta Am_{rms}(\alpha)$ variation using a two-sample variance F test. (c) The peak correlation r_p as a function of $\log_{10}(\tau)$.

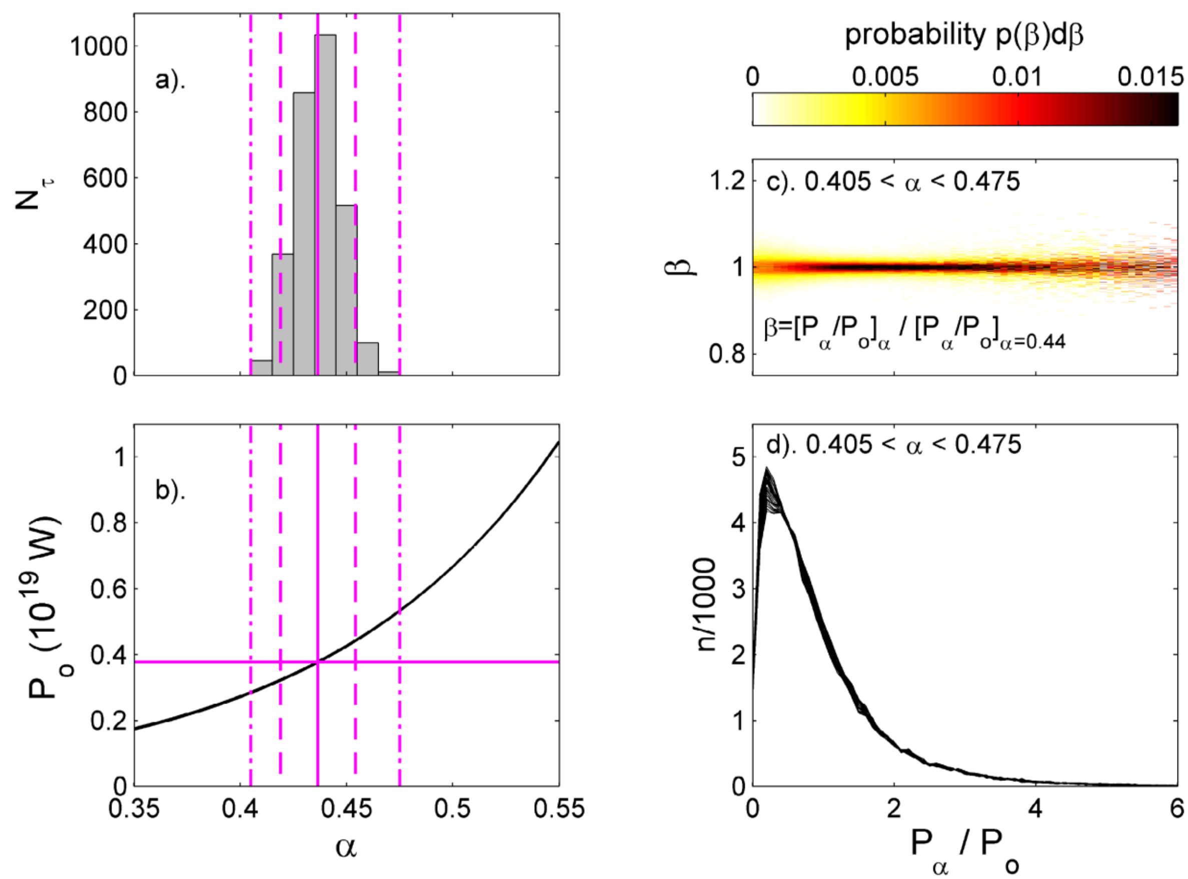


Figure 7. Analysis of the effects of the uncertainty in the best-fit α. (a). The distribution of best fit α values from the analysis of 2922 τ-values surveyed in figure 6. The vertical mauve dashed lines are the 2-σ points of the distribution and the dot-dash lines mark the extreme values. (b) The variation of the mean value of power input into the magnetosphere over the interval 1996-2017, P_o, as a function of α: the average α is 0.44 which gives $P_o = 0.38\times10^{19}$ W and the uncertainty range of α of 0.42- 0.45 at the 2-σ level (between the dashed lines) gives P_o between 0.32×10^{19} W and 0.44×10^{19} W. The extreme limits of α of 0.41 and 0.47 yield P_o of 0.29×10^{19} W and 0.53×10^{19} W, respectively. (c) Analysis of the effect of α on 3-hourly P_α/P_o values: the probability density is color-coded of values of $\beta = [P_\alpha/P_o]_\alpha/[P_\alpha/P_o]_{\alpha=0.44}$ for the full range of α of 0.41-0.47 (computed in steps of 0.01). (d) The distributions of P_α/P_o values for this full range of α.

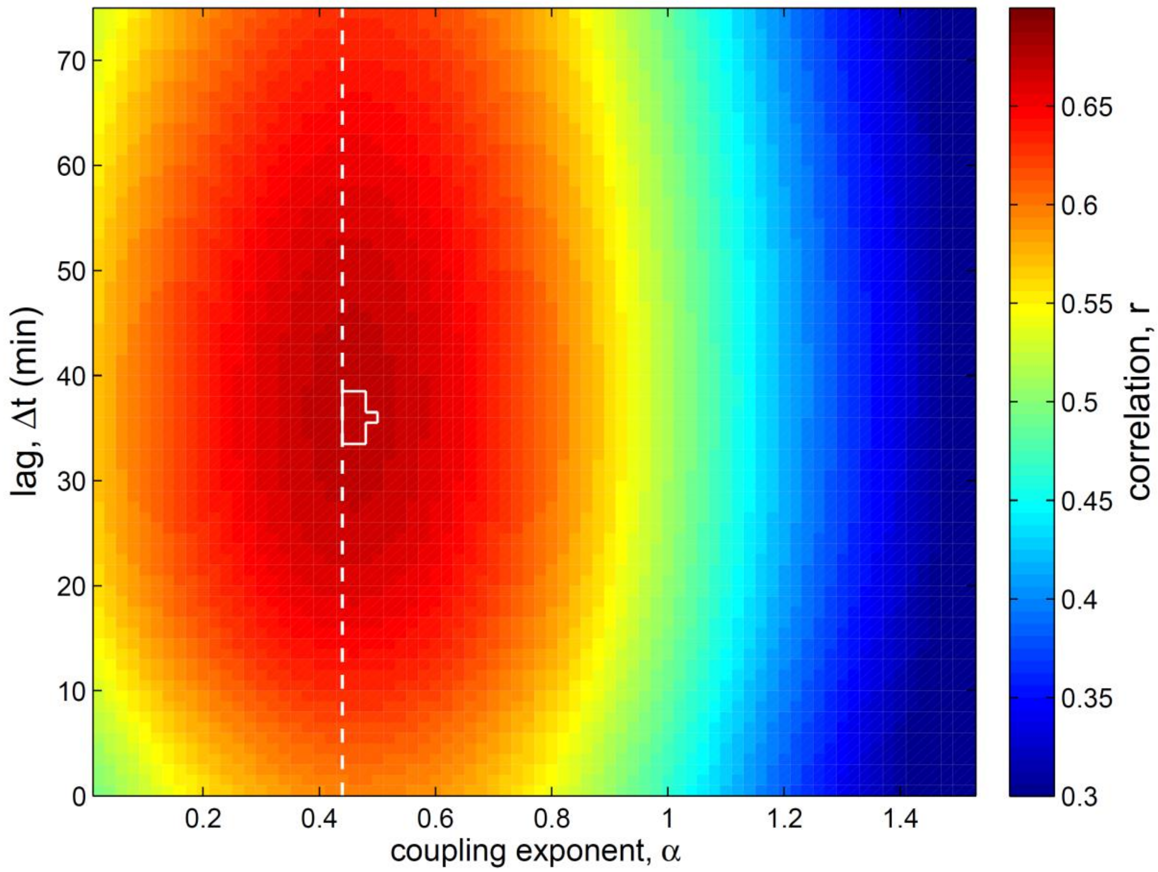


Figure 8. The correlation r of 1-minute P_α values with 1-minute values of the SuperMAG auroral electrojet index, SME, color-coded as a function of coupling exponent, α (horizontal axis) and time lag, Δt (vertical axis). Positive Δt is when P_α is lagged, i.e. the SME variation follows P_α. Inside the white contour the correlation is not significantly different from the peak value at the 1-σ level. The vertical white dashed line is at $\alpha = 0.44$.

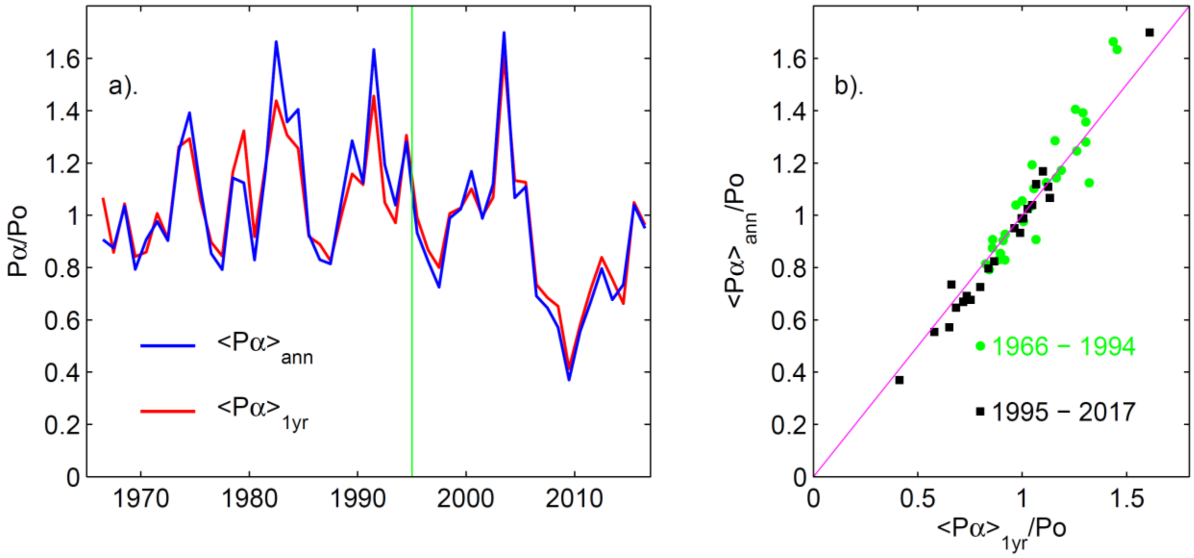


Figure 9. Comparison of "combine-then-average" and "average-then-combine" annual values of P_α (respectively, $<P_\alpha>_{1yr}$ and $<P_\alpha>_{ann}$). Note that the optimum coupling exponent is used which is $\alpha_p = 0.44$ for $<P_\alpha>_{1yr}$ and $\alpha_p = 0.48$ for $<P_\alpha>_{ann}$. (a) The time series of values (normalized to P_o, their mean value over the whole 1966-2016 data set: $P_o = 5.679 \times 10^{18}$ W for $<P_o>_{1yr}$ and $P_o' = 5.719 \times 10^{18}$ W for $<P_o>_{ann}$). The vertical green line is the start of almost continuous interplanetary data from ACE, WIND and DSCOVR. (b) Scatter plot of $< P_\alpha>_{1yr}/ P_o$ against $<P_\alpha>_{ann}/P_o'$. Points for 1995 and after are shown as black squares and points for before then are as green circles. The correlation coefficient for the 1995-2016 data is 0.994 and for the 1966-1994 data is 0.929. The mauve line is the line of perfect agreement.

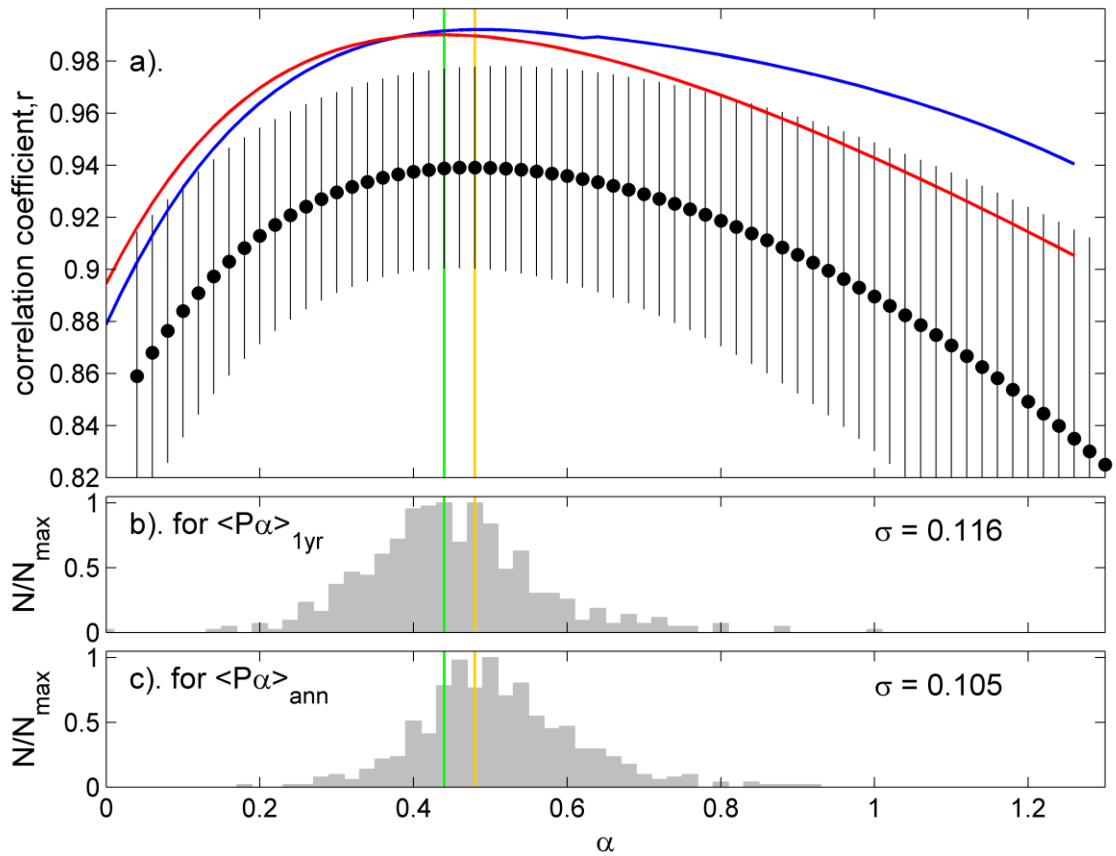


Figure 10. Analysis of the effect of adding the series of data gaps that exist in the P_α data series during 1966-1996 on the P_α-*am* correlogram (as a function of the coupling exponent, α,used) derived for annual means (i.e., τ = 1yr) from 1996-2016. In this case the effects of these gaps has been neglected (i.e., Method B, the most commonly employed). (a) The red and blue lines are the correlograms for all the 1996-2016 data for, respectively, $<P_\alpha>_{1yr}$ (the combine-then-average estimate) and $<P_\alpha>_{ann}$ (the average-then-combine estimate). The black points with error bars are the mean and standard deviations of the distribution of correlation coefficients, *r*, for $<P_\alpha>_{1yr}$ at a given α, obtained from N_s = 500 different masks applied to the 1996-2016 data to introduce extra, synthetic data gaps. These masks produce the time series of gaps as was observed (30–*D*) years earlier where *D* is varied between 0 and 10 years in steps of 175 hours. (So, for example, for *D* = 0, the mask produces data gaps in the same positions of the data series as for the 1966-

1986 data and for D = 10 years they are in the same positions as for 1976-1996). For the period from which the data gap masks are drawn (1966-1996) data availability is 23.9%. (b) The distribution of the optimum α values (which give the peak r between $\langle P_\alpha \rangle_{1yr}$ and $\langle am \rangle_{1yr}$ for the N_S = 500 different data gap masks applied to both series. (c) The corresponding distribution of α values giving the peak r between $\langle P_\alpha \rangle_{ann}$ and $\langle am \rangle_{1yr}$ for the 500 different data gap masks applied to both series. The green line in all three panels is the optimum α of 0.44 derived for $\langle P_\alpha \rangle_{1yr}$ and the orange line is the optimum α of 0.48 derived for $\langle P_\alpha \rangle_{1ann}$.

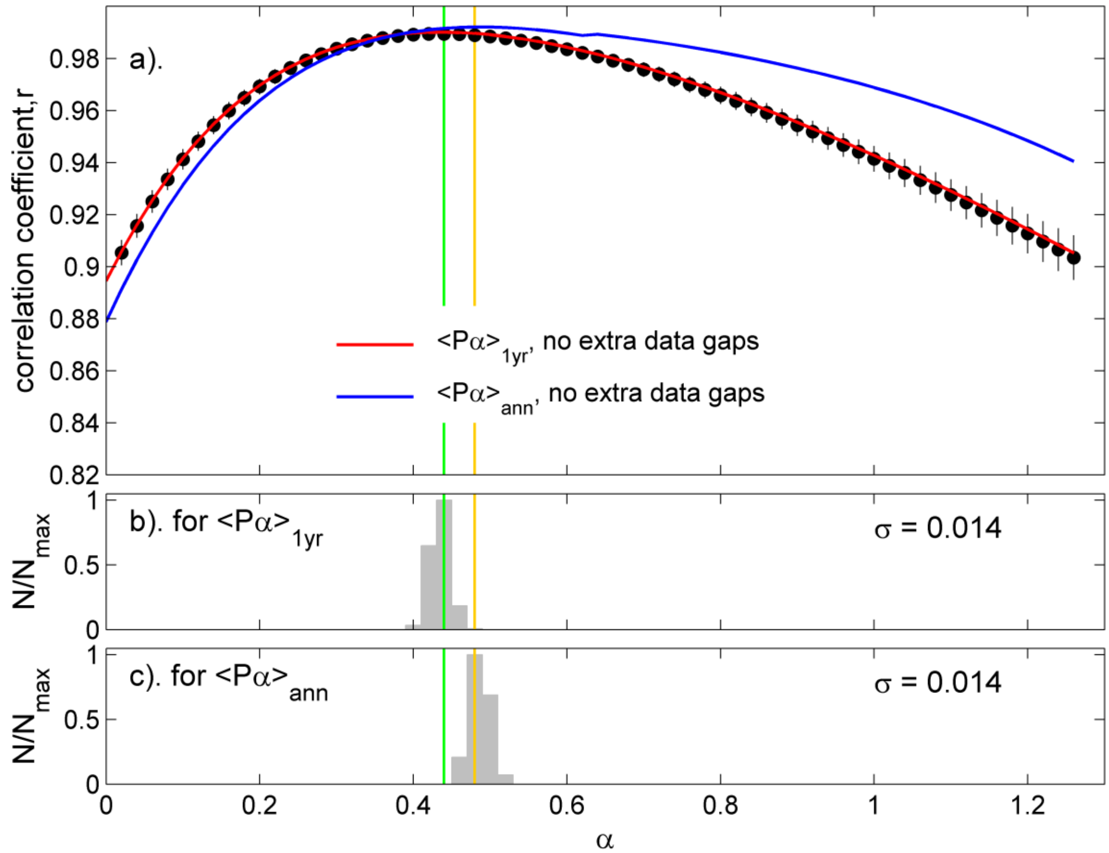


Figure 11. The same as Figure 10, but allowance for data gaps is made using the piecewise removal procedure of *Finch and Lockwood* [2007] (Method B) applied to the *am* data at times of the synthesized data gaps in P_α. In addition, in this case, the 500 synthetic data gap masks that have been applied to the P_α data are drawn from sequences of actual data gaps during 1995-2016, when data availability is 83.8%. Specifically, the 500 masks constructed for the data-gap series observed (1-*D*) years earlier (when the data are almost continuous) for *D* varied between 0 and 1 year in steps of 17.5 hours. (So for example, for *D* = 0, the mask produces extra, synthetic data gaps in the same positions of the data series as for the 1995-2015 data and for *D* = 1 year the mask reproduces the actual data gaps 1996-2016).