

*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., Bentley, S. N., Owens, M. J., Barnard, L. A., Scott, C. J., 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: <https://doi.org/10.1029/2018SW001856> Available at <http://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

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

3 Mike Lockwood¹, Sarah N. Bentley¹, Mathew J. Owens¹, Luke A. Barnard¹,
4 Chris J. Scott¹, Clare E. Watt¹ and Oliver Allanson¹

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

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

8 Key Points:

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

15 **Abstract**

16 Different terrestrial space weather indicators (such as geomagnetic indices, transpolar voltage,
 17 and ring current particle content) depend on different “coupling functions” (combinations of
 18 near-Earth solar wind parameters) and previous studies also reported a dependence on the
 19 averaging timescale, τ . We study the relationships of the *am* and SME geomagnetic indices to
 20 the power input into the magnetosphere P_α , estimated using the optimum coupling exponent α ,
 21 for a range of τ between 1 min and 1 year. The effect of missing data is investigated by
 22 introducing synthetic gaps into near-continuous data and the best method for dealing with them
 23 when deriving the coupling function, is formally defined. Using P_α , we show that gaps in data
 24 recorded before 1995 have introduced considerable errors into coupling functions. From the
 25 near-continuous solar wind data for 1996-2016, we find $\alpha = 0.44 \pm 0.02$ and no significant
 26 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
 27 Interplanetary Magnetic Field (IMF), N_{sw} the solar wind number density, m_{sw} its mean ion mass,
 28 V_{sw} its velocity and θ the IMF clock angle in the Geocentric Solar Magnetospheric reference
 29 frame. Values of P_α that are accurate to within $\pm 5\%$ for 1996-2016 have an availability of 83.8%
 30 and the correlation between P_α and *am* for these data is shown to be 0.990 (between 0.972 and
 31 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
 32 hours, respectively, and that between P_α and *SME* at τ of 1 min. is 0.7046 ± 0.0004 .

33 **1. Introduction**

34 **1.1 Coupling functions**

35 On short timescales, the coupled magnetosphere-ionosphere-thermosphere system responds to
 36 the magnitude of the southward component of the Interplanetary Magnetic Field (IMF, in a
 37 suitable frame oriented with respect to the geomagnetic field axis, such as Geocentric Solar
 38 Magnetospheric, GSM). There are two time constants of response, the “directly-driven” system
 39 responds on a timescale of order a few minutes [Nishida, 1968; Lockwood *et al.*, 1986; Etemadi
 40 *et al.*, 1988, Todd *et al.*, 1988]. The directly-driven flows store magnetic energy in the

41 magnetospheric tail and subsequently that energy is released and deposited in the inner
42 magnetosphere and nightside auroral ionosphere and thermosphere via the substorm current
43 wedge. This is the “storage-unloading” system, which generates a response after a delay of
44 typically between about 30 and 60 min. [*Schatten and Wilcox*, 1967; *Arnoldy*, 1971; *Baker et al.*,
45 1981, 1983; *Lockwood et al.*, 1990]. *Finch et al.* [2007] used data from the global network of
46 geomagnetic observatories to show how stations responded differently to the directly-driven and
47 storage/unloading systems depending on their position (in geomagnetic latitude and magnetic
48 local time (MLT) coordinates). Similar conclusions was reached by *Dods et al.* [2015; 2017],
49 using network analysis.

50 A great many combinations of near-Earth interplanetary parameters (so-called “coupling
51 functions”) have been proposed over many years to describe the transfer of energy, and/or mass,
52 and/or momentum, and/or electric field from the solar wind into the Earth’s magnetosphere-
53 ionosphere-thermosphere system [e.g., *Burton et al.*, 1975; *Crooker et al.*, 1977; *Feynman and*
54 *Crooker*, 1978; *Perreault and Akasofu*, 1978; *Kan and Lee*, 1979; *Reiff et al.*, 1981; *Murayama*,
55 1982; *Wygant et al.*, 1983; *Bargatze et al.*, 1985; 1986; *Vasyliunas et al.*, 1982; *Scurry and*
56 *Russell*, 1991; *Wing and Sibeck*, 1997; *Wu and Lundstedt* , 1997; *Papitashvili et al.* 2000;
57 *Temerin and Li*, 2002; 2006; *Newell et al.*, 2007; *Borovsky*, 2008; *Balikhin et al.*, 2010; *Spencer*
58 *et al.*, 2011; *Tenfjord and Ostgaard*, 2013, *McPherron et al.*, 2015]. These are derived and
59 tested by comparison with terrestrial space weather disturbance indices (sometimes in
60 combination) which respond to the energy, mass, electric field and/or momentum input to the
61 magnetosphere from the solar wind (in a combination thereof that depends on which terrestrial
62 index is used). *Gonzalez* [1990] examined the empirical coupling functions in use at the time
63 and concluded that almost all could be derived either from electric field or power transfer
64 associated with the reconnection process. Although they are related in a general sense [e.g. *Saba*
65 *et al.*, 1997], different terrestrial indices respond differently to different combinations of energy,
66 mass, momentum and/or electric field transfer into the magnetosphere (and on different
67 timescales): hence it is not surprising that there are a wide variety of derived coupling functions.
68 For example, *Svalgaard and Cliver* [2005] noted that different geomagnetic indices responded
69 differently to different coupling functions. This is also unsurprising as each is influenced by
70 different combinations of the currents in the terrestrial system: the ionospheric DP1 (auroral

71 electrojet and substorm current wedge); the ionospheric DP2 (growth phase convection) currents;
 72 the magnetospheric ring current; the magnetopause (Chapman-Ferraro) currents; the cross-tail
 73 current sheet; and (unless ionospheric conductivities are spatially uniform) the field-aligned
 74 (Birkeland) currents. *Svalgaard and Cliver [2005]* also noted that this is a very useful feature
 75 because it allows reconstruction of multiple solar wind parameters from combinations of indices
 76 for times before interplanetary spacecraft were available, as exploited by *Lockwood et al. [1999]*,
 77 *Lockwood [2003]*, *Rouillard et al. [2009]* and *Lockwood et al. [2014a]*. A comparison of the
 78 optimum exponents of interplanetary parameters for various commonly-used geomagnetic
 79 indices has been presented for annual timescales by *Lockwood [2013]*.

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

89 Many studies of coupling functions have concentrated on the exponents needed for each of the
 90 solar wind parameters that have been shown statistically to play a role in controlling the variation
 91 of the terrestrial response, and have used multivariate analysis to adjust the exact functional form
 92 of the combination of parameters used in the coupling function. This raises a problem. Including
 93 all of the factors with their own weighting factors and/or exponents could result in extremely
 94 good fits that are, nevertheless, statistically meaningless as each additional fit parameter reduces
 95 the statistical significance of the correlation. Such fits can have limited, and in extreme cases, no
 96 predictive capability. This pitfall is called “over-fitting” and it is a serious, but often under-
 97 appreciated, problem in multiple regression analysis of geophysical time series that have internal
 98 “geophysical” noise. Overfitting refers to the situation when a fit has too many degrees of
 99 freedom and starts to approximate the noise in the training subset, which is not robust throughout
 100 the whole dataset. This is a recognized pitfall in areas where quasi-chaotic behaviors give large

101 internal noise such as climate science and population growth (see, for example, *Knutti et al.*
102 [2006] and *Knape and de Valpine* [2011]) but is not recognized as widely in space physics,
103 possibly because systems may have been viewed as being somewhat more deterministic and with
104 lower internal variability. To help guard against over-fitting, we favor the physics-based
105 approach to coupling function derivation by *Vasyliunas et al.* [1982] which is described in the
106 next section. This approach yields a single fit parameter, the coupling exponent α , which arises
107 from an unknown dependence on the solar wind Alfvén Mach number but which influences the
108 dependence of the coupling function on solar wind speed, V_{sw} , number density N_{sw} , mean ion
109 mass, m_{sw} and the interplanetary field B – all in a consistent way. Despite having just the one
110 free fit parameter, extremely high correlations with geomagnetic indices (often exceeding 0.99
111 for annual timescales) can be derived for some geomagnetic indices.

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

123 The study presented here as also been repeated using the *ap* index, generated from *K* values from
124 a network of 11–13 stations at $|\Lambda_{CG}|$ between 44° and 60° . To generate this index, the station *K*
125 indices are first converted into standardized K_s values, using conversion tables for each
126 observatory that were defined by *Bartels* [1949, 1957]. The *Kp* index is the arithmetic mean of
127 the 3-hourly K_s values for the observatories employed. The *Kp* values are converted into *ap*
128 values using the standard K2aK scale that is constructed such that *ap* may be regarded as the
129 range of the most-disturbed of the two horizontal field components, expressed in the unit of 2nT,

130 at a station at dipole latitude of 50° . More details of the compilation of the *ap* and *am* indices are
 131 provided by *Mayaud* [1980], *Menvielle and Berthelier* [1991] and *Dieminger et al.* [1996].

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

142 The *am* index and its daily means, *Am*, responds strongly to substorm currents [*Saba et al.*,
 143 1997], such that the linear correlation coefficient between all available coincident 56 annual
 144 means of the auroral electrojet AE(12) index (from 1959-2017, inclusive) and *Am* is 0.92 ± 0.04
 145 (significant at $>99.99\%$ level), and the correlation between the 20301 coincident daily means of
 146 AE(12) and *Am* is 0.855 ± 0.004 (significant to at least the same level). One reason for the lower
 147 correlation for daily means is evident from the scatter plot which is linear at low disturbance
 148 levels but shows a marked non-linearity in large disturbances, with large *Am* values consistently
 149 exceeding the corresponding linearly-regressed values from AE(12) (see, for example, Figure 4
 150 of *Adebesin* [2016] and the scatter plots presented in the supplementary material for the present
 151 paper). This is consistent with the effect on AE(12) of an extremely expanded auroral oval
 152 migrating increasingly equatorward of the fixed latitude of the ring of 12 stations that it is
 153 compiled from.

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

159 *Lockwood* [2004]). Hence am is an index that we expect to depend on the energy input to the
160 magnetosphere and so correlate well with P_{α} .

161 The substorm current wedge (that am responds to) forms when the near-Earth edge of the cross-
162 tail current is disrupted at, or soon after, substorm expansion phase onset and, due to the shape of
163 the magnetosphere, the solar wind dynamic pressure constrains the cross-sectional area of the
164 geomagnetic tail lobe at such locations. This means that solar wind dynamic pressure plays a
165 role in the growth-phase rise in magnetic field intensity at and around the distances down the tail
166 (the X -coordinate in the GSM frame) of the inner edge of the cross tail current. Hence the
167 dynamic pressure (as well as the open magnetospheric flux in tail) controls the magnetic shear
168 and hence the cross-tail current at these locations [*Lockwood, 2013*]. (Note that this is not the
169 case in the far tail (greater $-X$) where the tail lobe area expands as the open flux increases and
170 the cross-tail current is constant, being set by the static pressure of the solar wind). This role of
171 solar wind dynamic pressure in setting the cross-tail current (the current that becomes deflected
172 into the substorm current wedge at substorm onset) means that the am index should show a
173 strong variation on V_{sw}^n , where the exponent n is close to 2 [*Lockwood, 2013*]. The K_p (and
174 hence ap) and am indices are also known to be good indicators of the strength of convection in
175 the magnetosphere-ionosphere system [*Thomsen, 2004*]. That am is good indicator of both
176 substorm currents and convection is also expected from the Expanding-Contracting Polar Cap
177 model in which convection is understood as the net effect of the opening of field lines by
178 magnetopause reconnection (associated with the directly-driven current system, and the storage
179 (growth) phase of the substorm cycle) and their re-closing by reconnection in the cross-tail
180 current sheet in expansion and recovery phases (associated with the unloading phase of the
181 substorm cycle) [*Lockwood et al., 1990; Lockwood and Cowley, 1992; Cowley and Lockwood,*
182 *1992; Milan et al., 2003; 2008*].

183 **1.2. Power Input into the magnetosphere**

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

$$188 \quad P_{\alpha} = (\pi L_o^2) \times (m_{sw} N_{sw} V_{sw}^3/2) \times (t_r) \quad (1)$$

189 where L_o is the mean cross-sectional radius of the magnetosphere, such that the geomagnetic
 190 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
 191 of the particles, which is the dominant energy flux in the solar wind (where m_{sw} is the mean ion
 192 mass, N_{sw} the number density and V_{sw} the speed). The term t_r is a dimensionless transfer
 193 function, the fraction of the incident energy flux that is transferred to inside the magnetosphere.

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

$$196 \quad L_o = k_1 (M_E^2/P_{sw} \mu_o)^{1/6} \quad (2)$$

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

$$199 \quad P_{sw} = m_{sw} N_{sw} V_{sw}^2 \quad (3)$$

200 *Vasyliunas et al.* [1982] adopted the dimensionless transfer function:

$$201 \quad t_r = k_2 M_A^{-2\alpha} \sin^4(\theta/2) \quad (4)$$

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

$$207 \quad M_A = V_{sw} (\mu_o m_{sw} N_{sw})^{1/2} / B \quad (5)$$

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

$$209 \quad P_{\alpha} = (k_1 k_2 \cdot \pi/2 \mu_o^{(1/3-\alpha)}) M_E^{2/3} m_{sw}^{(2/3-\alpha)} \times B^{2\alpha} N_{sw}^{(2/3-\alpha)} V_{sw}^{(7/3-\alpha)} \sin^4(\theta/2) \quad (6)$$

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

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

226 The helium abundance ratio, N_{He}/N_p , is relatively constant near 0.04 in high-speed streams, but
 227 varies in phase with the sunspot number in slow wind, between about 0.01 and 0.04 [Wang,
 228 2016]. Hourly means of N_{He}/N_p are available in the Omni2 dataset from 1972 and are here
 229 interpolated to the center times of the 1-minute samples of B , N_{sw} and V_{sw} using Piecewise Cubic
 230 Hermite Interpolating Polynomial (PCHIP) interpolation [Fritsch and Carlson, 1980]. As a
 231 check, linear interpolation was also used and the deviation of the results from the two methods
 232 quantified to check that it remained small (<10%) and hence that the interpolations were
 233 reasonable. The mean ion mass is then given by

$$234 \quad m_{sw} = m_p \{ 1 + (4N_{\text{He}}/N_p) \} / \{ 1 + (N_{\text{He}}/N_p) \} \quad (7)$$

235 where m_p is the mass of the proton.

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

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

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

257 **1.3. The effect of data gaps on studies of coupling functions**

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

263 Data gaps can influence correlation studies [George *et al.*, 2015]. For many applications such as
 264 spectral analysis, it is desirable or necessary to fill the data gaps and a number of methods for
 265 doing this are available, but one has to remain aware of the implications of the method used for
 266 application in question [Sturges, 1983; Wynn and Wickwar, 2007; Henn *et al.*, 2013, Munteanu
 267 *et al.*, 2016]. Such gap filling techniques have been applied to solar wind data by, for example,
 268 Kondrashov *et al.* [2010; 2014] but many require a proxy dataset for either interpolation or
 269 testing purposes. This usually involves using a terrestrial space weather disturbance index as the

270 proxy, which means the coupling function is assumed and used to help fill the data gap and so
 271 such techniques must not be used in the context of deriving a coupling function. Correlation
 272 studies, such as between a solar wind coupling function and a terrestrial space weather response,
 273 are not influenced by data gaps at basic time resolution, other than through the loss of correlation
 274 significance through the reduced number data point pairs available. The question then arises as to
 275 how best to deal with data gaps when the basic time resolution data are averaged into longer
 276 intervals τ . Some of the more recent studies of solar wind-magnetosphere coupling have been
 277 able to reduce the problem of data gaps by employing only data from after 1995 when data
 278 became much more continuous [e.g., *Luo et al.*, 2013; *Temerin and Li*, 2006], despite the
 279 drawback that this reduces the number of years covered by the study. Some studies have placed
 280 tight restrictions on data gap occurrence [e.g., *Teodorescu et al.*, 2015] but these tend to be
 281 severe and also greatly reduce the available data such that only some spectral studies are possible
 282 and coupling functions are studied through of a series of events [e.g. *Bargatze et al.*, 1985;
 283 1986]. This limits the potential for noise reduction by averaging (which makes correlation
 284 coefficients increase with averaging timescale τ) and reduction of the influence of others factors
 285 such as ionospheric conductivity [e.g., *Nagatsuma*, 2006; *Luo et al.*, 2013] . Many studies
 286 continue to use the pre-1995 data, despite its severe problem of many and long data gaps (as
 287 shown in Figures 2 and 3 and discussed in section 2). In a great many of these studies, the
 288 averages of all available data have been taken on the (usually tacit) assumption that the effect of
 289 data gaps will average out. This philosophy is often introduced by the use of the Omni or Omni2
 290 interplanetary datasets, which supply averages for a requested τ without a limit on the data
 291 availability (so in extreme cases a mean is given even if only one data sample is available in the
 292 interval). *Lockwood and Finch* [2007] made allowance for data gaps by introducing the solar
 293 wind data gaps into the geomagnetic data sequence that they correlated with and showed that the
 294 assumption that data gaps can be neglected is often invalid: hence they argued for piecewise
 295 removal of the geomagnetic data at times corresponding to the interplanetary data gaps (i.e.
 296 allowing for an appropriate lag) before the averaging of both into the intervals of duration τ . For
 297 P_{α} values this means that if any of N_{sw} , V_{sw} , m_{sw} , B or θ were missing, the corresponding
 298 geomagnetic data were piecewise masked out. This ensures that the averages contain only the
 299 corresponding data in the two data sets. Although they argued that this was the best approach,

300 [*Finch and Lockwood, 2007*] also cautioned that data gaps could still have an influence on the
 301 optimum correlations.

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

306 Other interpolation procedures have been employed to fill data gaps, such as in *Temerin and Li*
 307 [2002] in their analysis of a coupling function to predict the Dst index. However, the same
 308 authors soon after moved to using only data from after 1995 and assuming that the effects of
 309 gaps averaged out, with considerable improvement of the correlation between their coupling
 310 function and the Dst geomagnetic index [*Temerin and Li, 2006*]. Others have used interpolation
 311 with restrictions, for example *Wu and Lundstedt* [1997] deployed autocorrelation-based
 312 interpolation for data gaps of up to 5 hours, and intervals with longer data gaps excluded from
 313 the study.

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

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

341 2. Datasets

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

349 A great many coupling function studies have employed the Omni2 hourly means, which are
 350 generated for the Omni database even if only one sample is available within the hour (but note
 351 the number of samples used can always also be downloaded). In evaluating the full effect of data
 352 gaps, the question arises as to how much data within each hour is required to generate a valid
 353 hourly mean (i.e., one which is close to the value that would be obtained with continuous
 354 sampling). The answer to this depends strongly on the autocorrelation function of the parameter
 355 in question. Figure 1a shows the autocorrelation functions (a.c.f.s) of B , N_{sw} , V_{sw} , $\sin^4(\theta/2)$ and
 356 P_α (computed here for the optimum α found in section 3, $\alpha_p=0.5$) for lags up to 3 hours. The

357 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
 358 considerably whereas that of V_{sw} has hardly fallen at all. It follows that the number of samples
 359 required to make a valid mean to a given uncertainty is much greater for $\sin^4(\theta/2)$ than it is for
 360 V_{sw} .

361 To study the effect of data gaps on 1-hour averages, we here took 1-minute Omni2 data for 1996-
 362 2016 (11046240 samples, falling in 184104 hourly intervals). We note that for some of the solar
 363 wind measurements, the basic time resolution is greater than 1-minute: for example the ACE
 364 satellite spins with 64 s. period setting the time resolution of the solar wind data. In such cases,
 365 the long autocorrelation time constant of the solar wind parameters allows interpolation into 1-
 366 minute values. However, the same is not true for the IMF data. For each parameter we searched
 367 for hourly intervals in which all 60 minutes in the hour gave a 1-minute sample made up from at
 368 least one observation such that it was not classed as “missing data” in the Omni dataset. For B ,
 369 and $\sin^4(\theta/2)$ and such hours numbered 112855; for N_{sw} and V_{sw} the number was 121685 and for
 370 P_α it was 1083797. For these hours we removed N_g 1-minute sub-samples at random (this was
 371 done 10 times for each of the hours) where N_g was increased from 1 up to 59 and the percentage
 372 error in the mean of the reduced data for that parameter and hour was evaluated by comparing it
 373 to the mean for the full data (i.e., the effect of the synthesized data gaps was determined). The
 374 distributions of these fractional errors were Gaussian and in Figure 1b the standard deviations
 375 (i.e., the one sigma error in the hourly means), ϵ_{hr} , are shown for each parameter as a function of
 376 data availability within the hour, $f_{hr} = 1 - (N_g/60)$. As expected, the high a.c.f. of V_{sw} results in the
 377 average error introduced into hourly means, ϵ_{hr} , being very small and it is only 0.2% even if just
 378 one sample is available in the hour. For N_{sw} and B , this figure is larger but is still, respectively,
 379 only just above and just below 2%. Much larger errors arise from the IMF orientation factor
 380 $\sin^4(\theta/2)$, as expected for its greater variability at sub-hour timescales, shown by rapid decline in
 381 its a.c.f. with time in Figure 1a. Figure 1b shows that it is the variability in $\sin^4(\theta/2)$ that
 382 generates matching behavior for P_α , and hence sets a requirement for a sub-hour $\sin^4(\theta/2)$
 383 availability threshold for coupling studies. From this plot we can find the fraction of samples f_{hr}
 384 that is needed in the hour to give a specified 1- σ uncertainty in the hourly means for that
 385 parameter, ϵ_{hr} . We here study the 2% and 5% uncertainty levels. Figure 1b shows that $\epsilon_{hr} < 2\%$ is

386 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
 387 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
 388 just one sample is adequate.

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

398 In both Figures 2 and 3, and all subsequent places in the text, we employ hourly means that we
 399 have made from the downloaded 1-minute Omni data. In Figure 3 we require enough 1-minute
 400 samples in the hour to reduce the statistical uncertainty (as shown by Figure 1) to a given level
 401 whereas in Figure 2 we use the criterion of requiring just a single 1-minute sample in the hour to
 402 generate an hourly mean in order to investigate the effects of that criterion. To distinguish the
 403 two later in this paper, the latter cases are always identified using the phrase “using the Omni
 404 criterion” (because the Omni data give an hourly mean value even if only one sample in the hour
 405 is available). All other hourly means refer to the former definition. From the analysis presented
 406 in Figure 1, the Omni criterion is certainly adequate for V_{sw} and generates acceptable average
 407 uncertainties of $\varepsilon_{th} \approx \pm 2\%$ for B and N_{sw} . However, for the IMF orientation, and hence coupling
 408 function, the errors in using a single sample are $\varepsilon_{th} \approx \pm 15\%$. Comparison of Figures 2 and 3
 409 investigates the effect of criteria on the occurrence of data gaps because, whereas Figure 2 is for
 410 the Omni criterion, Figure 3 uses the criterion that $\varepsilon_{th} \leq 5\%$ (requiring $f_{hr} \geq 0.82$ for the IMF
 411 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
 412 the solar wind data).

413 The availability of hourly samples in calendar years, f , is shown for the Omni criterion in Figure
 414 2a. The large rise in f caused by the advent of data from ACE and WIND in 1995 is marked by

415 the vertical green line. Figure 2b shows the spectrum of data-gap durations by giving the
 416 probability $p(L)$ of being in a data gap of duration L , as a function of $\log_{10}(L)$, evaluated in bins 1
 417 hour wide. Note that at large L before 1995 (in blue) the spectrum become discrete as there were
 418 often just single occurrences of a data gap at that L , but $p(L)$ increases linearly with L for such
 419 cases. Figure 2b shows that data gaps at almost all L (but particularly at large L) were
 420 considerably rarer after 1995 than before. Figure 2c shows the information as the cumulative
 421 probability distribution, which shows clearly the durations of gaps that contribute most to the
 422 loss of data. Using the **Omni** criterion, the total probability for being in a data gap for 1966-1994
 423 is 68% (so the average availability is $\langle f \rangle = 32\%$) and for 1995-2016 it is 8% ($\langle f \rangle = 92\%$).
 424 Remember from Figure 1b, this gives average errors in the hourly energy input estimates, P_α , of
 425 $\varepsilon_{\text{th}} \approx \pm 15\%$.

426 Figure 3 is the same as Figure 2, but for two tighter criteria on what constitutes a valid mean,
 427 namely $\varepsilon_{\text{hr}} < 5\%$ and $\varepsilon_{\text{hr}} < 2\%$. The red and blue lines require IMF data to have an availability of
 428 82% within in each hour (giving $\varepsilon_{\text{th}} \leq 5\%$) and this raises the total probability for being in a data
 429 gap for 1966-1994 to 76.1% (so the average availability is $\langle f \rangle = 23.9\%$) and for 1995-2016 it is
 430 16.2% ($\langle f \rangle = 83.8\%$). Comparison of the red and blue lines in Figures 2a and 3a shows that the
 431 additional requirement reduces the availability in most years before 1995 and makes some years
 432 around 1980 unusable. After 1995, it amplifies somewhat the trend for loss of data towards the
 433 end of the period when the main source of data was the ACE satellite. Comparing parts (b) and
 434 (c) of the two figures show that data gaps of all durations are increased by the extra criterion
 435 before 1995 but that after then the main effect is the loss of individual hours with data gaps of
 436 duration $L < 1$ day. The pink and cyan lines in Figure 3 are for $\varepsilon_{\text{th}} \leq 2\%$. The cyan line in Figure
 437 3a shows that this relatively small (3%) gain in the accuracy has a relatively small effect on the
 438 availability of valid data before 1995 ($\langle f \rangle = 18.5\%$) but after then it has a somewhat greater
 439 effect, lowering the average availability to $\langle f \rangle = 70.0\%$. Figure 3b and 3c show that the
 440 dominant effect of the lower error threshold is to introduce many short-duration data gaps.
 441 Hence a compromise needs to be struck between the error in hourly P_α means, ε_{th} , and the
 442 average availability of those hourly means $\langle f \rangle$. We here adopt an uncertainty limit of $\varepsilon_{\text{th}} \leq 5\%$

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

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

449 3. Optimum coupling function as a function of timescale

450 The hourly interplanetary data formed from hours with more than 82% of 1-minute IMF
451 orientation samples (giving $\varepsilon_{th} \leq 5\%$) were averaged into independent intervals of duration τ . The
452 process was repeated 2922 times as τ was varied between 3hrs and 1average year (365.25 days =
453 8766hrs) in steps of 3hrs (in other words, the values of τ used were 3hrs, 6hrs, 9hrs ... up to
454 8766hrs). Before the *am* data were similarly averaged, values co-incident in time to data gaps in
455 the interplanetary data were masked out using the procedure of *Finch and Lockwood* [2007] and
456 not included in the means. Note that the data in the Omni2 dataset are lagged by the predicted
457 propagation time between the observing satellite and the nose of the magnetosphere (see
458 description of the procedure by *King and Papitashvili* available at
459 https://omniweb.gsfc.nasa.gov/html/omni_min_data.html).

460 The averages of the input power for a given τ , $\langle P_\alpha \rangle_\tau$ were computed using equation (6) for
461 values of the coupling exponent α between 0 and 1.5 in steps of 0.01. The linear correlation
462 coefficient between $\langle P_\alpha \rangle_\tau$ and $\langle am \rangle_\tau$ (with data gaps introduced into *am* to that match those in
463 the P_α data) at each α , $r(\alpha)$, was then determined. Figure 4 presents an example of the results for
464 annual means (i.e. for $\tau = 1$ yr). Figure 4(a) shows the correlogram of $r(\alpha)$ as a function of α .
465 The peak $r(\alpha)$, $r_p = 0.990$, is at $\alpha = \alpha_p = 0.44$ (marked by the vertical green line). Figure 4(b)
466 shows the significance $S(\alpha)$ of the difference between $r(\alpha)$ and its peak value, r_p , evaluated using
467 the Meng-Z test of the significance of the difference between two correlations [*Meng et al.*,
468 1992]. Sometimes also called Steiger's Z test, this is a variant of the Fisher-Z test [e.g. *Asuero et*
469 *al.*, 2006] that makes allowance for inter-correlation between the comparison time series. We
470 use a one-sided version of the test against the null hypothesis that $r(\alpha)$ is not lower than $r(\alpha_p)$.

471 By definition, $S(\alpha) = 0$ at $\alpha = \alpha_p$ and increases at α away from the peak as $r(\alpha)$ falls. The vertical
 472 red and blue lines are where $S(\alpha)$ rises to 0.68 and so mark the 1- σ uncertainties in the α_p value
 473 (giving the optimum α to be $\alpha_p = 0.42$, with 1- σ uncertainties of +0.11 and -0.08 in this case for
 474 $\tau = 1$ yr). Figure 4e shows the time series of $\langle am \rangle_{\tau=1\text{yr}}$ (in black) and the best-fit linear
 475 regression of $\langle P_\alpha \rangle_{\tau=1\text{yr}}$ to $\langle am \rangle_{\tau=1\text{yr}}$ (in mauve) for $\alpha = \alpha_p$, the best fit being $s \langle P_\alpha / P_o \rangle_{\tau=1\text{yr}} + c$,
 476 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
 477 the whole 1996-2016 dataset and is a convenient normalizing factor. The corresponding scatter
 478 plot is shown by the points in Figure 4f and the best-fit linear regression in mauve.

479 Figures 4(c) and 4(d) show an alternative procedure for determining α_p and its uncertainty. The
 480 r.m.s. deviation of am from the best-fit linear regression of $\langle P_\alpha / P_o \rangle_{\tau=1\text{yr}}$ (illustrated in Figures
 481 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
 482 a different metric of the same thing as the correlation coefficient (i.e., the level of agreement
 483 between Am and $\langle P_\alpha / P_o \rangle_{\tau=1\text{yr}}$), but can be used with different statistical tests. As expected, the
 484 two procedures give the same α_p to within ± 0.05 (the resolution of the tests deployed as α was
 485 incremented in steps of 0.1). Figure 4d looks at the significance S of the difference between
 486 ΔAm_{rms} at $\alpha = \alpha_p$ and the values at all other α . This is evaluated using a 2-sample F test of the
 487 variances of the distributions of the fit residuals [*Snedecor and Cochran*, 1989]: this allows us to
 488 estimate the α values at which ΔAm_{rms} is larger than the minimum by an amount that is
 489 significant at a specified level. We use the one-tailed version of the test against the null
 490 hypothesis that ΔAm_{rms} at general α is not greater than that $\alpha = \alpha_p$. The F -test is a parametric
 491 test (i.e., it assumes Gaussian distributions) and is particularly sensitive to the effect of
 492 distortions from that form. Furthermore, a non-normal distribution of fit residuals indicates that
 493 the linear regression may also be invalid as it is a violation of one of the assumptions of the
 494 regression. Hence, although this is a more direct test of the fit, it is particularly sensitive to the
 495 normality of the distributions and it is important to test if the fit residuals are normally
 496 distributed. This is done in the quantile-quantile (q-q) plot against a normal distribution in Figure
 497 4g. For a normal distribution, the points would lie along the diagonal line and Figure 4g shows
 498 the residuals' distribution is quite close to Gaussian in this case. (However, there is a slight “S-

499 shape” which is discussed below in relation to Figure 5 in which it is more pronounced). The
 500 pink and cyan vertical lines in Figure 4c and 4d are at $S = 0.68$ (the 1-sigma level) for this test
 501 and comparison with Figures 4a and 4b shows that the F -test gives slightly larger uncertainties in
 502 the best value of α than does the Meng-Z test. Note that the Meng-Z test is also parametric, but
 503 is not as sensitive to the assumption of Gaussian distributions and gives valid results even if the
 504 distributions tested are only approximately normal.

505 It is important to stress that we are testing the significance of the difference between the
 506 agreement between the optimum fit of $\langle P_\alpha/P_o \rangle_{\tau=1\text{yr}}$ to Am and the fits for other, less optimum,
 507 $\langle P_\alpha/P_o \rangle_{\tau=1\text{yr}}$ data series. This is different to testing the significance of the less-good fits.
 508 Because the optimum fits are so good (correlation coefficient very close to unity and r.m.s. fit
 509 residual close to zero), the agreements for less-good fits are still highly significant in themselves
 510 - even when the difference to the optimum fit is found to be significant. Notice also that we here
 511 use the 1-sigma level to quantify the uncertainty in α and the estimate would naturally be larger
 512 for the 2-sigma or 3-sigma level. This is because 1-sigma is a conservative (pessimistic)
 513 estimate for our application of the uncertainties because they are here used to gauge if α is
 514 constant with τ and using the smaller uncertainties is a stricter test of the constancy of α .

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

527 Figure 6a studies the full evolution of the $r(\alpha)$ correlograms with averaging timescale τ between
 528 3 hours and 1 year, (i.e. averaging intervals of duration $\tau = 3, 6, 9, \dots, 8766$ hours, a total of
 529 2922 different durations). The correlation coefficient is color-coded as a function of τ along the
 530 horizontal axis and α along the vertical axis. The middle white line gives the α_p value which
 531 yields the peak correlation r_p (given by the vertical green lines in the examples presented in
 532 Figures 4 and 5) and the two black lines give the $\pm 1\sigma$ uncertainty points from the Meng-Z test
 533 (given by the vertical red and blue lines in Figures 4 and 5). Figure 6b compares the uncertainty
 534 bands from the two methods or computing the uncertainty, as a function of averaging timescale,
 535 τ , and using the same line colors as Figures 4 and 5. It can be seen that they give very similar
 536 results, except at lower τ when there are concerns about the normality of the distributions,
 537 particularly for the F test. We here use the results from the Meng-Z test and the general similarity
 538 of the F test results gives us confidence in these uncertainty estimates. Figure 6c plots the
 539 corresponding variation with τ of the peak correlation r_p .

540 Range indices, such as *am*, cannot easily be generated for $\tau < 3$ hrs. In addition, at such
 541 timescales the correlations become complicated by lag times with the directly-driven system
 542 responding within about 5 min. of changes in P_α arriving at the dayside magnetopause and the
 543 storage-release system responding, typically, after between 30 min. and an hour (see section 1).
 544 Hence the analysis becomes more complex and depends critically on the combination of current
 545 systems to which a given index responds. Nevertheless the near-constant α found for τ between
 546 3 hours and 1 year in figure 6 is found to apply at $\tau < 3$ hrs for indices that respond primarily to
 547 the auroral electrojet and the substorm current wedge. Figure 8 shows an example of the analysis
 548 used to test this at $\tau = 1$ min. We use the 1-min. resolution SME index generated by the
 549 SuperMAG project [Newell and Gjerloev, 2011]. This index is equivalent to the AE index and
 550 responds strongly to the nightside auroral electrojet, but is compiled from the large SuperMAG
 551 magnetometer network and so does not suffer the nonlinear effect due to the limited latitudinal
 552 coverage of the AE(12) stations (see supplementary materials). Figure 8 shows the correlation
 553 between P_α and SME for $\tau = 1$ min. as a function of coupling exponent, α (horizontal axis) and
 554 time lag, Δt (vertical axis). Positive Δt is when P_α is lagged, i.e. the SME variation follows P_α .
 555 The peak correlation is $r_p = 0.705$ at $\alpha = 0.46$ and $\Delta t = 37$ min. The plot suggests that α for this

556 τ is slightly higher than 0.44 (marked by the vertical white dashed line) but inside the white
557 contour the difference between r and r_p is not significant at even the $1-\sigma$ level. Hence we cannot
558 regard the difference between r_p for $\tau = 1\text{min.}$ and 0.44 as significant. However, it is quite
559 possible that α is indeed slightly greater than 0.44 in this case because SME is influenced by
560 both the directly-driven and the storage/release systems and analysis of indices such as AU that
561 respond primarily to the directly-driven system find a dependence on V_{sw}^n with a smaller
562 exponent n than for the storage-release system. We do know that the dominant response of SME
563 is due to the storage-release system because the optimum lag is $\Delta t = 37 \pm 3$ min., which we
564 equate to the average duration of substorm growth phases for strong solar wind forcing [Li *et al.*,
565 2013]. Note that the correlation is large (>0.65) for a wide range of lags Δt (between about 15
566 min and 1 hour) and this great variability in the length of the substorm growth phase will have
567 reduced the peak correlation at $\Delta t = 37\text{min}$ whereas it would have had no effect on 3-hourly
568 means. This, and equivalent tests at other τ , show that $\alpha = 0.44$ is consistent with the best
569 correlation at all τ between 1 min and 1 year.

570 Figure 6 shows that, at all τ , the optimum α is consistent with a constant value of 0.44. This
571 contrast somewhat with the results of *Finch and Lockwood* [2007], who found an optimum α of
572 $\alpha_p = 0.4$ at $\tau = 3$ for the *am* index hours which fell to $\alpha_p = 0.3$ at $\tau = 1$ year. The major
573 difference between the present study and that by *Finch and Lockwood* is that, as shown by
574 Figure 2, much of the interplanetary data used by *Finch and Lockwood* contained many data gaps
575 of a wide range of durations whereas we here employ data from 20 years that has a relatively
576 small number of relatively short data gaps. P_α is designed to be an estimate of the power input
577 that drives the substorm cycle and the geomagnetic *am* index has been shown to be an excellent
578 indicator of the geomagnetic response to that energy input on timescales longer than the
579 substorm cycle of energy storage and release (typically 1-2 hours) [*Finch et al.*, 2007; *Finch*,
580 2008]. Hence, there is no physical reason why the optimum α should vary with averaging
581 timescale and so *Finch and Lockwood* considered the variation of α with τ a somewhat
582 unsatisfactory result that may have arisen from the presence of large data gaps in interplanetary
583 data. In the next section we introduce synthetic data gaps into the near-continuous data since the
584 start of 1996 to investigate their effects and see if this was indeed the case.

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

587 **4.1. Simulation of the effects of the interplanetary data gaps before 1995**

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

- 597 A. The interplanetary data were averaged into a mean over interval of length $\tau = 3$ hours
 598 only if there are 3 valid hourly means available (by our criterion of sufficient one-minute
 599 samples in the hour to give an error below 5%). The simultaneous *am* data during gaps in
 600 these 3-hourly P_α data series were piecewise removed using the procedure of *Finch and*
 601 *Lockwood* [2017] before both *am* and P_α were averaged into annual means. Similar
 602 piecewise removal of Dst data during solar wind data gaps was used by *Temerin and Li*
 603 [2006].
- 604 B. The full *am* data series (with no piecewise data removal) and all the available hourly P_α
 605 data were used (i.e. the presence of data gaps in P_α was ignored).
- 606 C. The full *am* data series were used and the available 1-hour P_α data were linearly
 607 interpolated to fill the data gaps before both *am* and P_α were averaged. This method was
 608 used (at 10-minute resolution) in the coupling function study by *Temerin and Li* [2002].
- 609 D. The full *am* data series were used and the P_α data interpolated to fill data gaps using the
 610 scheme adopted by *Svalgaard and Cliver* [2005]. Specifically, the hourly means were
 611 calculated and combined into daily means (in Universal Time), even if only one hourly
 612 mean was available; the 27-day Bartels rotation mean was calculated from available daily

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

616 In all cases both the “combine-then-average” and the “average-then-combine” procedures were
 617 studied. This is an additional complication that applies to all 4 methods. We here term annual
 618 values of P_α generated by the former $\langle P_\alpha \rangle_{1\text{yr}}$ and by the latter as $\langle P_\alpha \rangle_{\text{ann}}$. Figure 9a shows the
 619 time series of $\langle P_\alpha \rangle_{1\text{yr}}$ (in red) and $\langle P_\alpha \rangle_{\text{ann}}$ (in blue) (both normalized to ratio of P_0 , their
 620 average value over the whole period) and Figure 9b the scatter plot of $\langle P_\alpha \rangle_{\text{ann}} / P_0$ as a function
 621 of $\langle P_\alpha \rangle_{1\text{yr}} / P_0$. The agreement is very close indeed for after 1995 (the linear correlation
 622 coefficient is 0.99, with a 2- σ uncertainty range of 0.97-0.995) but the increased scatter of the
 623 green points shows it is not so close for the data before 1996 (linear correlation coefficient 0.93,
 624 with a 2- σ uncertainty range of 0.86-0.97). Hence it appears that data gaps also play a role in
 625 creating a difference between these two ways of generating annual estimates of this coupling
 626 function.

627 Figure 10 shows the results of using Method B to deal with data gaps. In this method the
 628 (synthetic) data gaps are effectively ignored and the annual mean P_α data series constructed
 629 from the available data (after removal of the synthetic data gaps), whereas $\langle am \rangle_{1\text{yr}}$ is
 630 constructed using all of the (continuous) am data. This method assumes that the effects of data
 631 gaps average out, and this is the most common way of dealing with the missing data. In Figure
 632 10a, the red and blue lines are the correlograms (correlation coefficient r as a function of
 633 assumed coupling exponent, α) for all the 1996-2016 data (i.e. no synthetic data gaps are
 634 introduced) showing, in red, $r(\alpha)$ for am and $\langle P_\alpha \rangle_{1\text{yr}}$ (the combine-then-average annual
 635 estimates) and, in blue, $r(\alpha)$ for am and $\langle P_\alpha \rangle_{\text{ann}}$ (the average-then-combine annual estimates).
 636 As found by *Lockwood et al.* [2017a] the peak, r_p , is at slightly higher α for $\langle P_\alpha \rangle_{\text{ann}}$ than for
 637 $\langle P_\alpha \rangle_{1\text{yr}}$ ($\alpha_p = 0.44$ and $\alpha_p = 0.48$, marked by the vertical green and orange lines, respectively).
 638 Note that this is despite the fact that, as shown by figure 9, $\langle P_\alpha \rangle_{\text{ann}}$ and $\langle P_\alpha \rangle_{1\text{yr}}$ are very
 639 highly correlated ($r = 0.99$) for this interval. The solid points are the means of the values of $r(\alpha)$
 640 for am and $\langle P_\alpha \rangle_{1\text{yr}}$ from the runs for the $N_S = 500$ different data masks, and the error bars are

641 ± 1 standard deviation. It can be seen that correlations at a given α are generally reduced
 642 compared to that for the full available data series. The mean correlation at the peak (at $\alpha = \alpha_p =$
 643 0.44, marked by the vertical green line) is reduced from 0.99 to $r_{avp} = 0.94$ with a standard
 644 deviation for the N_S simulations of $\sigma_{avp} = 0.04$ and the optimum α is increased to 0.48, compared
 645 to the 0.44 obtained without data gaps. The gray histogram in Figure 10(b) shows the distribution
 646 of the derived α_p values (giving the peak $r = r_p$ between $\langle P_\alpha \rangle_{1yr}$ and $\langle am \rangle_{1yr}$) for the $N_S = 500$
 647 different data gap masks. It can be seen that the mode value of this distribution is $\alpha_m = 0.48$ and
 648 the median $\alpha_{0.5} = 0.43$. It must be remembered that we have introduced 500 sets of data gaps but
 649 only one would have existed in the actual pre-1996 data and so it is really the range of possible
 650 individual values in Figure 10b that we need to consider, rather than the mean, median or mode
 651 of the distribution. The standard deviation of the distribution of the 500 α_p values is $\sigma = 0.116$,
 652 the minimum to maximum range is 0.14 - 1.00 and the lower and upper 2- σ points are $\alpha_{0.05} =$
 653 0.24 and $\alpha_{0.95} = 0.67$. It must be remembered that the only difference between the 500
 654 simulations is when the data gaps happen, by chance, to fall and neglecting the effect of data
 655 gaps (the most commonly used procedure) could generate any one of the α values in this
 656 distribution.

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

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

669 The results or the piecewise removal of *am* data using the procedure of *Finch and Lockwood*
670 [2007] (Method A) are surprisingly similar to those shown in Figure 10 for Method B. On
671 average, the correlations are very slightly increased (r_{avp} is 0.95 instead of 0.94) but the
672 distributions α_p are actually slightly broader for Method A for both $\langle P_\alpha \rangle_{1\text{yr}}$ and $\langle P_\alpha \rangle_{\text{ann}}$.
673 However the distribution medians are closer to the “true” values (data gap free, i.e. 0.44 and
674 0.48) in both cases. Hence we can say that the procedure of *Finch and Lockwood* [2007] is an
675 improvement over neglecting data gaps and averaging over all available data; however, it is only
676 a very small improvement in the context of improving the peak correlation and is slightly more
677 likely to give the correct estimate of the optimum α .

678 The most unsatisfactory method for dealing with the (synthetic) gaps in 1-hour P_α data is to fill
679 them using interpolation (method C). This is here implemented using linear interpolation but
680 results are even worse if cubic splines or PCIP are used. The peak correlations are lower, the
681 distributions of α_p are very wide and the median and mode values of those distributions shifted
682 away from the “true” (i.e., data gap free) value. We note this method was used (on pre-1995
683 data) by *Temerin and Li* [2002]; however, we also note that the same authors soon after moved to
684 using a variant of method A and only data from after 1995 with considerable improvement of the
685 correlation between their coupling function and the Dst geomagnetic index [*Temerin and Li*,
686 2006]. There are many variants of interpolation procedures. Others have used interpolation with
687 restrictions, for example *Wu and Lundstedt* [1997] deployed autocorrelation-based interpolation
688 for data gaps of up to 5 hours, and intervals with longer data gaps excluded from the study (a
689 variant of method A).

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

697 **4.2. For the availability of interplanetary data after 1995**

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

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

710 Figure 11 shows the results in the same format as Figure 10. It can be seen that the correlations
 711 for any of the 500 masks are only slightly reduced compared to the values for no additional data
 712 gaps. The optimum α is unchanged and the 500 masks give a standard deviation σ of only
 713 0.014. It is instructive to compare this to the uncertainty in the α of peak correlation (α_p)
 714 between am and $\langle P_\alpha \rangle_{1\text{yr}}$ for this τ (1 year) with the $+0.11$ to -0.08 $1\text{-}\sigma$ error from the Meng-Z
 715 test, as shown in figure 4. Hence data gaps are contributing to the uncertainty in α_p when data
 716 availability is 83% (post 1995) but are far from the dominant factor and so uncertainties
 717 introduced by the data gaps are smaller than that inherent in the correlations due to instrumental
 718 and geophysical noise in both P_α and am . In contrast, the test of Method A with data gap masks
 719 drawn from before 1995 (when average data availability is 30%) shows that these uncertainties
 720 rise to ± 0.13 and so errors due to data gaps would be dominant.

721 **5. Discussion and Conclusions**

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

725 The first is that the low auto-correlation time constant of the IMF orientation factor means that
 726 hourly samples that do not have 82% data availability within the hour cause errors greater than
 727 5% (at the 1- σ level) in hourly values of coupling functions, such as P_α , and should be treated as
 728 bad data and removed. Note that this error could be reduced by requiring a higher availability,
 729 but this would generate a great many and longer data gaps.

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

739 Using the dataset with the requirement that IMF availability in each hour is 82% (which limits
 740 the uncertainty in hourly P_α values to $\pm 5\%$ at the 1- σ level) we find that the optimum coupling
 741 exponent of $\alpha = \alpha_p = 0.44 \pm 0.02$ at all averaging timescales τ between 1 minute and 1 year. We
 742 find no significant variation of this value with τ . From equation (6), this yields dependencies on
 743 $B^{0.88}$, $V_{sw}^{1.90}$, and $(m_{sw}N_{sw})^{0.23}$. A dependence on B has been found in a great many proposed
 744 coupling functions [e.g., *Kan and Lee, 1979; Wygant et al., 1983; Scurry and Russell, 1991;*
 745 *Temerin and Li, 2006*]. The 1- σ uncertainty on the optimum α found here is ± 0.02 which allows
 746 a variation of anywhere between $B^{0.84}$ and $B^{0.92}$; however, this is not enough to explain the
 747 discrepancy with a dependence on B , nor with the result of *Newell et al. [2007]* who find a $B^{0.67}$
 748 dependence (probably because these authors were attempting to match a wide range of terrestrial
 749 response measures). The dependence found here is close to BV_{sw}^2 and similar to that found in a

750 great many other studies. *Lockwood* [2013] studied the optimum values of n for functions of the
 751 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
 752 optimum value deduced here for *am* is $n = 1.90$ and allowing for the uncertainty in the peak α
 753 (± 0.02) the, n derived here is between 1.75 and 1.91. However, bearing in mind that *Lockwood*
 754 [2013] used annual means of data from 1966 onwards with the *Finch and Lockwood* [2007]
 755 method (A) to deal with data gaps, Table 1 shows a 1- σ uncertainty in α of ± 0.12 would apply to
 756 the *Lockwood* [2013] values which means that n could be between about 1.86 and 2.24 for AL
 757 and between 1.55 and 2.04 for *am*. Hence results of *Lockwood* [2013] for the mid-latitude
 758 station range indices *am*, *ap* and *aa* and the auroral electrojet index AL are not inconsistent with
 759 the result found here for *am* when we allow for the potential effect of data gaps.

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

770 The derived dependence on $(m_{sw}N_{sw})^{0.23}$ is interesting. The exponent of this term would be zero
 771 if $\alpha = 2/3$, removing any dependence on the solar wind mass or number flux. This is because for
 772 $\alpha = 2/3$ the effect of an increase in mass flux ($m_{sw}N_{sw}$) on the particle kinetic energy flux in the
 773 solar wind ($m_{sw}N_{sw}V_{sw}^3$) would be cancelled by the compressional effect of the solar wind
 774 dynamic pressure ($m_{sw}N_{sw}V_{sw}^2$) and the reduction of the target area presented to the solar wind
 775 flow by the geomagnetic field. The lower α of 0.44 means that the first effect is the slightly
 776 larger of the two and there is a (weak) dependence on solar wind mass flux ($m_{sw}N_{sw}$), variations
 777 in which are dominated by the number flux (N_{sw}). Several studies have found that increased N_{sw}

778 increases terrestrial space weather response to IMF changes [e.g., *Lopez et al.* 2004; *Weigel,*
 779 2010] but many coupling functions do not include a term in either $(m_{sw}N_{sw})$ or N_{sw} . *Newell et al.*
 780 [2007] also found a $(m_{sw}N_{sw})^{1/6}$ dependence, but chose to omit it from their coupling function in
 781 the interests of making it match a wide range of terrestrial responses.

782 Plots corresponding to those presented in this paper, but using the geomagnetic ap index rather
 783 than am , are given in the supplementary materials. These plots are very similar indeed to those
 784 for am in all cases. Table 2 shows how ap correlates slightly better with $P\alpha$ on annual averaging
 785 timescales, but am performs slightly better at 27 days, 1 day and 3 hours. This is consistent with
 786 an analysis of the two indices that have carried out showing that the empirical, tabular K -to- K_s
 787 conversions used to make K_p (and hence ap) introduce spurious diurnal and annual variations,
 788 but these are averaged out on annual timescales, on which ap gains an advantage in noise
 789 suppression by averaging data from a more concentrated cluster of stations in Europe. The best
 790 fit α values are higher for ap than for am by between 0.02 and 0.06. Figure 7 investigates the
 791 implications of the precise value of α . Parts (c) and (d) of Figure 7 demonstrate that the
 792 difference in α_p for the two indices makes almost no difference to the distribution of the
 793 normalized power input to the magnetosphere, P_α/P_o , but part (b) shows it does influence the
 794 average value over the whole interval 1996-2016, P_o , and hence the absolute magnitude of the
 795 power input which is estimated. Because am has a more uniform time-of-day-time-of-year
 796 response pattern, we prefer the value obtained for am , which is $P_o = (0.38\pm 0.06)\times 10^{19}$ W (at the
 797 2σ uncertainty level). However, the sensitivity of this value to the index used illustrates how
 798 difficult it is to get an absolute power input value accurately even though the variation in P_α/P_o is
 799 well defined.

800 *Lockwood et al.* [2017] have estimated the annual mean power input into the magnetosphere for
 801 all years back to 1612 from the reconstructed solar wind and interplanetary field parameters
 802 derived by *Owens et al.* [2017], and from this *Lockwood et al.* [2018a] have derived the annual
 803 means of Ap and AE back to this date. In the two subsequent papers in the present series
 804 [*Lockwood et al.*, 2018b; c] we begin to construct a space weather climatology by studying the
 805 distributions of space weather parameters about these averages and, in particular, how these
 806 distributions evolve with timescale. The present paper is an important first step in this because

807 it shows that the formula for the optimum coupling function does not significantly evolve with
808 timescale. Previous studies that suggested there was such a variation had been influenced by data
809 gaps which have different effects on different timescales. Removing this potential complication
810 (by showing that the optimum coupling exponent α is, in fact, independent of timescale) is a
811 valuable first step in the construction of a useful space weather climatology.

812 **Acknowledgments and Data**

813 The authors are grateful to the staff of Space Physics Data Facility, NASA/Goddard
814 Space Flight Center, who prepared and made available the OMNI2 dataset used. The data were
815 downloaded from <http://omniweb.gsfc.nasa.gov/ow.html>. They are also grateful to the staff of
816 L'École et Observatoire des Sciences de la Terre (EOST), a joint of the University of Strasbourg
817 and the French National Center for Scientific Research (CNRS) and the International Service of
818 Geomagnetic Indices (ISGI) for making the *am* index data available from
819 http://isgi.unistra.fr/data_download.php. We also thank GeoForschungsZentrum (GFZ)
820 Potsdam, Adolf-Schmidt-Observatorium für Geomagnetismus, Niemegk, Germany who generate
821 the *ap* data. The *ap* data were downloaded from the UK Solar System Data Centre from
822 <https://www.ukssdc.ac.uk/> with updating of recent data from BGS Edinburgh
823 http://www.geomag.bgs.ac.uk/data_service/data/magnetic_indices/apindex.html. For the
824 SuperMAG indices data we gratefully acknowledge: Intermagnet; USGS, Jeffrey J. Love;
825 CARISMA, PI Ian Mann; CANMOS; The S-RAMP Database, PI K. Yumoto and Dr. K.
826 Shiokawa; The SPIDR database; AARI, PI Oleg Troshichev; The MACCS program, PI M.
827 Engebretson, Geomagnetism Unit of the Geological Survey of Canada; GIMA; MEASURE,
828 UCLA IGPP and Florida Institute of Technology; SAMBA, PI Eftyhia Zesta; 210 Chain, PI K.
829 Yumoto; SAMNET, PI Farideh Honary; The institutes who maintain the IMAGE magnetometer
830 array, PI Eija Tanskanen; PENGUIN; AUTUMN, PI Martin Connors; DTU Space, PI Dr. Rico
831 Behlke; South Pole and McMurdo Magnetometer, PI's Louis J. Lanzarotti and Alan T.
832 Weatherwax; ICESTAR; RAPIDMAG; PENGUIn; British Antarctic Survey; McMac, PI Dr.
833 Peter Chi; BGS, PI Dr. Susan Macmillan; Pushkov Institute of Terrestrial Magnetism,
834 Ionosphere and Radio Wave Propagation (IZMIRAN); GFZ, PI Dr. Juergen Matzka; MFGI, PI
835 B. Heilig; IGFPAS, PI J. Reda; University of L'Aquila, PI M. Vellante; BCMT, V. Lesur and A.
836 Chambodut; Data obtained in cooperation with Geoscience Australia, PI Marina Costelloe;
837 SuperMAG, PI Jesper W. Gjerloev. The work presented in this paper is supported by STFC
838 consolidated grant number ST/M000885/1, the work of ML and MJO is also supported by the
839 SWIGS NERC Directed Highlight Topic Grant number NE/P016928/1/ and of OA by NERC
840 grant NE/P017274/1. SB is supported by an NERC PhD studentship.

841 **References**

- 842 Adebessin, B.O. (2016) Investigation into the linear relationship between the AE, Dst and ap
843 indices during different magnetic and solar activity conditions, *Acta Geod Geophys*, 51
844 (2), 315–331, doi: 10.1007/s40328-015-0128-2
- 845 Arnoldy, R.L. (1971), Signature in the interplanetary medium for substorms, *J. Geophys. Res.*,
846 76 (22), 5189–5201, doi:10.1029/JA076i022p05189.
- 847 Asuero, A.G., A. Sayago & A. G. González (2006) The Correlation Coefficient: An Overview,
848 *Critical Reviews in Analytical Chemistry*, 36 (1), 41-59, DOI:
849 10.1080/10408340500526766
- 850 Baker, D. N., E. W. Hones, Jr., J. B. Payne, and W. C. Feldman (1981) A high time resolution
851 study of interplanetary parameter correlations with AE, *Geophys. Res. Lett.*, 8, 179-182,
852 doi: 10.1029/gl008i002p00179
- 853 Baker, D. N., R. D. Zwickl, S. J. Bame, E. W. Hones Jr., B. T. Tsurutani, E. J. Smith, and S.-I.
854 Akasofu (1983) An ISEE 3 high time resolution study of interplanetary parameter
855 correlations with magnetospheric activity, *J. Geophys. Res.*, 88 (A8), 6230–6242,
856 doi:10.1029/JA088iA08p06230.
- 857 Balikhin, M. A., R. J. Boynton, S. A. Billings, M. Gedalin, N. Ganushkina, D. Coca, and H. Wei
858 (2010), Data based quest for solar wind–magnetosphere coupling function, *Geophys.*
859 *Res. Lett.*, 37, L24107, doi:10.1029/2010GL045733.
- 860 Bargatze, L. F., D. N. Baker, R. L. McPherron, and E. W. Hones Jr. (1985), Magnetospheric
861 impulse response for many levels of geomagnetic activity, *J. Geophys. Res.*, 90 (A7),
862 6387–6394, doi:10.1029/JA090iA07p06387.
- 863 Bargatze, L. F., R. L. McPherron, and D. N. Baker (1986), Solar wind-magnetosphere energy
864 input functions, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A.
865 Slavin, pp. 93–100, Terrapub/Reidel, Tokyo, Japan, doi: 10.1007/978-94-009-4722-1_7
- 866 Bartels, J., Heck, N.H. and Johnston, H.F. (1939) The three-hour-range index measuring
867 geomagnetic activity, *Terr. Magn. Atmos. Electr.*, 44, 411–454, doi:
868 10.1029/TE044i004p00411
- 869 Bartels, J. (1949) The standardized index Ks and the planetary index Kp, IATME Bulletin 12b,
870 97.
- 871 Bartels, J. (1957) The geomagnetic measures for the time-variations of solar corpuscular
872 radiation, described for use in correlation studies in other geophysical fields, *Ann. Intern.*
873 *Geophys. Year 4*, 227-236
- 874 Bochsler, P. (1987) Solar-wind ion composition, *Physica Scripta*, T18, 55-60, doi:
875 10.1088/0031-8949/1987/T18/007
- 876 Borovsky, J. E. (2008) The rudiments of a theory of solar wind/ magnetosphere coupling derived
877 from first principles, *J. Geophys. Res.*, 113, A08228, doi:10.1029/2007JA012646.

- 878 Burton, R.K., R.L. McPherron, and C.T. Russell (1975), An empirical relationship between
879 interplanetary conditions and Dst, *J. Geophys. Res.*, 80, 4204-4214, doi:
880 10.1029/ja080i031p04204.
- 881 Couzens, D.A. and J.H., King (1986) Interplanetary Medium Data Book – Supplement 3, 1977 –
882 1985, NSSDC/WDC-A/R&S, 86-04, NASA, Greenbelt, MD, 1986,
883 http://www.archive.org/details/nasa_techdoc_19890001401
- 884 Cowley, S.W.H. (1991) Acceleration and heating of space plasmas - Basic concepts, *Annales*
885 *Geophys.*, 9, 176-187.
- 886 Cowley, S.W.H. and M. Lockwood (1992) Excitation and decay of solar-wind driven flows in
887 the magnetosphere-ionosphere system, *Annales Geophys.*, 10, 103-115
- 888 Crooker, N. U., J. Feynman, and J. T. Gosling (1977) On the high correlation between long-term
889 averages of solar wind speed and geomagnetic activity, *J. Geophys. Res.*, 82, 1933-1937,
890 doi: 10.1029/ja082i013p01933.
- 891 Dieminger, W., G.K. Hartmann, and R. Leitinger (1996) Geomagnetic Activity Indices, in *The*
892 *Upper Atmosphere* (Eds. W. Dieminger et al.), 887-911, Springer, Berlin Heidelberg, doi:
893 10.1007/978-3-642-78717-1_26
- 894 Dods, J., S. C. Chapman, and J. W. Gjerloev (2015), Network analysis of geomagnetic substorms
895 using the SuperMAG database of ground-based magnetometer stations, *J. Geophys. Res.*
896 *Space Physics*, 120, 7774–7784, doi: 10.1002/2015JA021456.
- 897 Dods, J., S. C. Chapman, and J. W. Gjerloev (2017), Characterizing the ionospheric current
898 pattern response to southward and northward IMF turnings with dynamical SuperMAG
899 correlation networks, *J. Geophys. Res. Space Physics*, 122, 1883–1902,
900 doi:10.1002/2016JA023686.
- 901 Etemadi, A., S.W.H. Cowley, M. Lockwood, B.J.I. Bromage, D.M. Willis, and H. Lühr (1988)
902 The dependence of high-latitude dayside ionospheric flows on the north-south component
903 of the IMF: a high time resolution correlation analysis using EISCAT "POLAR" and
904 AMPTE UKS and IRM data, *Planet. Space Sci.*, 36, 471-498, doi: 10.1016/0032-
905 0633(88)90107-9
- 906 Farrugia, C.J., M.P. Freeman, S.W.H. Cowley, D.J. Southwood, M. Lockwood and A. Etemadi
907 (1989) Pressure-driven magnetopause motions and attendant response on the ground,
908 *Planet. Space Sci.*, 37, 589-608, doi: 10.1016/0032-0633(89)90099-8
- 909 Feynman, J. and N. U. Crooker (1978) The solar wind at the turn of the century, *Nature*, 275,
910 626 – 627, doi: 10.1038/275626a0
- 911 Finch, I.D. (2008) The use of geomagnetic activity observations in studies of solar wind-
912 magnetosphere coupling and centennial solar change, *PhD thesis, University of*
913 *Southampton*, British Library Ethos ID: uk.bl.ethos.485008
- 914 Finch, I.D., and M. Lockwood (2007), Solar wind-magnetosphere coupling functions on
915 timescales of 1 day to 1 year, *Annales Geophys.*, 25, 495-506, doi: 10.5194/angeo-25-
916 495-2007

- 917 Finch, I.D., M. Lockwood, A. P. Rouillard (2008) , The effects of solar wind magnetosphere
918 coupling recorded at different geomagnetic latitudes: separation of directly-driven and
919 storage/release systems, *Geophys. Res. Lett.*, 35, L21105, doi: 10.1029/2008GL035399
- 920 Fritsch, F.N. and R.E. Carlson (1980) Monotone Piecewise Cubic Interpolation, *SIAM Journal*
921 *on Numerical Analysis.*, 17, .238–246, doi: 10.1137/0717021
- 922 George, S.V., G. Ambika, R. Misra (2015) Effect of data gaps on correlation dimension
923 computed from light curves of variable stars *Astrophys. and Space Sci.*, 360 (1), 5, doi:
924 10.1007/s10509-015-2516-z
- 925 Gonzalez, W.D. (1990) A unified view of solar wind-magnetosphere coupling functions, *Planet.*
926 *Space Sci.*, 38(5) , 627-632, 10.1016/0032-0633(90)90068-2
- 927 Hapgood, M.A., G. Bowe, M. Lockwood, D.M. Willis, and Y. Tulunay (1991) Variability of the
928 interplanetary magnetic field at 1 A.U. over 24 years: 1963 – 1986, *Planet. Space Sci.*,
929 39, 411-423, doi: 10.1016/0032-0633(91)90003-S
- 930 Henn, B., M.S. Raleigh, A. Fisher, and J.D. Lundquist (2013) A Comparison of Methods for
931 Filling Gaps in Hourly Near-Surface Air Temperature Data, *J. Hydromet.*, 14(3) on pages
932 929-945, doi: 10.1175/jhm-d-12-027.1
- 933 Kan, J. R., and L. C. Lee (1979), Energy coupling and the solar wind dynamo, *Geophys. Res.*
934 *Lett.*, 6, 577-580, doi: 10.1029/gl006i007p00577
- 935 Kasper, J.C., M.L. Stevens, A.J. Lazarus, J.T. Steinberg, K.W. Ogilvie (2007) Solar Wind
936 Helium Abundance as a Function of Speed and Heliographic Latitude: Variation through
937 a Solar Cycle, *Astrophys. J.*, 660 (1), 901 – 910, doi: 10.1086/510842
- 938 King, J. H., and N. E. Papitashvili (1994) Interplanetary Medium Data Book, Supplement 5,
939 1988-1993, NSSDC/WDC-A-R&S 94-08, NASA/National Space Science Data Center,
940 GSFC, Greenbelt, Maryland, 1994.
- 941 King, J.H., and N.E. Papitashvili (2005), Solar wind spatial scales in and comparisons of hourly
942 Wind and ACE plasma and magnetic field data, *J. Geophys. Res.*, 110, A02104,
943 doi:10.1029/2004JA010649
- 944 Knape, J., and P. de Valpine (2011) Effects of weather and climate on the dynamics of animal
945 population time series. *Proc. Roy. Soc. London B: Biological Sciences*, 278 (1708), 985-
946 992, 2011, doi: 10.1098/rspb.2010.1333
- 947 Knutti, R., G.A. Meehl, M.R. Allen, and D.A. Stainforth (2006) Constraining climate sensitivity
948 from the seasonal cycle in surface temperature. *J. Climate*, 19 (17), 4224-4233, 2006,
949 doi: 10.1175/JCLI3865.1
- 950 Kondrashov, D., Y. Shprits, and M. Ghil (2010) Gap filling of solar wind data by singular
951 spectrum analysis, *Geophys. Res. Lett.*, 37, L15101, doi:10.1029/2010GL044138.
- 952 Kondrashov, D., R. Denton, Y. Y. Shprits, and H. J. Singer (2014), Reconstruction of gaps in the
953 past history of solar wind parameters, *Geophys. Res. Lett.*, 41, 2702–2707,
954 doi:10.1002/2014GL059741.

- 955 Li, H., C. Wang, and Z. Peng (2013), Solar wind impacts on growth phase duration and substorm
956 intensity: A statistical approach, *J. Geophys. Res. Space Physics*, 118, 4270–4278,
957 doi:10.1002/jgra.50399.
- 958 Lockwood, N. (2003) Twenty-three cycles of changing open solar flux, *J. Geophys. Res.*, 108
959 (A3), 1128, doi: 10.1029/2002JA009431
- 960 Lockwood, M. (2004) Solar Outputs, their variations and their effects of Earth
961 in "The Sun, Solar Analogs and the Climate", Proc. Saas-Fee Advanced Course, 34 by
962 J.D. Haigh, M.Lockwood and M.S. Giampapa, eds. I. Rüedi, M. Güdel, and W. Schmutz,
963 pp107-304, Springer, ISBN: 3-540-23856-5
- 964 Lockwood, M. (2013) Reconstruction and Prediction of Variations in the Open Solar Magnetic
965 Flux and Interplanetary Conditions, *Living Rev. Solar Physics*, 10, 4, 2013. doi:
966 10.12942/lrsp-2013-4
- 967 Lockwood, M. and S.W.H. Cowley (1992) Ionospheric Convection and the substorm cycle, in
968 *Substorms I, Proceedings of the First International Conference on Substorms, ICS-1*, ed
969 C. Mattock, ESA-SP-335, 99-109, European Space Agency Publications, Noordwijk, The
970 Netherlands
- 971 Lockwood, S.W.H. Cowley, and M.P. Freeman (1990) The excitation of plasma convection in
972 the high latitude ionosphere, *J. Geophys Res.*, 95, 7961-7971, doi:
973 10.1029/JA095iA06p07961
- 974 Lockwood, M., A.P. van Eyken, B.J.I. Bromage, D.M. Willis, and S.W.H. Cowley (1996)
975 Eastward propagation of a plasma convection enhancement following a southward
976 turning of the interplanetary magnetic field, *Geophys. Res. Lett.*, 13, 72-75, doi:
977 10.1029/GL013i001p00072
- 978 Lockwood, M., R. Stamper and M.N. Wild (1999) A doubling of the sun's coronal magnetic field
979 during the last 100 years, *Nature*, 399, 437-439, doi: 10.1038/20867, 1999
- 980 Lockwood, M., A.P. Rouillard, I.D. Finch, and R. Stamper (2006) Comment on "The IDV
981 index: its derivation and use in inferring long-term variations of the interplanetary
982 magnetic field strength" by Svalgaard and Cliver, *J. Geophys. Res.*, 111, A09109,
983 doi:10.1029/2006JA011640
- 984 Lockwood, M., H. Nevanlinna, L. Barnard, M.J. Owens, R.G. Harrison, A.P. Rouillard, and C.J.
985 Scot (2014) Reconstruction of Geomagnetic Activity and Near-Earth Interplanetary
986 Conditions over the Past 167 Years: 4. Near-Earth Solar Wind Speed, IMF, and Open
987 Solar Flux, *Annales. Geophys.*, 32, 383-399, doi:10.5194/angeo-32-383-2014
- 988 Lockwood, M., M.J. Owens, L.A. Barnard S. Bentley, C.J. Scott, and C.E. Watt, (2016) On the
989 Origins and Timescales of Geoeffective IMF, *Space Weather*, 14, 406–432, , doi:
990 10.1002/2016SW001375
- 991 Lockwood, M., M.J. Owens, L.A. Barnard, C.J. Scott, and C.E. Watt (2017) Space Climate and
992 Space Weather over the past 400 years: 1. The Power input to the Magnetosphere, *J.*
993 *Space Weather Space Clim.*, 7, A25, doi: 10.1051/swsc/2017019

- 994 Lockwood, M., M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt and S. Bentley (2018a) Space
995 Climate and Space Weather over the past 400 years: 2. Proxy indicators of geomagnetic
996 storm and substorm occurrence, *J. Space Weather Space Clim.*, in press, doi:
997 10.1051/swsc/2017048
- 998 Lockwood, M., S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt and O. Allanson
999 (2018b) The development of a space climatology: 2. the variation of space weather
1000 parameters with timescale, *Space Weather*, submitted with present paper
- 1001 Lockwood, M., S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt and O. Allanson
1002 (2018c) The development of a space climatology: 3. The evolution of distributions of
1003 space weather parameters with timescale
- 1004 Lopez, R. E., M. Wiltberger, S. Hernandez, and J. G. Lyon (2004), Solar wind density control of
1005 energy transfer to the magnetosphere, *Geophys. Res. Lett.*, 31, L08804,
1006 doi:10.1029/2003GL018780.
- 1007 Luo, B., X. Li, M. Temerin, and S. Liu (2013), Prediction of the AU, AL, and AE indices using
1008 solar wind parameters, *J. Geophys. Res. Space Physics*, 118, 7683–7694,
1009 doi:10.1002/2013JA019188.
- 1010 Mayaud, P.-N. (1980) Derivation, Meaning and Use of Geomagnetic Indices, Geophysical
1011 Monograph, 22, American Geophysical Union, Washington, DC, doi: 10.1029/GM022
- 1012 Meng, X.-I., R. Rosenthal and D. B. Rubin (1992), Comparing Correlated Correlation
1013 Coefficients, *Psych. Bulletin*, 111 (1), 172-175, doi: 10.1037//0033-2909.111.1.172
- 1014 Menvielle, M., and A. Berthelier (1991) The K-derived planetary indices: Description and
1015 availability, *Rev. Geophys.*, 29 (3), 415–432, doi:10.1029/91RG00994.
- 1016 Milan S.E., M. Lester, S.W.H. Cowley, K. Oksavik, M. Brittnacher, R.A. Greenwald, G.Sofko,
1017 and J.-P. Villain (2003) Variations in polar cap area during two substorm cycles, *Ann.*
1018 *Geophys.*, 21, 1121-1140, doi: 10.5194/angeo-21-1121-2003
- 1019 Milan S.E., P.D. Boakes, and B. Hubert (2008) Response of the expanding/contracting polar cap
1020 to weak and strong solar wind driving: implications for substorm onset, *J. Geophys. Res.*,
1021 113, A09215, doi: 10.1029/2008JA013340.
- 1022 McPherron R.L., G. Siscoe G and C.N. Arge (2004) Probabilistic forecasting of the 3-h Ap
1023 index, *IEEE Trans. Plasma Sci.* 32 1425, doi: 10.1109/TPS.2004.833387
- 1024 McPherron, R. L., T.-S. Hsu, and X. Chu (2015) An optimum solar wind coupling function for
1025 the AL index, *J. Geophys. Res. Space Physics*, 120, 2494–2515,
1026 doi:10.1002/2014JA020619.
- 1027 Munteanu, C., C. Negrea, M. Echim, and K. Mursula (2016) Effect of data gaps: comparison of
1028 different spectral analysis methods, *Ann. Geophys.*, 34, 437–449, doi: 10.5194/angeo-34-
1029 437-2016
- 1030 Murayama, T. (1982) Coupling function between solar-wind parameters and geomagnetic
1031 indexes, *Rev. Geophys.*, 20 (3), 623–629, doi: 10.1029/RG020i003p00623.

- 1032 Nagatsuma, T. (2006) Diurnal, semiannual, and solar cycle variations of solar wind
1033 magnetosphere-ionosphere coupling, *J. Geophys. Res.*, 111, A09202, doi:
1034 10.1029/2005JA011122, 2006
- 1035 Newell, P. T., and J. W. Gjerloev (2011), Evaluation of SuperMAG auroral electrojet indices as
1036 indicators of substorms and auroral power, *J. Geophys. Res.*, 116, A12211,
1037 doi:10.1029/2011JA016779.
- 1038 Newell, P.T., T. Sotirelis, K. Liou, C.-I. Meng, and F. J. Rich (2007) A nearly universal solar
1039 wind-magnetosphere coupling function inferred from 10 magnetospheric state variables,
1040 *J. Geophys. Res.*, 112, A01206, doi:10.1029/2006JA012015.
- 1041 Nishida, A. (1968), Coherence of geomagnetic DP 2 fluctuations with interplanetary magnetic
1042 variations, *J. Geophys. Res.*, 73 (17), 5549–5559, doi:10.1029/JA073i017p05549
- 1043 Owens, M.J., M. Lockwood, and P. Riley (2017) Global solar wind variations over the last four
1044 centuries, *Nature Scientific Reports*, 7, Article number 41548, doi: 10.1038/srep41548
- 1045 Papitashvili, V.O., N.E. Papitashvili, and J.H. King (2000), Solar cycle effects in planetary
1046 geomagnetic activity: Analysis of 36-year long OMNI dataset, *Geophys. Res. Lett.*, 27,
1047 2797-2800, doi: 10.1029/2000gl000064.
- 1048 Perreault, W. K., and S-I. Akasofu (1978) A study of geomagnetic storms, *Geophys. J. Int.*, 54,
1049 547-573, 10.1111/j.1365-246x.1978.tb05494.x.
- 1050 Reiff, P.H., R.W. Spiro, and T.W. Hill (1981) Dependence of polar cap potential drop on
1051 interplanetary parameters, *J. Geophys. Res.*, 86, 7639-7648, doi:
1052 10.1029/ja086ia09p07639.
- 1053 Rouillard, A.P., M. Lockwood and I.D. Finch (2007) Centennial changes in the solar wind speed
1054 and in the open solar flux. *J. Geophys. Res.*, 112, A05103, doi: 10.1029/2006JA012130.
- 1055 Saba, M.M.F., W.D. Gonzalez, A.L. Clua de Gonzalez (1997) Relationships between the AE, ap
1056 and Dst indices near solar minimum (1974) and at solar maximum (1979), *Annales*
1057 *Geophys.*, 15 (10), 1265-1270.
- 1058 Schatten, K. H., and J. M. Wilcox (1967) Response of the geomagnetic activity index Kp to the
1059 interplanetary magnetic field, *J. Geophys. Res.*, 72 (21), 5185–5191, doi:
1060 10.1029/JZ072i021p05185.
- 1061 Scurry, L., and C.T. Russell (1991), Proxy studies of energy transfer to the magnetopause, *J.*
1062 *Geophys. Res.*, 96 (A6), 9541–9548, doi:10.1029/91JA00569.
- 1063 Snedecor, G.W. and W.G. Cochran (1989) *Statistical Methods*, 8th edition, Ames: Iowa State
1064 University Press, ISBN 0 8138 1561 6
- 1065 Spencer, E., P. Kasturi, S. Patra, W. Horton, and M. L. Mays (2011), Influence of solar wind–
1066 magnetosphere coupling functions on the Dst index, *J. Geophys. Res.*, 116, A12235, doi:
1067 10.1029/2011JA016780.
- 1068 Stamper, R., M. Lockwood, M.N. Wild, and T.D.G. Clark (1999) Solar Causes of the Long Term
1069 Increase in Geomagnetic Activity, *J. Geophys Res.*, 104, 28325-28342, doi:
1070 10.1029/1999JA900311

- 1071 Sturges, W. (1983) On interpolating gappy records for time series analysis, *J. Geophys. Res.*
1072 (*Oceans*), 88 0736-9740, doi: 10.1029/jc088ic14p09736
- 1073 Svalgaard, L. and E.W. Cliver (2005) The IDV index: Its derivation and use in inferring long-
1074 term variations of the interplanetary magnetic field strength, *J. Geophys. Res.*, 110,
1075 A12103, 2005, doi: 10.1029/2005JA011203
- 1076 Temerin, M., and X. Li, A new model for the prediction of Dst on the basis of the solar wind , *J.*
1077 *Geophys. Res.*, 107 (A12), 1472, doi:10.1029/2001JA007532, 2002.
- 1078 Temerin, M., and X. Li (2006) Dst model for 1995– 2002, *J. Geophys. Res.*, 111, A04221,
1079 doi:10.1029/2005JA011257.
- 1080 Tenfjord, P., and N. Ostgaard (2013), Energy transfer and flow in the solar wind-magnetosphere-
1081 ionosphere system: A new coupling function, *J. Geophys. Res. Space Physics*, 118 (9),
1082 5659–5672, doi:10.1002/Jgra.50545.
- 1083 Teodorescu, E., M. Echim, C. Munteanu, T. Zhang, R. Bruno and P. Kovacs (2015) Inertial
1084 range turbulence of fast and slow solar wind at 0.72 AU and solar minimum, *Astrophys.*
1085 *J. Lett.*, 804, doi:10.1088/2041-8205/804/2/L41, 2015.
- 1086 Thébault, E., C.C. Finlay, C.D. Beggan, P. Alken, J. Aubert, O. Barrois, F. Bertrand, T. Bondar,
1087 A. Boness, L. Brocco, E. Canet, A. Chambodut, A. Chulliat, P. Coïsson, F. Civet, A. Du,
1088 A. Fournier, I. Fratter, N. Gillet, B. Hamilton, M. Hamoudi, G. Hulot, T. Jager, M. Korte,
1089 W. Kuang, X. Lalanne, B. Langlais, J.-M. Léger, V. Lesur, F.J. Lowes, S. Macmillan, M.
1090 Manda, C. Manoj, S. Maus, N. Olsen, V. Petrov, V. Ridley, M. Rother, T.J. Sabaka, D.
1091 Saturnino, R. Schachtschneider, O. Sirol, A. Tangborn, A. Thomson, L. Tøffner-Clausen,
1092 P. Vigneron, I. Wardinski and T. Zvereva (2015) International Geomagnetic Reference
1093 Field: the 12th generation, *Earth, Planets and Space*, 67, 79, doi: 10.1186/s40623-015-
1094 0228-9
- 1095 Thomsen, M. F. (2004), Why Kp is such a good measure of magnetospheric convection?, *Space*
1096 *Weather*, 2, S11004, doi:10.1029/2004SW000089.
- 1097 Todd, H., S.W.H. Cowley, M. Lockwood, D.M. Willis, and H. Lühr (1988) Response time of the
1098 high-latitude dayside ionosphere to sudden changes in the north-south component of the
1099 IMF, *Planet. Space Sci.*, 36, 1415-1428, doi: 10.1016/0032-0633(88)90008-6
- 1100 Vassiliadis, D., A. J. Klimas, D. N. Baker, and D. A. Roberts (1995) A description of the solar
1101 wind-magnetosphere coupling based on nonlinear filters, *J. Geophys. Res.*, 100 (A3),
1102 3495–3512, doi:10.1029/94JA02725.
- 1103 Vasyliunas, V. M., J. R. Kan, G. L. Siscoe, and S.-I. Akasofu (1982) Scaling relations governing
1104 magnetospheric energy transfer, *Planet. Space Sci.*, 30, 359–365, doi: 10.1016/0032-
1105 0633(82)90041-1
- 1106 Wang, C., J. P. Han, H. Li, Z. Peng, and J. D. Richardson (2014), Solar wind-magnetosphere
1107 energy coupling function fitting: Results from a global MHD simulation, *J. Geophys. Res.*
1108 *Space Physics*, 119, 6199–6212, doi:10.1002/2014JA019834.
- 1109 Wang, Y.-M. (2016) Role of the coronal Alfvén speed in modulating the solar-wind helium
1110 abundance, *Astrophys. J. Lett.*, 833 (2) L21, doi: 10.3847/2041-8213/833/2/L21

- 1111 Weigel, R. S. (2010), Solar wind density influence on geomagnetic storm intensity, *J. Geophys.*
1112 *Res.*, 115, A09201, doi:10.1029/2009JA015062
- 1113 Wing, S., and D. G. Sibeck (1997) Effects of interplanetary magnetic field z component and the
1114 solar wind pressure on the geosynchronous magnetic field, *J. Geophys. Res.*, 102, 7207-
1115 7216, doi: 10.1029/97ja00150
- 1116 Wu, J.-G., and H. Lundstedt (1997), Neural network modeling of solar wind-magnetosphere
1117 interaction, *J. Geophys. Res.*, 102 (A7), 14457–14466, doi:10.1029/97JA01081.
- 1118 Wygant, J. R., R. B. Torbert, and F. S. Mozer (1983), Comparison of S3-3 polar cap potential
1119 drops with the interplanetary magnetic field and models of magnetopause reconnection,
1120 *J. Geophys. Res.*, 88 (A7), 5727–5735, doi:10.1029/JA088iA07p05727.
- 1121 Wynn T.A. and V.B. Wickwar (2007) The effects of large data gaps on estimating linear trend in
1122 autocorrelated data, in *Annual Fellowship Symposium of the Rocky Mountain NASA*
1123 *Space Grant Consortium*, Salt Lake City, Utah, May 2007.
1124 http://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1000&context=atmlidar_conf
- 1125 Xie, H., N. Gopalswamy, O. C. St. Cyr, and S. Yashiro (2008), Effects of solar wind dynamic
1126 pressure and preconditioning on large geomagnetic storms, *Geophys. Res. Lett.*, 35,
1127 L06S08, doi:10.1029/2007GL032298.
- 1128 Xu, D., T. Chen, X.X. Zhang, and Z. Liu (2009) Statistical relationship between solar wind
1129 conditions and geomagnetic storms in 1998–2008, *Planet. Space Sci.*, 57 (12), 1500–
1130 1513. doi:10.1016/j.pss.2009.07.015
1131

1132 **Table 1.** Comparison of the performance of methods to handle data gaps or annual data. The
 1133 rows, from top to bottom, give: (1) the ranking order of the accuracy the method; (2) the interval
 1134 used to give data gap masks; (3) the number of masked data simulations, N_s ; (4) the peak value
 1135 of the average correlation coefficient, r_{avp} , from the N_s simulations, plus and minus (when $N_s > 1$)
 1136 the standard deviation at that peak, σ_{avp} (shown in Figure 10a for the case of Method C by the
 1137 black dot and error bar aligned with the vertical green line); (5) and (10) the value of α giving
 1138 that peak, α_{avp} (marked by the vertical green line); (6) and (11) the mode value α_m of the
 1139 distribution (given in Figure 10b or 10c for Method C) of the N_s individual α values giving peak
 1140 correlation; (7) and (12) the standard deviation of that distribution, σ ; (8) and (13) the median
 1141 value of that distribution, $\alpha_{0.5}$; (9) and (14) the range between the upper and lower 2- σ points of
 1142 that distribution, $(\alpha_{0.05}-\alpha_{0.95})$. Rows (4) to (9) apply to the combine-then-average annual means,
 1143 $\langle P_\alpha \rangle_{1yr}$ (the distribution shown in Figure 10b for Method C) and rows (10) to (14) apply to the
 1144 combine-then-average annual means, $\langle P_\alpha \rangle_{ann}$ (the distribution shown in Figure 10c for Method
 1145 C). Row (15) gives the relevant figure. The columns are for: the single analysis with no
 1146 simulated data gaps (the ideal case that the methods are trying to reproduce) and Methods A-D.
 1147 Method A is applied with data masks drawn from both 1995-2015 and from 1966-1994. The
 1148 other methods are only used with masks from 1966-1994.

1149

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, $\langle P_\alpha \rangle_{1yr}$	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
	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, $\langle P_\alpha \rangle_{ann}$	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	

1150

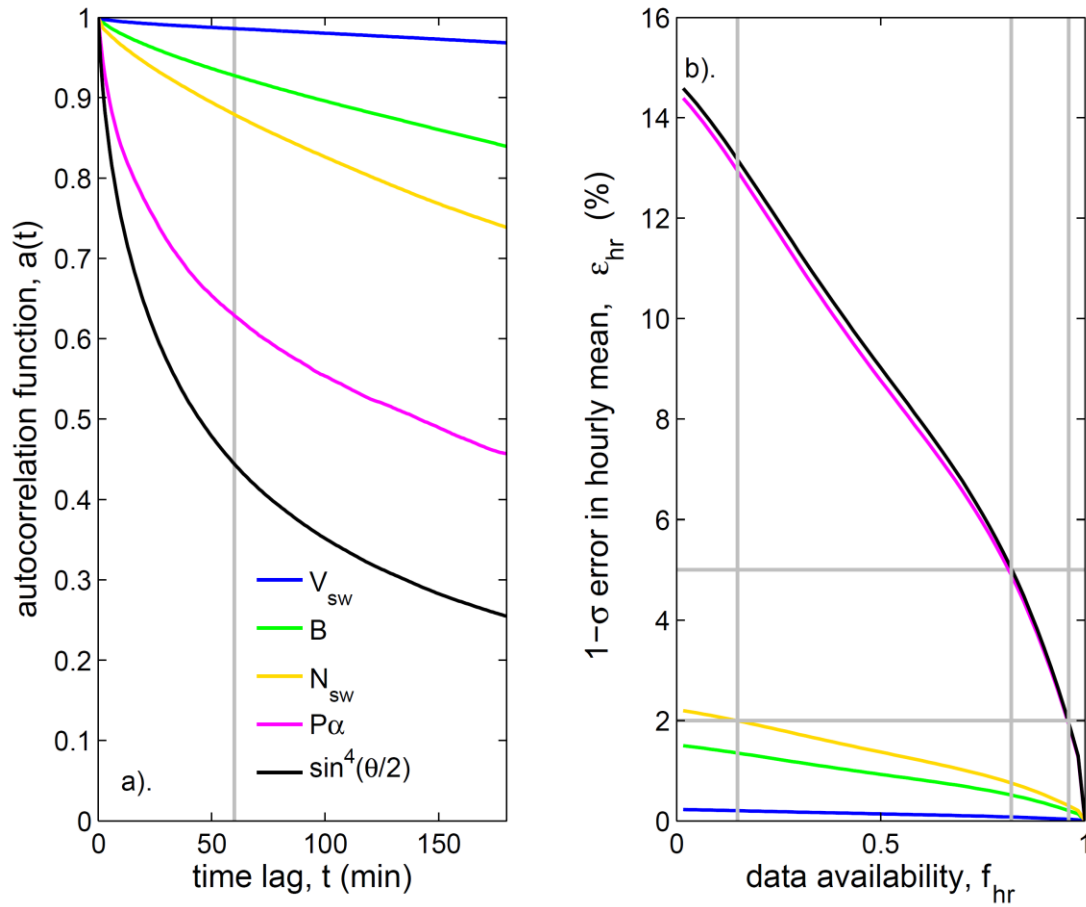
1151 **Table 2.** Correlations between $P\alpha$ and the ap and am indices

averaging timescale, τ	correlation coefficient, r		$2\text{-}\sigma$ range in r		optimum coupling exponent, α_p	
	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

1152

1153 **Figures**

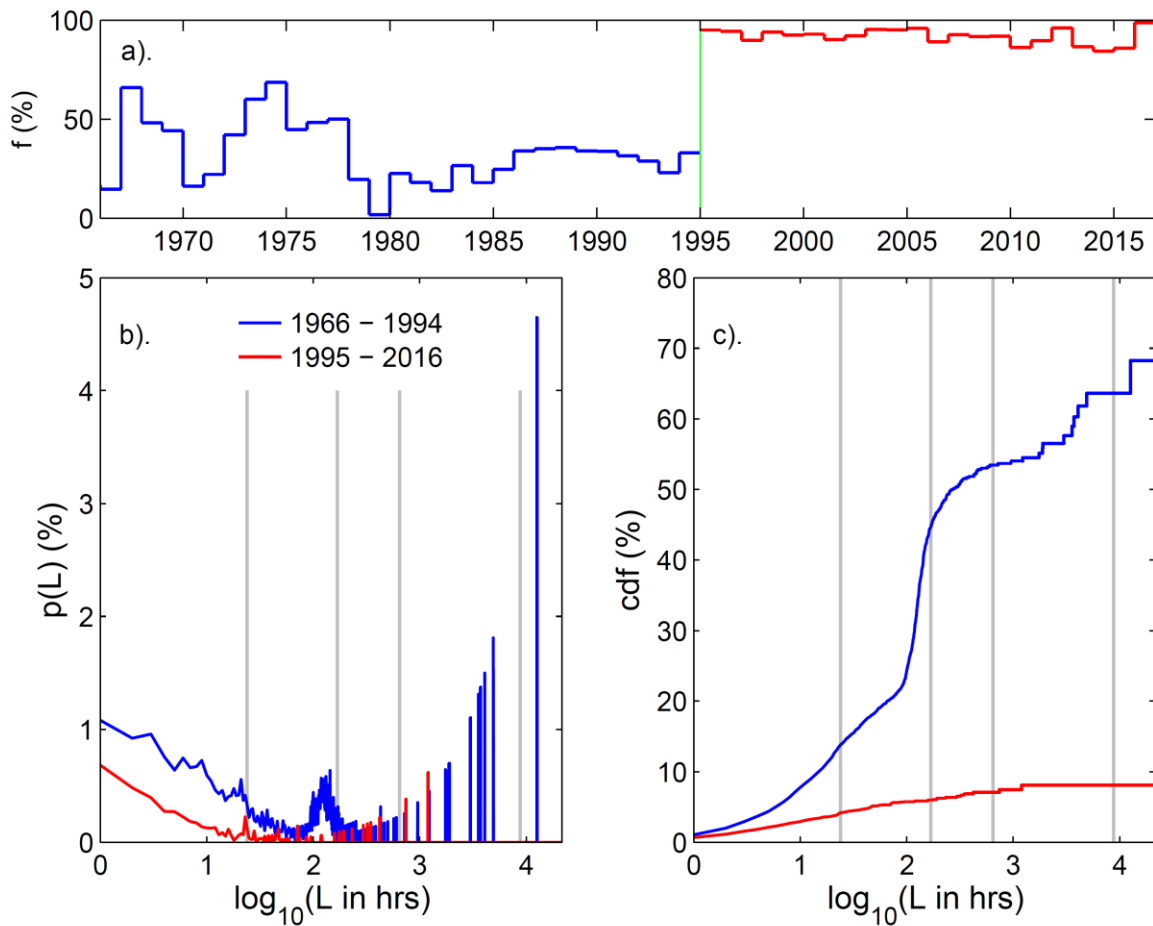
1154



1155 **Figure 1.** (a) Autocorrelation functions, $a(t)$ as a function of lag time, t , for 1-minute samples
 1156 of: (blue) solar wind speed, V_{sw} ; (orange) solar wind number density, N_{sw} ; (green) interplanetary
 1157 magnetic field (IMF) , B ; (mauve) power input to the magnetosphere (for $\alpha = 0.5$), P_α ; and
 1158 (black) IMF orientation factor, $\sin^4(\theta/2)$. The vertical gray line marks $t = 1$ hour. (b) 1- σ
 1159 percentage errors in hourly means, ϵ_{hr} , as a function of data availability within the hour, f_{hr} , for
 1160 the same parameters (modulated by introducing synthetic data gaps, see text), shown using the
 1161 same color scheme. The horizontal gray lines in (b) are uncertainties ϵ_{hr} of 5% and 2% which set
 1162 threshold requirements for IMF orientation data availability of $f_{hr} \geq 0.82$ and $f_{hr} \geq 0.96$,
 1163 respectively. The $\epsilon_{hr} \leq 2\%$ condition sets an additional availability requirement on N_{sw} data of f_{hr}

1164 ≥ 0.15 . These f_{hr} thresholds are shown by the vertical gray lines. Plots are constructed using 1-
1165 minute Omni samples for 1996-2016.

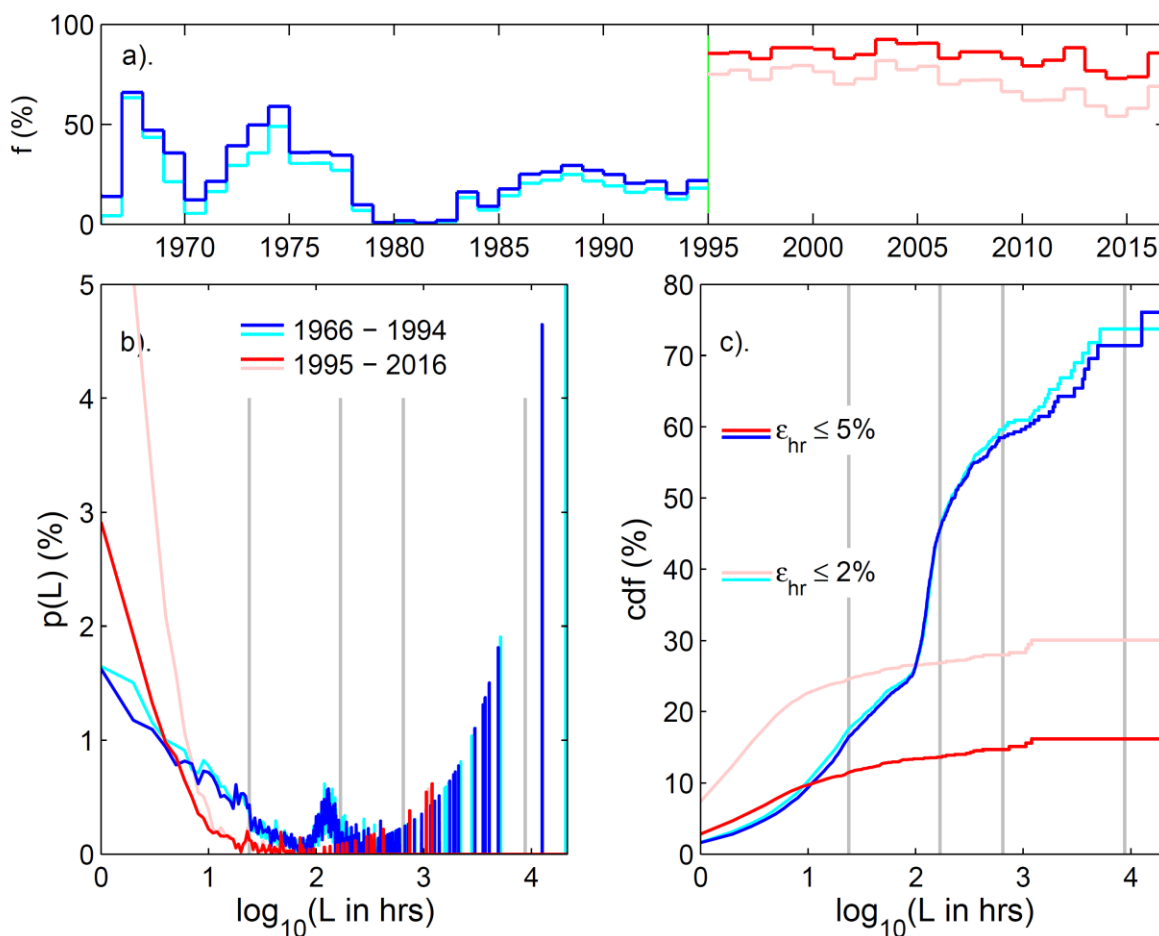
1166



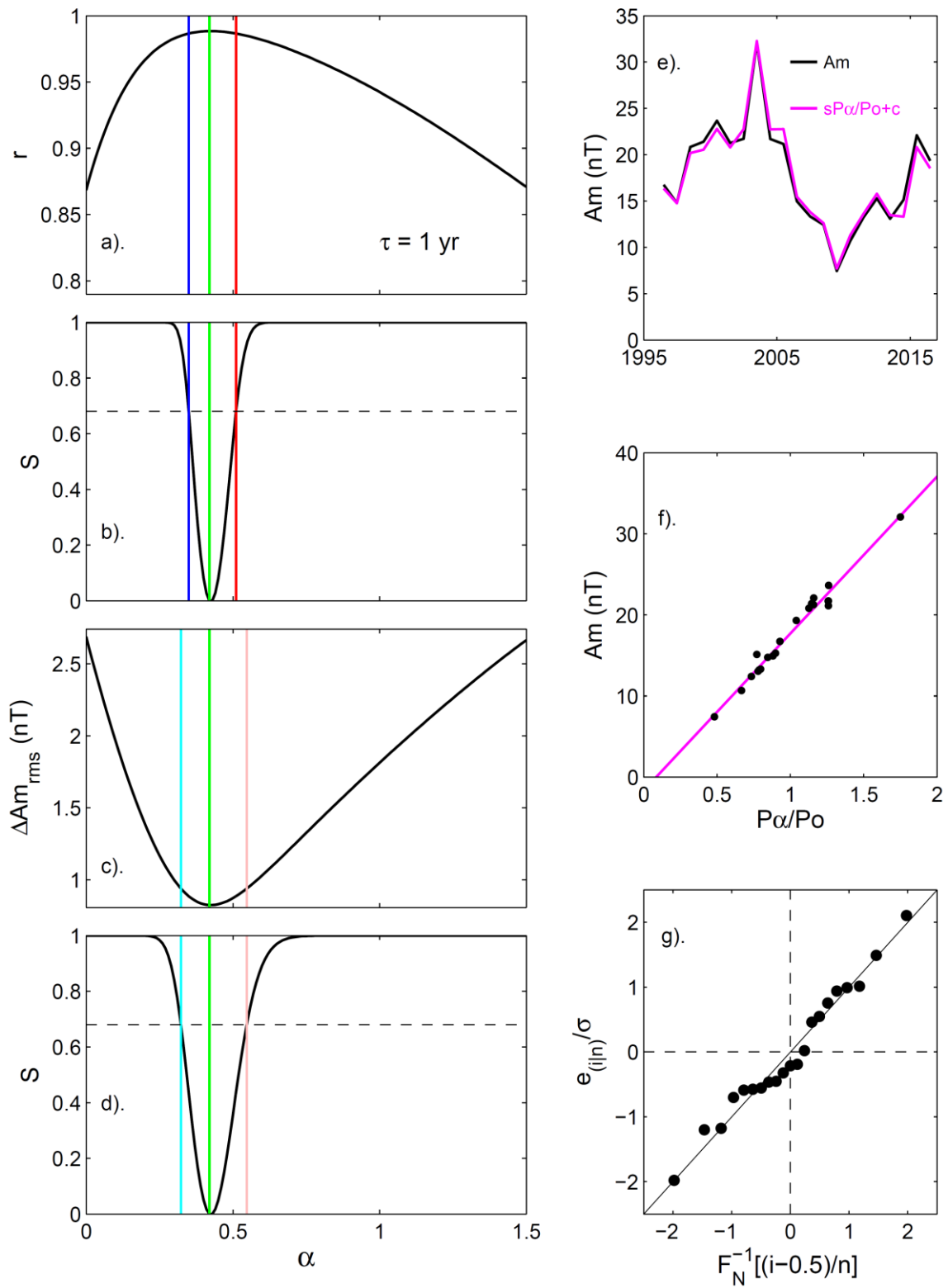
1167

1168 **Figure 2.** Analysis of data gap occurrence in the P_α interplanetary data. Data in blue are for
 1169 1966-1994 (inclusive) and in red are for 1995-2016 (inclusive). Data gaps are here defined by the
 1170 absence of any sample in one-hour intervals, as used in compiling Omni2 data hourly samples.
 1171 (a) the availability f of 1-hour means in calendar years. (b) The probability of being in a data gap
 1172 of duration L , $p(L)$, shown as a function $\log_{10}(L)$ for L between 1 hour and 1 year, for before and
 1173 after 1 January 1995 (blue and red lines, respectively). (c) The cumulative probability
 1174 distribution as a function of $\log_{10}(L)$. The total probability for being within a data gap for 1966-
 1175 1994 is 67% (so the average data availability is $\langle f \rangle = 33\%$) and for 1995-2016 it is 8% ($\langle f \rangle =$
 1176 92%). Gray vertical lines in (b) and (c) are at 1 day, 7 days, 27 days and 1 year.

1177

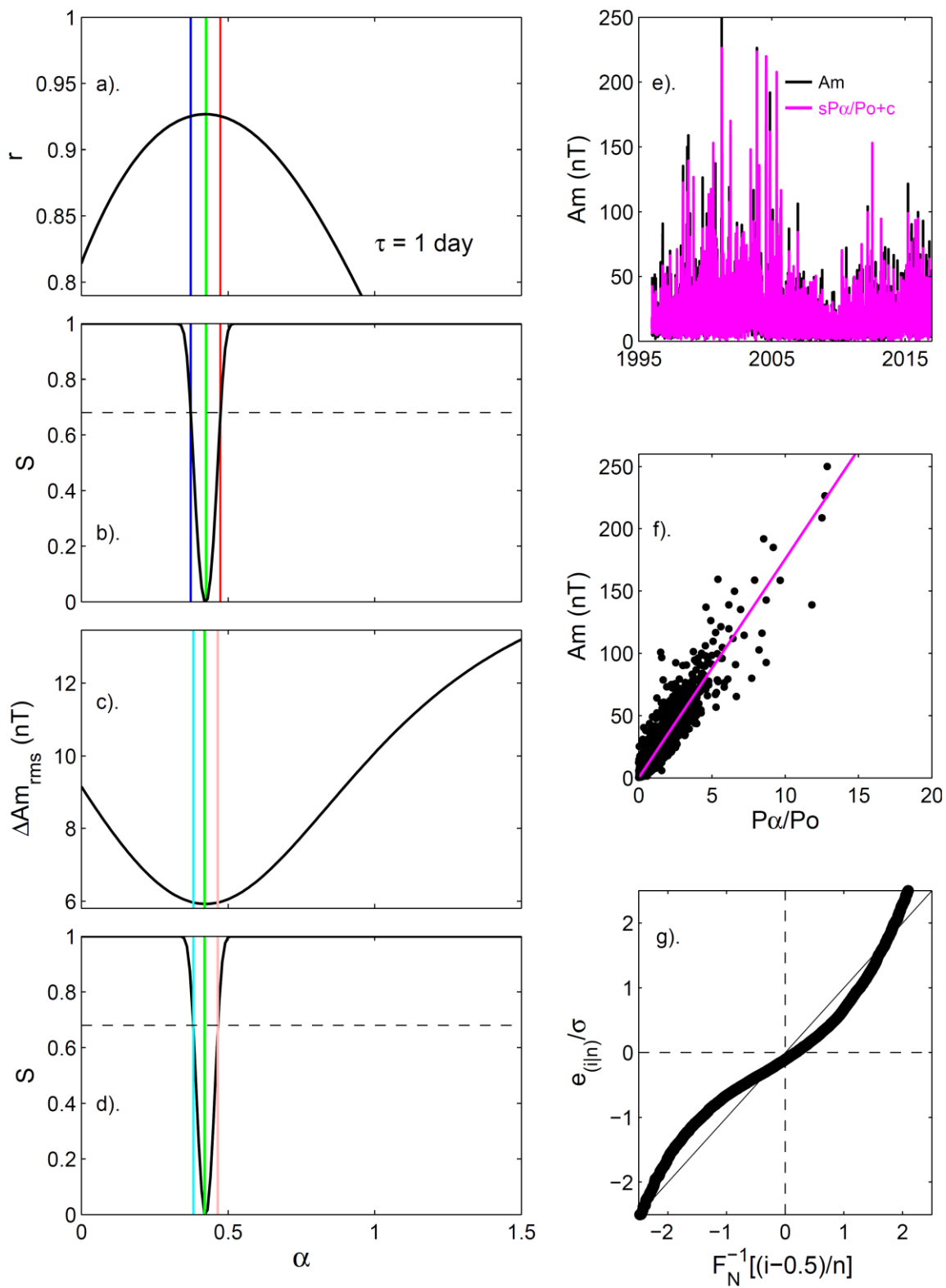


1178 **Figure 3.** Same as Figure 2 for more stringent criteria as to what constitutes a valid hourly
 1179 sample. Specifically, for the red and blue lines we require that solar wind data availability gives
 1180 an average uncertainty in the hourly coupling function P_α values of $\epsilon_{hr} \leq 5\%$ and for the pink
 1181 and cyan lines we require $\epsilon_{hr} \leq 2\%$. Blue and cyan lines are for before 1995: $\epsilon_{hr} \leq 5\%$ gives the
 1182 blue lines and 23.9% availability of hourly sample; lowering the uncertainty limit to $\epsilon_{hr} \leq 2\%$
 1183 lowers this availability to 18.9% (cyan lines). Red and pink lines are for 1995 and after: $\epsilon_{hr} \leq 5\%$
 1184 gives the red lines with 83.8% availability of hourly samples; lowering the uncertainty limit to ϵ_{hr}
 1185 $\leq 2\%$ lowers this availability to 70.0% (pink lines).



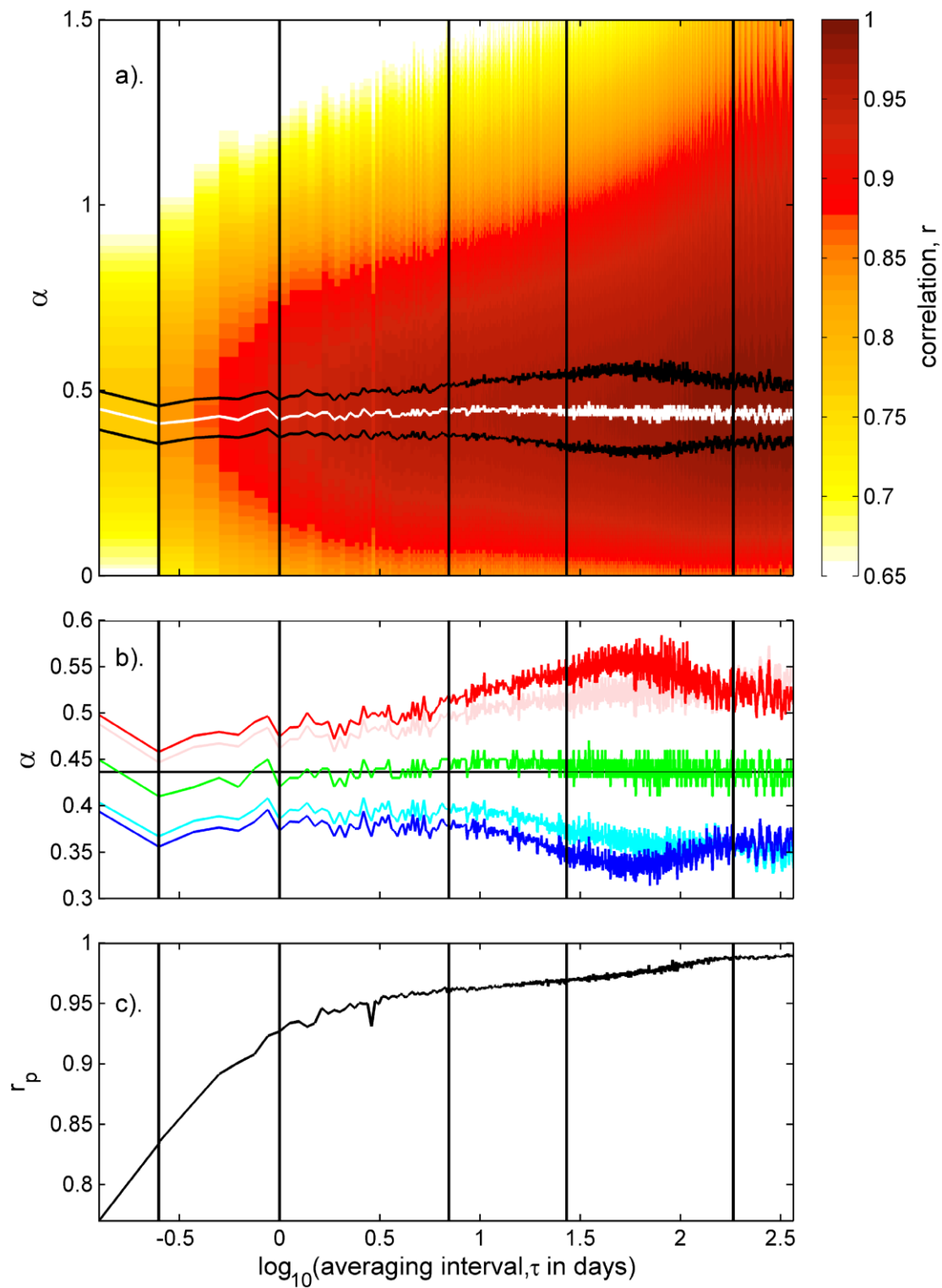
1186 **Figure 4.** Analysis of correlation between power input to the magnetosphere, P_α , and the am

1187 geomagnetic index for averaging timescale $\tau = 1$ yr over the interval 1996-2016 (inclusive).
 1188 Panels (a)-(d) show, as a function of the coupling exponent α , (a) the correlation coefficient r
 1189 between $\langle Am \rangle_{1\text{yr}}$ and $\langle P_\alpha \rangle_{1\text{yr}}$; (b) the significance, S , of the difference between the peak value
 1190 of r , r_p (marked by the vertical green line at $\alpha = \alpha_p = 0.42$) and the r at any other α , $r(\alpha)$,
 1191 computed using the Meng-Z test (see text for details): the vertical red and blue lines are where S
 1192 $= 0.68$ and so mark where r is significantly lower than r_p at the $1-\sigma$ level; (c) the r.m.s.
 1193 difference, ΔAm_{rms} between $\langle Am \rangle_{1\text{yr}}$ and the best fit (for $\alpha = \alpha_p$) of $\langle P_\alpha / P_o \rangle_{1\text{yr}}$, $Am_{\text{fit}} =$
 1194 $s \cdot \langle P_\alpha / P_o \rangle_{1\text{yr}} + c$, where P_o is the average of P_α for all the data and the linear best-fit regression
 1195 coefficients $s = 19.37 \pm 0.42$ and $c = -1.64 \pm 0.31$. (d) shows the significance S of the difference
 1196 between ΔAm_{rms} for $\alpha = \alpha_p$ and that at other α , computed using a 2-sample variance F test (see
 1197 text for details): pink and cyan vertical lines are at $S = 0.68$. (e) shows the time series of
 1198 $\langle Am \rangle_{1\text{yr}}$ (in black) and the best-fit Am_{fit} (in mauve);. (f) Scatter plot of $\langle Am \rangle_{1\text{yr}}$ against $\langle P_\alpha / P_o$
 1199 $\rangle_{1\text{yr}}$ for $\alpha = \alpha_p$ with the best-fit linear regression line shown in mauve. (g) is a quantile-quantile
 1200 (q-q) plot of the best-fit residuals of Am_{fit} and $\langle Am \rangle$ used to test for the normality of their
 1201 distribution: $e_{(i|n)}/\sigma$ are the ordered standardized residuals and $F_N^{-1}[(i-0.5)/n]$ are the quantiles of
 1202 the standard normal distribution.



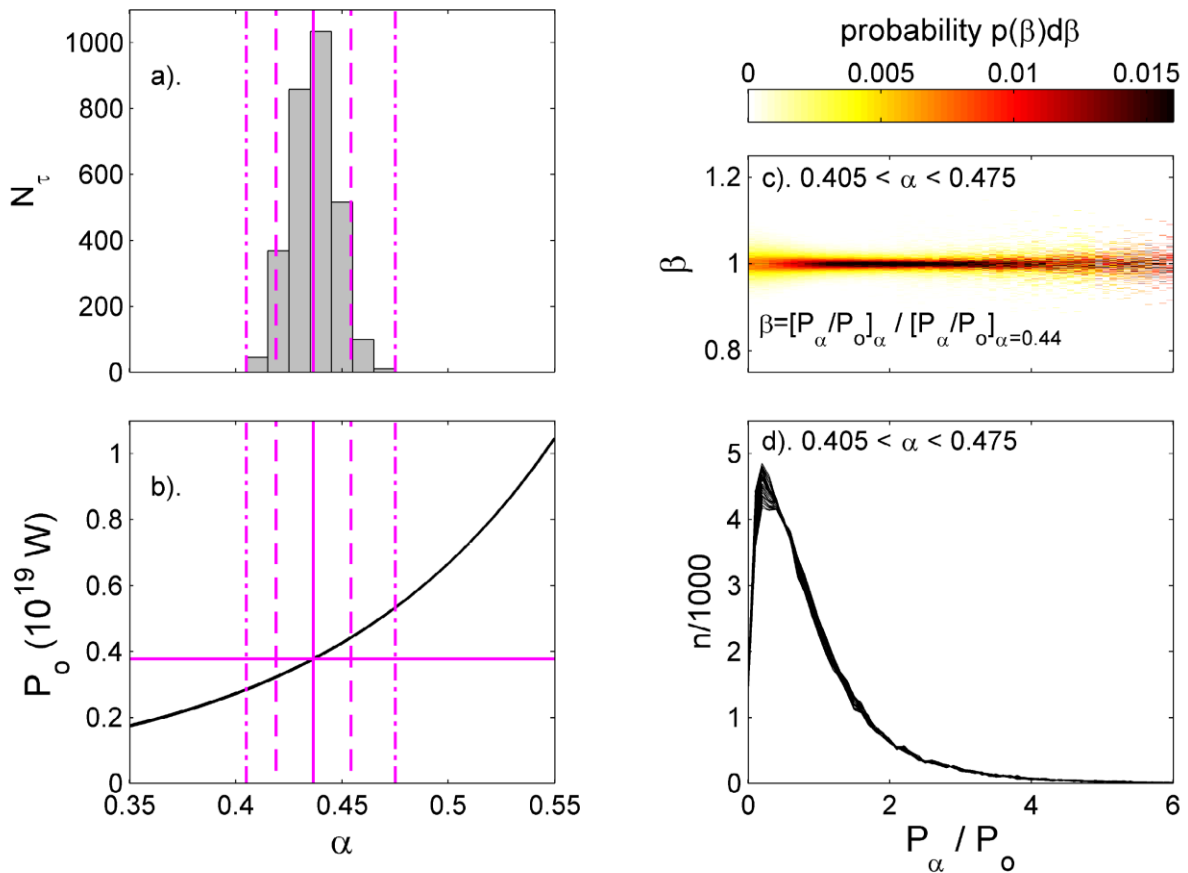
1203 **Figure 5.** Same as Figure 4 for an averaging timescale $\tau = 1$ day. In this case, $\alpha_p = 0.42$ and the

1204 best-fit linear regression coefficients are $s = 17.58 \pm 0.89$ and $c = 0.02 \pm 0.40$.



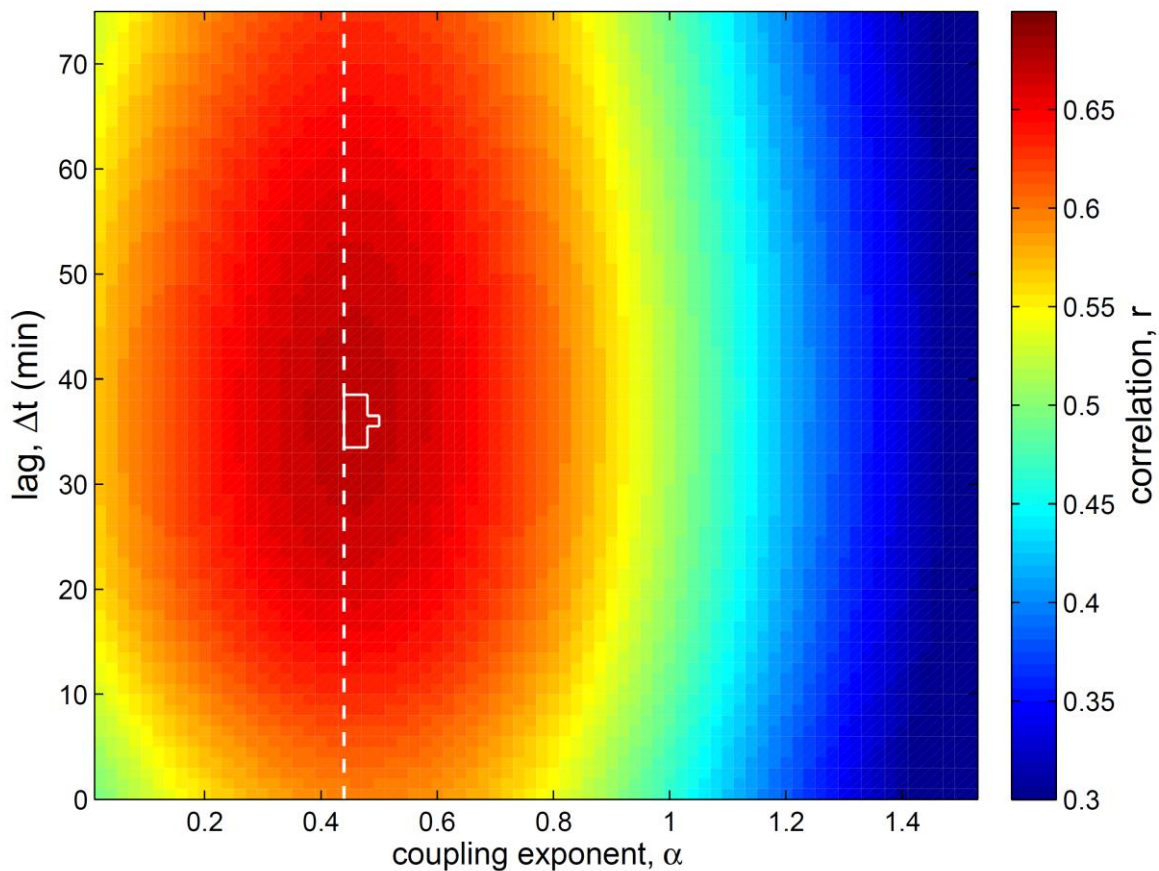
1205 **Figure 6.** (a). Correlation coefficients between $\langle am \rangle_\tau$ and $\langle P_\alpha \rangle_\tau$ for the interval 1996-2016
 1206 (inclusive), r , color-coded as a function the logarithm of averaging timescale, $\log_{10}(\tau)$, and the

1207 coupling exponent, α . The middle white line gives the peak of each vertical slice, i.e. the
 1208 optimum α , α_p (for which r has a maximum value r_p) (as shown by the vertical green lines in
 1209 Figures 4a and 5a for $\tau = 1$ yr and $\tau = 1$ day, respectively). The upper and lower black lines give
 1210 the $\pm 1\sigma$ uncertainty of α_p from the Meng-Z test (as shown by the vertical red and blue lines in
 1211 Figure 4a for τ of 1 year and in Figure 5a for τ of 1 day, respectively). The left-hand edge of the
 1212 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,
 1213 7 days, 27 days and 0.5 year. (b) Variation of best α estimates (α_p) and uncertainties. Line
 1214 colors are as are used in Figures 4 and 5. The green line give the results for α_p from both the
 1215 peak correlation ($r = r_p$) and the minimum r.m.s. fit deviation ($\Delta A_{m_{rms}}$), which are always
 1216 exactly the same. The blue and red lines are the $\pm 1\text{-}\sigma$ uncertainties in α_p computed from the $r(\alpha)$
 1217 variation at each τ using the Meng-Z test. The cyan and pink are the 1-sigma uncertainties in α_p
 1218 computed from the fit residuals $\Delta A_{m_{rms}}(\alpha)$ variation using a two-sample variance F test. (c) The
 1219 peak correlation r_p as a function of $\log_{10}(\tau)$.



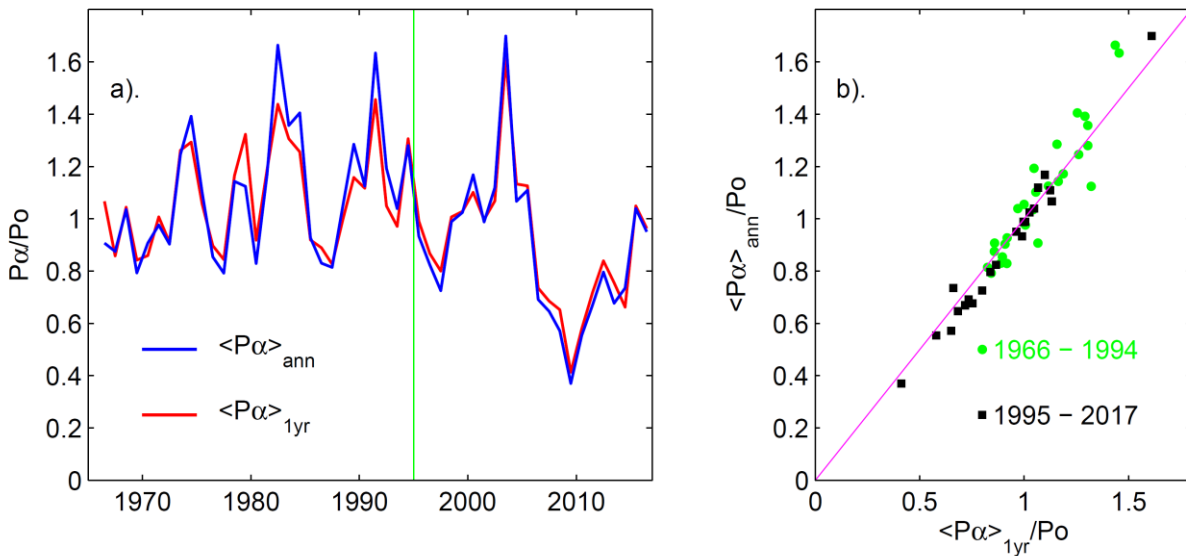
1220

1221 **Figure 7.** Analysis of the effects of the uncertainty in the best-fit α . (a). The distribution of
 1222 best fit α values from the analysis of 2922 τ -values surveyed in figure 6. The vertical mauve
 1223 dashed lines are the 2- σ points of the distribution and the dot-dash lines mark the extreme values.
 1224 (b) The variation of the mean value of power input into the magnetosphere over the interval
 1225 1996-2017, P_o , as a function of α : the average α is 0.44 which gives $P_o = 0.38 \times 10^{19}$ W and the
 1226 uncertainty range of α of 0.42- 0.45 at the 2- σ level (between the dashed lines) gives P_o between
 1227 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
 1228 and 0.53×10^{19} W, respectively. (c) Analysis of the effect of α on 3-hourly P_α / P_o values: the
 1229 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
 1230 0.41-0.47 (computed in steps of 0.01). (d) The distributions of P_α / P_o values for this full range of
 1231 α .



1232

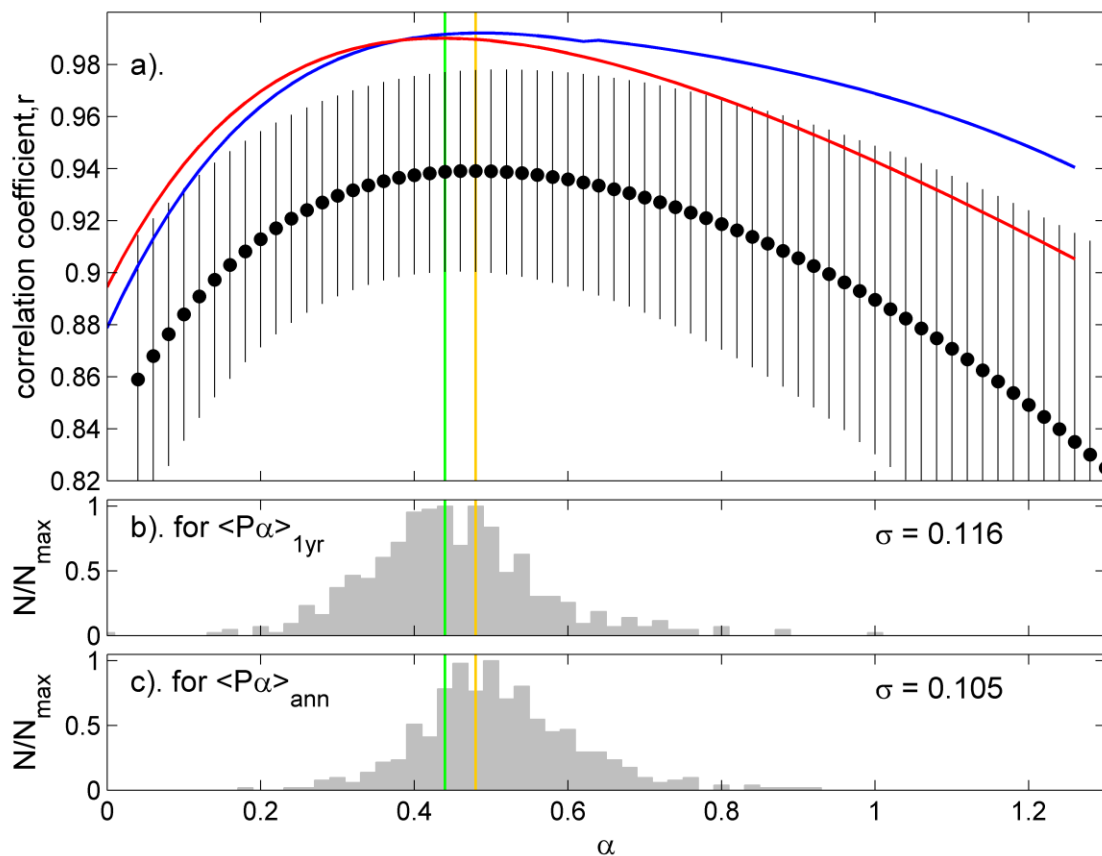
1233 **Figure 8.** The correlation r of 1-minute P_α values with 1-minute values of the SuperMAG
 1234 auroral electrojet index, SME, color-coded as a function of coupling exponent, α (horizontal
 1235 axis) and time lag, Δt (vertical axis). Positive Δt is when P_α is lagged, i.e. the SME variation
 1236 follows P_α . Inside the white contour the correlation is not significantly different from the peak
 1237 value at the 1- σ level. The vertical white dashed line is at $\alpha = 0.44$.



1238

1239 **Figure 9.** Comparison of “combine-then-average” and “average-then-combine” annual values of
 1240 P_α (respectively, $\langle P_\alpha \rangle_{1yr}$ and $\langle P_\alpha \rangle_{ann}$). Note that the optimum coupling exponent is used which
 1241 is $\alpha_p = 0.44$ for $\langle P_\alpha \rangle_{1yr}$ and $\alpha_p = 0.48$ for $\langle P_\alpha \rangle_{ann}$. (a) The time series of values (normalized to
 1242 P_o , their mean value over the whole 1966-2016 data set: $P_o = 5.679 \times 10^{18}$ W for $\langle P_o \rangle_{1yr}$ and $P_o' = 5.719 \times 10^{18}$ W for $\langle P_o \rangle_{ann}$). The vertical green line is the start of almost continuous
 1243 interplanetary data from ACE, WIND and DSCOVR. (b) Scatter plot of $\langle P_\alpha \rangle_{1yr}/P_o$ against
 1244 $\langle P_\alpha \rangle_{ann}/P_o'$. Points for 1995 and after are shown as black squares and points for before then are
 1245 as green circles. The correlation coefficient for the 1995-2016 data is 0.994 and for the 1966-
 1246 1994 data is 0.929. The mauve line is the line of perfect agreement.

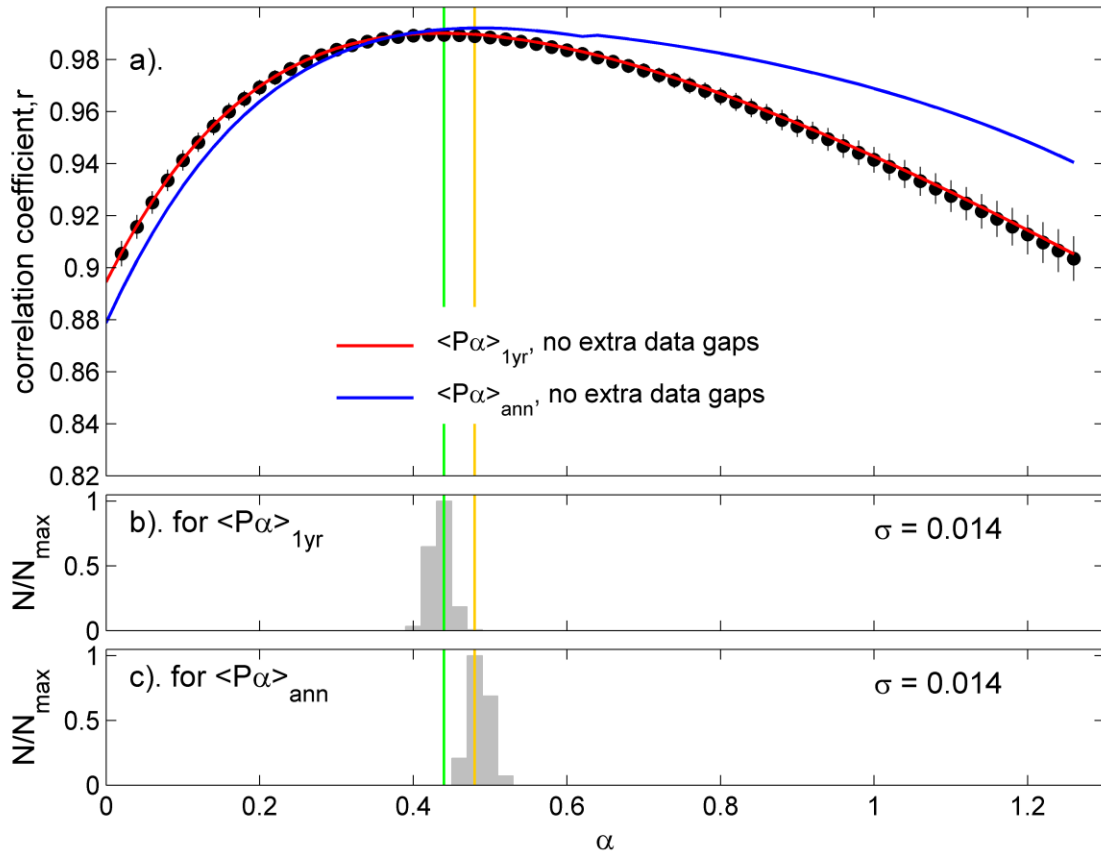
1248



1249

1250 **Figure 10.** Analysis of the effect of adding the series of data gaps that exist in the P_α data series
 1251 during 1966-1996 on the P_α -am correlogram (as a function of the coupling exponent, α , used)
 1252 derived for annual means (i.e., $\tau = 1\text{yr}$) from 1996-2016. In this case the effects of these gaps **has**
 1253 been neglected (i.e., Method B, the most commonly employed). (a) The red and blue lines are
 1254 the correlograms for all the 1996-2016 data for, respectively, $\langle P_\alpha \rangle_{1\text{yr}}$ (the combine-then-average
 1255 estimate) and $\langle P_\alpha \rangle_{\text{ann}}$ (the average-then-combine estimate). The black points with error bars are
 1256 the mean and standard deviations of the distribution of correlation coefficients, r , for $\langle P_\alpha \rangle_{1\text{yr}}$ at
 1257 a given α , obtained from $N_s = 500$ different masks applied to the 1996-2016 data to introduce
 1258 extra, synthetic data gaps. These masks produce the time series of gaps as was observed ($30-D$)
 1259 years earlier where D is varied between 0 and 10 years in steps of 175 hours. (So, for example,
 1260 for $D = 0$, the mask produces data gaps in the same positions of the data series as for the 1966-

1261 1986 data and for $D = 10$ years they are in the same positions as for 1976-1996). For the period
1262 from which the data gap masks are drawn (1966-1996) data availability is 23.9%. (b) The
1263 distribution of the optimum α values (which give the peak r between $\langle P_\alpha \rangle_{1\text{yr}}$ and $\langle am \rangle_{1\text{yr}}$ for
1264 the $N_S = 500$ different data gap masks applied to both series. (c) The corresponding distribution
1265 of α values giving the peak r between $\langle P_\alpha \rangle_{\text{ann}}$ and $\langle am \rangle_{1\text{yr}}$ for the 500 different data gap
1266 masks applied to both series. The green line in all three panels is the optimum α of 0.44 derived
1267 for $\langle P_\alpha \rangle_{1\text{yr}}$ and the orange line is the optimum α of 0.48 derived for $\langle P_\alpha \rangle_{1\text{ann}}$.
1268



1269

1270 **Figure 11.** The same as Figure 10, but allowance for data gaps is made using the piecewise
 1271 removal procedure of *Finch and Lockwood* [2007] (Method B) applied to the *am* data at times of
 1272 the synthesized data gaps in P_α . In addition, in this case, the 500 synthetic data gap masks that
 1273 have been applied to the P_α data are drawn from sequences of actual data gaps during 1995-2016,
 1274 when data availability is 83.8%. Specifically, the 500 masks constructed for the data-gap series
 1275 observed ($1-D$) years earlier (when the data are almost continuous) for D varied between 0 and 1
 1276 year in steps of 17.5 hours. (So for example, for $D = 0$, the mask produces extra, synthetic data
 1277 gaps in the same positions of the data series as for the 1995-2015 data and for $D = 1$ year the
 1278 mask reproduces the actual data gaps 1996-2016).