

*Does adding solar wind Poynting flux
improve the optimum solar wind -
magnetosphere coupling function?*

Article

Accepted Version

Lockwood, M. ORCID: <https://orcid.org/0000-0002-7397-2172>
(2019) Does adding solar wind Poynting flux improve the
optimum solar wind - magnetosphere coupling function?
Journal of Geophysical Research: Space Physics, 124 (7). pp.
5498-5515. ISSN 2169-9402 doi: 10.1029/2019JA026639
Available at <https://centaur.reading.ac.uk/84889/>

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/10.1029/2019JA026639>

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

Does adding solar wind Poynting flux improve the optimum solar wind - magnetosphere coupling function?

Mike Lockwood* (1)

(1) Department of Meteorology, University of Reading UK

Abstract. We study the contribution of the solar wind Poynting flux \vec{S}_{sw} , to the total power input into the magnetosphere. The dominant power delivered by the solar wind is the kinetic energy flux of the particles which is larger than S_{sw} by a factor of order M_A^2 , where M_A is the Alfvén Mach number. The currents \vec{j} flowing in the bow shock and magnetosheath and the electric field \vec{E} of the solar wind give regions where $\vec{j} \cdot \vec{E} < 0$, which are sources of Poynting flux, generated from the kinetic energy flux. For southward IMF, \vec{E} is duskward and the currents in the high-latitude tail magnetopause are also sources of Poynting flux. We show transfer of kinetic energy into the magnetosphere is less efficient than direct entry of \vec{S}_{sw} by a factor M_A . Because M_A is typically of order 10, this means that although the power density in the solar wind due to SSW is typically only 1%, it is responsible for of order 10% of the energy input to the magnetosphere. To investigate the effect of this, we add a term to the solar wind-magnetosphere energy coupling function that allows for \vec{S}_{sw} which increases the correlation with the geomagnetic am index for 1995-2017 (inclusive) from 0.908 to 0.924 for 1-day averages and from 0.978 to 0.979 for annual means. The increase for means on daily or smaller timescales is a small improvement but is significant (at over the 3- σ level), whereas the improvement for annual or Carrington-rotation means is not significant.

1. Introduction

The basic physics of energy flow into the magnetosphere was elegantly summarized using Poynting's theorem by *Cowley* (1991). He considered only the steady-state case when the near-Earth Interplanetary Magnetic Field (IMF) points southward. This was generalized to cover the substorm phases and northward IMF conditions (which are all inherently non-steady state cases) by *Lockwood* (2004). Global magnetohydrodynamic (MHD) numerical

30 simulations give a unique way of studying the details of this energy flow into the
31 magnetosphere, the storage and deposition in the magnetosphere, and its return from the
32 magnetosphere to the interplanetary medium (*Palmroth et al., 2003; Ebihara et al. (2019)*).
33 These studies confirm the fundamental physical expectation that the energy density in the
34 solar wind is dominated by the kinetic energy flux of the bulk flow of the particles, it being
35 roughly 2 orders of magnitude larger than the magnitude of the Poynting flux vector $|\vec{S}_{SW}| =$
36 S_{SW} , (and larger than both the magnetic energy flux and the thermal energy flux by about the
37 same factor). The simulations also confirm the expectation that between about 2% and 7% of
38 the solar wind kinetic energy that is incident on the effective cross-sectional area that the
39 magnetosphere presents to the solar wind enters the magnetosphere (*Koskinen and*
40 *Tanskanen, 2002*).

41 An interesting point that emerges from the simulations by *Ebihara et al. (2019)* is that,
42 although the fraction of total energy flux in the solar wind that is in the form of Poynting flux
43 is very small, the fraction of power that is delivered to the magnetosphere that originates from
44 that solar wind Poynting flux may not be as small because of the relative inefficiency with
45 which kinetic energy of the solar wind is converted into Poynting flux by currents flowing in
46 the bow shock, magnetosheath and magnetopause (*Cowley, 1991*).

47 *Vasyliunas et al. (1982)* used dimensional analysis of power input into the magnetosphere to
48 derive a coupling function between the solar wind and the magnetosphere. This coupling
49 function has just one free fit parameter (the coupling exponent α) which means it is much
50 less prone to errors associated with statistical “overfitting”. *Lockwood et al. (2019a)* have
51 shown that data gaps in the interplanetary data series cause considerable noise in solar wind –
52 magnetosphere correlation studies and this noise can cause major overfitting problems
53 because fits with too many free parameters are fitting the noise and therefore are not robust
54 when considering general datasets. In short, overfitting is damaging predictive power. The
55 theory by *Vasyliunas et al. (1982)* introduces the one free fit parameter, α , through a
56 dimensionless term $M_A^{-2\alpha}$ which means that it appears in the exponents for terms in the
57 resulting expression for magnetospheric power input in solar wind velocity (V_{sw}), mean ion
58 mass (m_{sw}), number density (N_{sw}), and the IMF field strength (B). The formulation by
59 *Vasyliunas et al. (1982)* was shown to be optimum over a very large range of timescales by
60 *Finch and Lockwood (2004)*, although for long averaging timescales (approaching 1 year) the
61 simpler form $V_{sw}^2 B$ (with no IMF orientation factor) performs equally well. This formulation

62 has been used a great many times in diverse areas, for example: to investigate which coupling
 63 function best predicts geomagnetic storms (*Gonzalez et al.*, 1989); to compute long term
 64 change in open solar flux (*Lockwood et al.*, 1999); to understand transpolar voltage saturation
 65 (*Siscoe et al.*, 2002) and many others. Recently, the realisation of the problems associated
 66 with overfitting has made the *Vasyliunas et al.* formulation, with its single free fit parameter,
 67 important and it has been used to study remote sensing of the geoeffectiveness of CMEs
 68 (*Owens et al.*, 2018) and to reconstruct the numbers of storms and substorms back to (and
 69 including) the Maunder minimum (*Lockwood et al.*, 2017, 2018a).

70 The *Vasyliunas et al.* (1982) theory, and so the applications that employed it, is based on the
 71 approximation that all energy fluxes in the solar wind can be neglected except the kinetic
 72 energy flux. This has been very successful: for example, by careful analysis to minimise the
 73 effect of data gaps, *Lockwood et al.* (2019a) have shown that the correlation with
 74 geomagnetic activity is 0.990 ± 0.007 for annual means, 0.897 ± 0.004 for daily means, 0.79
 75 ± 0.03 , for 3-hourly means, and 0.7046 ± 0.0004 for one-minute means (the uncertainties
 76 being at the 2σ level and the lower correlation for 1-minute data is largely due to the
 77 variability around the average of the substorm growth-phase lag). Given that the study of
 78 *Ebihara et al.* (2019) shows that the fraction of the solar wind power entering the
 79 magnetosphere due to S_{sw} is larger than the fraction of the energy flux that it carries in the
 80 solar wind, it is interesting to see if solar wind coupling functions based on energy input into
 81 the magnetosphere can be improved by adding a term to allow for solar wind Poynting flux,
 82 S_{sw} . This is investigated.

83 2. Total Power into the magnetosphere

84 Poynting flux in a plasma (for which $\mu_r = 1$) is given by

$$85 \quad \vec{S} = \vec{E} \times \vec{H} = \frac{\vec{E} \times \vec{B}}{\mu_0} . \quad (1)$$

86 Assuming that the solar wind flows radially to Earth (i.e. the solar wind velocity \vec{V}_{sw} is in the
 87 $-X$ direction of the Geocentric Solar Magnetospheric (GSM) reference frame) and using
 88 ideal Magnetohydrodynamics (MHD) so that $\vec{E}_{sw} = -\vec{V}_{sw} \times \vec{B}$, where \vec{B} is the interplanetary
 89 magnetic field, IMF), the Earthward-directed Poynting flux is

$$90 \quad S_{sw} = \frac{V_{sw} B_{\perp}^2}{\mu_0} = \frac{V_{sw} B^2}{\mu_0} \cos^2(\varphi) \quad (2)$$

91 where B_{\perp} is the component of \vec{B} that is transverse to the flow and so φ is the angle between
 92 \vec{V}_{sw} and \vec{B} . Hence the ratio of the kinetic energy flux of particles to the Poynting flux in the
 93 solar wind is:

$$94 \quad \frac{F_{KE}}{S_{sw}} = \frac{\frac{1}{2} m_{sw} N_{sw} V_{sw}^3}{\frac{1}{\mu_0} V_{sw} B^2 \cos^2(\varphi)} = \frac{M_A^2}{2 \cos^2(\varphi)} \quad (3)$$

95 where m_{sw} , N_{sw} and M_A are the mean ion mass, number density and Alfvén Mach number of
 96 the solar wind, respectively. Because M_A is typically 10 and $2 \cos^2(\varphi) \approx 1$, this means that
 97 the Poynting flux in the solar wind is of order 1% of the kinetic energy flux.

98 Poynting's theorem for a plasma (in which there is no displacement current and no permanent
 99 magnetism) is

$$100 \quad - \int_{A_1} \vec{S} \cdot d\vec{A} = dW_B/dt + \int_{\tau_1} \vec{E} \cdot \vec{J} d\tau \quad (4)$$

101 where: W_B is the energy density stored in the magnetic field, $W_B = B^2/(2\mu_0)$; A_1 is a
 102 surface area surrounding a unit volume τ_1 ; \vec{E} , \vec{J} , and \vec{S} are the electric field, current density
 103 and Poynting flux vectors, $d\vec{a}$ is an element of the surface area A_1 , $d\tau$ is an element of
 104 volume and τ_1 is the volume inside the surface area A_1 . This means that the negative of the
 105 divergence of the Poynting flux equals the sum of the rates at which energy is given to the
 106 magnetic field and to the particles (it can be shown that the last term in (4) is the sum of
 107 ohmic heating and the work done by/against the so-called $\vec{J} \times \vec{B}$ force). Hence this is a
 108 statement of conservation of energy. Regions with current that is aligned with the electric
 109 field, and/or with an increasing the magnetic field, are sinks of Poynting flux and conversely
 110 regions with a current anti-parallel to the electric field and/or a falling magnetic field, are
 111 sources of Poynting flux. In the steady-state discussed by *Cowley* (1991), $dW_B/dt = 0$ and
 112 so sinks of Poynting flux are regions where particles are accelerated or heated ($\vec{E} \cdot \vec{J} > 0$) and
 113 sources of Poynting flux are where particles are slowed or cooled ($\vec{E} \cdot \vec{J} < 0$).

114 **2.1 Poynting Flux for Southward IMF**

115 Figure 1a illustrates the steady-state case for when the IMF is pointing southward in the GSM
 116 frame (in the $-Z_{GSM}$ direction) so the motional electric field of the solar wind in the Earth's
 117 frame, E_{sw} , points from dawn to dusk in the $+Y_{GSM}$ direction (after *Cowley*, 1991). In this

118 steady-state case, the magnetic field is constant everywhere and so, by Faraday's law, the
119 electric field is curl-free which means it is the same at all points in the noon-midnight plane
120 shown in Figure 1a.

121 Before entering the magnetosphere, the kinetic energy flux is converted into Poynting flux by
122 the currents \vec{J} that flow in the bow shock, magnetosheath and tail magnetopause in regions
123 where $\vec{J} \cdot \vec{E}_{sw} < 0$, as shown in in Figure 1a. This is added to the pre-existing solar wind
124 Poynting flux. Because \vec{S} is perpendicular to the magnetic field, the draped IMF in the
125 magnetosheath deflects Poynting flux towards the magnetosphere. This occurs irrespective
126 of the IMF orientation. The major extraction of energy by the magnetosphere from the
127 magnetosheath takes place during intervals of southward IMF along the north and south
128 flanks of the tail lobes where the magnetopause currents are from dusk to dawn and so anti-
129 parallel to the electric field (i.e., $\vec{J} \cdot \vec{E} < 0$). Note that these magnetopause currents are
130 orthogonal to the electric field at low magnetospheric latitudes (close to the equatorial plane)
131 and so $\vec{J} \cdot \vec{E} = 0$ and no energy is extracted there.

132 Figures 1b and 1c show how this is modified by non-steady conditions during the substorm
133 cycle. (after *Lockwood, 2004*). During the growth phase, energy is stored in the increasing
134 field in the tail lobes which are therefore sinks of Poynting flux; in the expansion phase this
135 stored energy is released and deposited in the plasma sheet, ring current and ionosphere,
136 whilst some is propagated down the far tail and returned to the interplanetary medium. Note
137 in these cases, the changing magnetic field in the tail lobes means that the electric field is not
138 curl-free and this induction effect decouples the electric field at high latitudes in the tail lobe
139 from that in the cross-tail current sheet: in the growth/expansion phase the electric field at the
140 cross-tail current sheet is lower/higher than that at the high-latitude magnetopause,
141 respectively, and hence the sink of Poynting flux in the cross-tail current sheet is
142 smaller/greater than the source at the high-latitude magnetopause for growth/expansion
143 phases as tail lobe energy is stored/released.

144 **2.2 Poynting Flux for Northward IMF**

145 Figure 1d shows the situation for persistent northward IMF (\vec{B} points in the $+Z_{GSM}$
146 direction). This reversal in the field direction compared to figure 1a, 1b and 1c also reverses
147 the electric field in the solar wind and the direction of the currents that flow in the bow shock

148 and magnetosheath. Hence these regions remain sources of Poynting flux. The draping of
149 the interplanetary magnetic field again deflects Poynting flux towards the magnetosphere.
150 The major difference in Figure 1d compared to the southward IMF cases is that the
151 Chapman-Ferraro currents do not reverse in direction when the IMF points northward
152 because the dominant cause of the magnetic shear across the magnetopause is the difference
153 in the magnitudes of the terrestrial and magnetosheath fields. Hence during northward IMF
154 these magnetopause currents may weaken but do not reverse in direction. Another key point
155 is that the geomagnetic tail never disappears because the timescales for that to happen are
156 very much longer than any intervals between the periods of southward IMF that generate
157 open lobe magnetic flux. This means that the dawnward electric field in the solar wind and
158 magnetosheath during northward IMF makes the north and south flanks of the tail sinks of
159 Poynting flux in this case ($\vec{J} \cdot \vec{E} > 0$). The magnetic shear persists across the cross-tail current
160 sheet in which reconnection will still occur, albeit at a reduced rate. This curl in the electric
161 field shows the tail lobe field is reducing, the stored energy being released to both the
162 magnetopause and cross-tail current sheet sinks of Poynting flux. Hence this northward IMF
163 case is an inherently non-steady-state situation and the curl in the electric field associated
164 with the decay of the field in the lobe allows \vec{E} to be in opposite directions in the centre and
165 flanks of the tail. This would persist for as long as the tail does not disappear, and as the
166 timescale for its loss is much, much greater than the duration of any interval of northward
167 IMF, this never occurs. In Figure 1d, the dayside magnetopause currents are drawn as
168 weakened because field has built up in the magnetosheath plasma depletion layer, reducing
169 the magnetic shear across the nose of the magnetosphere. The dayside acts as a sink of
170 Poynting flux partly because the magnetic flux that is lost from the tail accumulates on the
171 dayside, pushing the magnetopause at the nose of the magnetosphere in the $+X_{GSM}$ direction
172 (i.e. the stand-off distance increases). However, the key point in Figure 1d is that energy is
173 no longer extracted at the high-latitude flanks of the long geomagnetic tail. As the effect of
174 the IMF orientation influences all Poynting flux, whether it was generated from the solar
175 wind kinetic energy flux by the bow shock/magnetosheath or was present in the incident solar
176 wind, we should expect the extraction of energy from these two sources by the
177 magnetosphere to share the same IMF orientation dependence.

178 ***2.3 Adding Solar Wind Poynting Flux***

179 The efficiency of the transfer of the solar wind kinetic energy flux into the magnetosphere is
 180 accounted for in the dimensional analysis of *Vasyliunas et al.* (1982) using the dimensionless
 181 transfer function (the fraction of incident flux extracted by the magnetosphere)

$$182 \quad t_{KE} = M_A^{-2\alpha} \sin^4(\theta/2) \quad (5)$$

183 where θ is the IMF “clock angle” in the GSM frame and α is called the “coupling exponent”
 184 and is determined empirically. The clock angle θ is defined as $\arctan(B_{YM}/B_{ZM})$, where
 185 B_{YM} and B_{ZM} are the Y and Z components of the IMF in the GSM frame of reference. From
 186 careful analysis that minimises the effect of data gaps in the interplanetary data, *Lockwood et al.*
 187 *(2019a)* derived a value for α of 0.44 for all averaging timescales between 1 minute and 1
 188 year for the mid-latitude *am* geomagnetic index. The term $\sin^4(\theta/2)$ allows for the effect of
 189 IMF orientation on the transfer of energy into the magnetosphere. The supporting information
 190 for the paper by *Lockwood et al. (2019b)* confirms that this is the optimum IMF orientation
 191 factor to use when considering the extraction of the kinetic energy flux, F_{KE} , by the
 192 magnetosphere. The use of the $\sin^4(\theta/2)$ factor to quantify the effect of IMF orientation is
 193 optimum because it is a function that is not discontinuous in gradient and which allows for a
 194 low rate of reconnection taking place in the dayside magnetopause (thereby generating some
 195 open magnetospheric flux) when the IMF points northward (i.e., when the clock angle θ
 196 factor is less than 90 degrees): this has been demonstrated to be the case in a number of
 197 studies, most conclusively by observations of O^+ ions of ionospheric origin flowing out
 198 through the dayside magnetopause (*Chandler et al., 1999*).

199 The solar wind Poynting flux needs no conversion at the bow shock/magnetosheath
 200 equivalent to that needed for kinetic energy flux but, as discussed in the last section, its
 201 ability to enter the magnetosphere should also depend on the same IMF orientation factor.
 202 Hence the $M_A^{-2\alpha}$ term is not required whereas an IMF orientation factor is required. From
 203 the above arguments we expect the $\sin^4(\theta/2)$ term to apply to S_{sw} as it did to the transfer of
 204 kinetic energy flux. Later in this paper we will confirm that this is true to a good
 205 approximation and hence the transfer function for the Poynting flux can be written as

$$206 \quad t_S = f_s \sin^4(\theta/2), \quad (6)$$

207 where f_s is the fraction of the total solar wind Poynting flux power (incident on the same
 208 cross-sectional area as used to compute the total kinetic energy flux) that enters the
 209 magnetosphere. This area is $c\pi L_o^2$, where L_o is the stand-off distance of the nose of the

210 magnetosphere and c is an approximately constant area factor that allows for the shape of the
 211 dayside magnetosphere. Note that f_s can be greater than unity because the relevant cross-
 212 sectional area of the bow shock may be greater than the area $c\pi L_0^2$ relevant to kinetic energy
 213 capture by the magnetosphere. It is very difficult to estimate f_s directly without using a
 214 global MHD model of the magnetosphere: as mentioned above, the degree to which Poynting
 215 flux is deflected towards the magnetosphere depends on the degree of draping of the IMF
 216 over the nose of the magnetosphere in the magnetosheath. Furthermore, the X_{GSE} at which the
 217 Poynting flux arrives in the tail matters because if it arrives beyond the relevant tail
 218 reconnection point (i.e. the one that is closing open flux) it will almost certainly be returned
 219 to the interplanetary medium and not contribute to near-Earth phenomena such as the
 220 substorm current wedge, and hence to geomagnetic indices such as ap and am (Lockwood,
 221 2013). However, these geometric considerations apply equally, and in the same way, to
 222 Poynting flux generated by the bow shock currents and magnetosheath from the kinetic
 223 energy density flux of the solar wind. Later in this paper we empirically find optimum values
 224 for f_s that range from 0.74 for hourly data to 0.3 for annual means (although the results for
 225 timescales great than a day are found not to be statistically significant). From equations (3),
 226 (4) and (6), we can compute the ratio of total power inputs into the magnetosphere

$$227 \quad \psi = \frac{(c\pi L_0^2) S_{sw} t_S}{(c\pi L_0^2) F_{KE} t_{KE}} = \frac{2f_s \cos^2(\varphi)}{M_A^{(2-2\alpha)}} \quad (7)$$

228 using $2f_s \cos^2(\varphi) = 1$ and $\alpha = 0.44$ gives a value of this ratio $M_A^{1.12} \sim M_A$. Hence a typical
 229 M_A value of 10 means that the total power entering the magnetosphere due to solar Poynting
 230 flux may be of order a tenth of that due to the kinetic power, even though the ratio of the flux
 231 densities in the solar wind is of order 1/100.

232 The formulation of *Vasyliunas et al.* (1982) computes the area (πL_0^2) by assuming the dayside
 233 magnetopause is hemispheric in shape and in equilibrium so area L_0 is the equilibrium stand-
 234 off distance of the dayside magnetopause. *Lockwood et al.* (2019a) generalised this by using a
 235 constant multiplicative area factor c so the area presented to the solar wind is ($c\pi L_0^2$). (This
 236 does not add to the number of free fit parameters because it is eventually cancelled by
 237 normalising the power input to its overall average value). Assuming the dayside
 238 magnetopause is in equilibrium, pressure balance at the dayside magnetopause gives and
 239 expression for L_0 . The power input in to the magnetosphere due to kinetic energy density of

240 the solar wind can then be estimated (see *Vasyliunas et al.*, 1982; *Lockwood et al.* (2019a;
 241 2019b):

$$\begin{aligned}
 242 \quad P_\alpha &= (c\pi L_o^2) F_{KE} t_{KE} \\
 243 \quad &= (\pi c k_2 k_1^2 M_E^{2/3} \mu_0^{-1/3}) m_{sw}^{(2/3-\alpha)} N_{sw}^{(2/3-\alpha)} V_{sw}^{(7/3-2\alpha)} B^{2\alpha} \sin^4(\theta/2) \quad (8)
 \end{aligned}$$

244 where k_1 and k_2 are constants and M_E is the magnetic moment of the Earth which can be
 245 computed for a given time using the IGRF-15 Model (*Thébault et al.*, 2015). Because the
 246 variation of M_E with time is small and approximately linear we can treat the term in brackets
 247 as a constant that we can later cancel out by normalising P_α to its average value over the
 248 whole period P_o to give P_α/P_o .

249 From (6) we can add the power input due to solar wind Poynting flux to get the combined
 250 power input

$$251 \quad P_{\alpha 1} = P_\alpha \{1 + \psi\} = P_\alpha \left\{1 + 2f_s \cos^2(\varphi) M_A^{(2\alpha-2)}\right\} \quad (9)$$

252 Equations (7) and (8) can be used to compute the normalised total power into the
 253 magnetosphere, from both solar wind kinetic energy and solar wind Poynting flux, ($P_{\alpha 1}/P_{o1}$)
 254 from measured solar wind parameters, with two free parameters that need to be derived
 255 empirically, α and f_s .

256 Later in this paper we test the $\sin^4(\theta/2)$ IMF orientation factor used by employing a factor
 257 G defined by:

$$258 \quad G = (\pi c k_2 k_1^2 M_E^{2/3} \mu_0^{-1/3}) m_{sw}^{(2/3-\alpha)} N_{sw}^{(2/3-\alpha)} V_{sw}^{(7/3-2\alpha)} B^{2\alpha} \left\{1 + 2f_s \cos^2(\varphi) M_A^{(2\alpha-2)}\right\}. \quad (10)$$

259 To remove the constants and produce a simpler coupling function, we normalise G to its
 260 overall mean G_o .

261 As pointed out by *Vasyliunas et al.* (1982), if a plot of the observed ratio $\{am/(G/G_o)\}$
 262 against the observed $\sin^4(\theta/2)$ and get a proportional variation with slope s , we can write

$$263 \quad am = s (G/G_o) \sin^4(\theta/2) = (s/G_o) P_{\alpha 1} = s_1 P_{\alpha 1} = s_2 (P_{\alpha 1}/P_{o1}) \quad (11)$$

$$264 \quad \text{where} \quad s_2 = s (P_o/G_o) \quad (12)$$

265 Hence if we find proportionality s is constant and, because P_o and G_o are both constants, s_2 is
266 also a constant and hence $(P_{\alpha 1}/P_{o1})$ is a proportional predictor of am .

267 3. Observations

268 The study presented here is based on 12,097,440 1-minute samples of interplanetary
269 parameters observed in the interval 1995-2017 (inclusive), downloaded from the Omni
270 database compiled and maintained by the Space Physics Data Facility at NASA's Goddard
271 Space Flight Center. If any of the parameters used in compiling the power input estimates
272 are missing, this generates a data gap and, as a result, the 1-minute dataset contains 9,530,831
273 valid samples of both the old and new power input estimates P_α and $P_{\alpha 1}$, an availability of
274 78.78%. These were then used to generate hourly means. We require the number of minute
275 samples within the hour, N , to exceed a limit N_{lim} which depends on the parameter in
276 question and the accuracy required. *Lockwood et al.* (2019a) carried out a Monte-Carlo study
277 on data from 22-years in which they introduced data gaps at random into data for hours when
278 all 60 1-minute samples were available, and so derived the N_{lim} values needed to give
279 uncertainties of 2% and 5%. We here adopt the 5% criteria. The autocorrelation time of the
280 solar wind speed, V_{sw} , is sufficiently large that a single sample in the hour meets the 5%
281 requirement and so $N_{lim} = 1$. On the other hand, the IMF orientation clock angle, θ , has a
282 very short autocorrelation time that gives $N_{lim} = 56$. Applying the $N \geq N_{lim}$ criteria (to give
283 uncertainties in hourly means of all parameters in P_α and $P_{\alpha 1}$, that are below 5%), gave
284 161,627 hourly mean estimates of P_α and $P_{\alpha 1}$, out of a possible 201,624, an availability of
285 80.16%. (Because the criteria have the same effect on P_α and $P_{\alpha 1}$ this number applies to the
286 power input estimates with and without the solar wind Poynting flux. We then made 3-hourly
287 means only if all three of the hourly means are available. The interval covers 8401 days. We
288 use a mean value for a day if just one 3-hourly mean is available from the day, but employ
289 the piecewise removal of all the geomagnetic am index data for times when the
290 corresponding P_α value is missing (*Finch and Lockwood, 2007*), this gives 8401 daily
291 samples, and availability of 99.69%. The available hourly means were also averaged to give
292 23 annual means and 309 averages over Carrington synodic solar rotation periods of 27.26
293 days (654 hours).

294 The *am* geomagnetic index data are generated and made available by The International
 295 Service of Geomagnetic Indices (ISGI), France and collaborating institutes. The stations used
 296 to compile the *am* index (Mayaud, 1980) are situated at sub-auroral latitudes close to
 297 corrected geomagnetic latitude $\Lambda_{CG} = 50^\circ$. There are 15 stations in current use in the northern
 298 hemisphere and 10 in the southern. They are grouped into longitude sectors, with 5 such
 299 groups in the Northern hemisphere, and 4 in the Southern. The *K* indices for stations in a
 300 longitude sector are averaged together and the result is converted into a sector a_K value using
 301 the standard K2aK scale. Weighted averages of these sector a_K values are then generated in
 302 each hemisphere giving a_n and a_s , the weighting factors accounting for the differences in the
 303 longitude extents of the sectors. The index *am* is equal to $(a_n+a_s)/2$. Note that we here
 304 employ all available *am* data up to the end of 2017 and that after the end of 2014 these data
 305 are classed as “provisional” which means they have passed initial quality checks and can be
 306 used, but not yet been through the final review that defines them as “definitive”. Lockwood *et al.*
 307 (2018b) have developed a model of the sensitivity of a geomagnetic station to solar wind
 308 forcing that can be used to derived the time-of-day/time-if-year response of any geomagnetic
 309 index provided that, like *am*, is compiled using an analytic algorithm. This was used by
 310 Lockwood *et al.* (2019d) to show that the *am* index response is exceptionally uniform, the
 311 standard deviation being just 0.65% of the mean value. Analysis has demonstrated how mid-
 312 latitude “range” indices such as *am* respond primarily to the substorm current wedge of
 313 substorm expansion phases (see Finch *et al.*, 2008; Lockwood, 2013; and the Supporting
 314 Information file associated with Lockwood *et al.*, 2019a).

315 Figure 2 summaries the datasets used. The left hand plots show daily means (averaging
 316 timescale $\tau = 1$ day) and the right hand plots show annual means ($\tau = 1$ year) for 1995-2017
 317 (inclusive). From top to bottom: (a and b) the *am* index; (c and d) the normalised power
 318 input to the magnetosphere (P_α/P_o , where P_o is the mean of P_α for the whole interval); (e and
 319 f) the additional Poynting flux term for unit f_S , (ψ/f_S) (see equation 9 of text); (g and h) the
 320 solar wind Alfvén Mach number, M_A ; and (i and j) the cosine of the angle φ between the
 321 IMF vector \vec{B} and the solar wind velocity vector \vec{V}_{sw} . It can be seen that (ψ/f_S) is highly
 322 anti-correlated with M_A .

323 4. Correlation Analysis

324 4.1 Daily and Annual Averaging Timescales

325 Figure 3 shows correlations between the *am* index and the power input to the magnetosphere
 326 estimate, $P_{\alpha 1}$ which allows for the solar wind Poynting flux (using equation 9). The
 327 correlation coefficients are colour-contoured as function of the coupling exponent α (vertical
 328 axis) and the transfer fraction for solar wind Poynting flux, f_S (horizontal axis). Note that the
 329 left hand edge of these plots is for $f_S = 0$, when $P_{\alpha 1}$ reduces to P_α , the power input from solar
 330 wind kinetic energy flux alone. The plot on the left is for daily data (averaging timescale, $\tau =$
 331 1 day) and the right hand plot is for annual data ($\tau = 1$ year). Both plots show that the α
 332 giving peak correlation falls as f_S is increased and that for $f_S > 1$ the peak correlation also
 333 falls. The black dot in (a) marks the peak correlation for $\tau = 1$ day which is at $\alpha = 0.36$ and f_S
 334 $= 0.68$ for which the correlation is $r = 0.924$. The cyan diamond is for $f_S = 0$ and is at $\alpha =$
 335 0.42 for which the correlation is slightly lower, being $r = 0.908$. The mauve dot in (b) marks
 336 the peak correlation for $\tau = 1$ year which is at $\alpha = 0.32$ and $f_S = 0.30$ for which the
 337 correlation is $r = 0.980$. The green diamond is for $f_S = 0$ and is at $\alpha = 0.34$ for which the
 338 correlation is very close to the peak value, being $r = 0.979$.

339 The question arises if these small increases in correlation, achieved by allowing for the solar
 340 wind Poynting flux, are statistically significant. This question is addressed by Figure 4. We
 341 here look at the significance of the difference between two correlations by computing the *p*-
 342 value of the null hypothesis that they are the same. We use the Meng-Z test for the difference
 343 between the correlations between A and B and between A and C which allows for the inter-
 344 correlation of B and C (Meng *et al.*, 1992). We test against the AR1 red-noise model by using
 345 the effective number of independent data pairs, N_{eff} given by:

$$346 \quad N_{eff} = N \frac{(1 - a_1)}{(1 + a_1)} \quad (13)$$

347 where N is the actual number of data pairs and a_1 is the autocorrelation of A at lag 1 (Wilks,
 348 1995).

349 The top panels of Figure 4 are correlograms showing linear correlation coefficient r as a
 350 function of α for the (fixed) best-fit value of f_S - hence they are vertical slices through Figure
 351 3. As in Figure 3, the plot on the left is for $\tau = 1$ day and the right hand plot is for $\tau = 1$ year.
 352 The line colours are coded in the same way as the points in Figure 3, so the black and mauve
 353 lines are for the optimum f_S whereas the cyan and green lines are for $f_S = 0$. In each case, the
 354 α giving peak correlation is marked with a vertical dashed line and the symbols used in
 355 Figure 3. The second row of panels uses the same colour scheme to show the *p*-values for

356 the null hypothesis that the correlation is not significantly higher if the solar wind Poynting
357 flux is added (using the optimum combination of α and f_S), as computed using the Meng-Z
358 test. Hence at the peak r , the black curve in the middle-left panel gives $p = 1$ but the p -value
359 falls either side of this as r falls below its peak value. The horizontal dashed lines mark the
360 level at which this drop in r becomes significant at the 1σ and 2σ levels. The cyan line gives
361 the corresponding p -value for the difference between r (at a given α and for $f_S = 0$) and the
362 peak r for the optimum f_S ($= 0.68$) and α ($= 0.36$). These p -values are very small and so they
363 have been plotted times 100 in Figure 4c: there is a small peak at the peak r for $f_S = 0$ at $\alpha =$
364 0.42 , but even at that peak, the r is significantly lower than for $f_S = 0.68$ at greater than the
365 3σ level. Hence adding the solar wind Poynting flux has significantly improved the
366 correlation for $\tau = 1$ day, even though the increase in r is actually rather small. This is
367 stressed by the similarity of the scatter plots shown in the bottom left panel for the α of peak
368 correlation for $f_S = 0.68$ (black points) and for $f_S = 0$ (cyan points).

369 The right hand panels are the corresponding plots for $\tau = 1$ year. In this case the difference
370 for $f_S = 0$ is not significant at even the 1σ level at almost all α (green line in the middle-right
371 panel). Hence although significant improvement can be made for 1-day data by adding the
372 solar wind Poynting flux, this is not true for annual data.

373 **4.2 Three-Hourly and Hourly Averaging Timescales**

374 The power input to the magnetosphere has a characteristic variation with both time-of-year F
375 and UT because of the Russell-McPherron effect of the angle of rotation between the
376 Geocentric Solar Ecliptic (GSE) and GSM reference frames (*Russell and McPherron, 1973*).
377 In addition, there are dipole tilt effects on the geomagnetic response which means that am
378 shows a marked equinoctial variation (*Cliver et al., 2000*). (See *Lockwood et al. (2016)* for a
379 discussion of the proposed mechanisms). Hence there are strong UT dependencies in the
380 solar wind-magnetosphere-ionosphere system that make it desirable to repeat the study for
381 sub-daily averaging timescales. The problem is that there is a variable lag in the am response
382 that becomes an increasingly significant factor as the averaging timescale is reduced. For the
383 3-hourly resolution of the am data, neglecting a response lag of 1 hour would mean that a
384 third of the data in the interplanetary average (33%) was not data relevant to the
385 corresponding three-hourly am data. By way of comparison, neglecting the same lag for daily
386 averages would mean that (1/24) of the data were not relevant (4.1%), and this falls to 0.15%
387 for averages over Carrington Rotation (CR) intervals and (4×10^{-4})% for annual means. There

388 are high time geomagnetic resolution indices, such as the auroral electrojet indices (*AE*, *AU*
389 and *AL*) and their SuperMAG equivalents (*SME*, *SMU* and *SML*) which have 1-minute
390 resolution that could be used in this context, but these are all generated from northern
391 hemisphere stations only which means that they have a uneven response and give a spurious
392 annual modulation which *am*, on the other hand, reduces to very low levels by optimum
393 choice of stations and the use of longitude sector weightings (*Lockwood et al.*, 2019d).

394 Rather than move to a different geomagnetic index in order to study sub-daily timescales, we
395 here adopt a different approach that enables us to continue to use the *am* index. The Omni
396 interplanetary dataset has been lagged from the time and place of observation to the nose of
397 Earth's bow shock (*Case and Wild*, 2012). The *am* index responds after a lag *dt* that is the
398 sum of the propagation delay from the nose of the bow shock to the dayside magnetopause,
399 the duration of the substorm growth phase (when energy accumulates in the near-Earth tail
400 lobe magnetic field), the propagation time from current disruption in the near-Earth cross tail
401 current sheet to the nightside auroral oval, and the time between substorm onset and the peak
402 response of *am*.

403 To derive the optimum lag *dt* we use the *am* data, assigned to the times t_{am} of the mid-points
404 of the 3-hour intervals ($\tau = 3$ hours) over which each *am* estimate is made (i.e., 1.5 UT for the
405 0-3 UT interval, 4.5 UT for 3-6 hrs, up to 22.5 UT for 21-24 hrs UT). Three-hourly means of
406 one-minute interplanetary data were generated over intervals between $(t_{am} - \tau/2 - dt)$ and $(t_{am}$
407 $+ \tau/2 - dt)$ for a given lag *dt*. To construct the 3-hourly means of interplanetary data we make 3
408 one-hour means and use the criteria derived by *Lockwood et al.*, (2019a) of the numbers of
409 samples in the hour that are required to make the uncertainty in the hourly mean of that
410 parameter 5% or less. These three one-hour means are then averaged to give the 3-hourly
411 mean and valid data points require that all three one-hour means are available. The
412 correlations between $P_{\alpha}(t_{am} - dt)$ and $am(t_{am})$ for each *F*-UT bin were then evaluated for lags
413 *dt* which was varied between zero and 300 min (5 hours). For each *F*-UT bin, the lag giving
414 peak correlation was determined and the distribution of *dt* for the 160 bins is shown in Figure
415 5a. The peak correlations were always above 0.65 and most correlations near the mode *dt*
416 value exceeded 0.95.

417 To look for an effect of activity level on the lag *dt*, the *am* data were then further sub-divided
418 into quantile ranges. We use the notation that $q(n)$ is the $(100 \times n)\%$ quantile of the
419 distribution of *am* values. Five of the quantile ranges used each contained 20% of the *am*

420 data being: $0 \leq am \leq q(0.2)$; $q(0.2) < am \leq q(0.4)$; $q(0.4) < am \leq q(0.6)$; $q(0.6) < am \leq q(0.8)$;
 421 and $q(0.8) < am \leq q(1)$. In addition, we looked at three quantile ranges to study the most
 422 active periods covering the top 10% am values, $q(0.9) < am \leq q(1)$; the top 5%, $q(0.95) < am$
 423 $\leq q(1)$; and the top 1%, $q(0.99) < am \leq q(1)$. The results are shown in Figure 5b. The peak of
 424 the distribution is near $dt = 60$ min for $0 \leq am \leq q(0.2)$; and rises with am to near 70 min for
 425 $q(0.2) < am \leq q(0.4)$; $q(0.4) < am \leq q(0.6)$; and $q(0.6) < am \leq q(0.8)$. However, for the largest
 426 am values the peak dt falls again and is in the range 40-50 min. In all cases there are only a
 427 few examples that give $dt < 0$: these are almost non-existent for low am levels and occur
 428 mostly for high activity levels when persistent solar wind forcing is likely to generate them
 429 from chance occurrences in the data series. A number F -UT combinations and am activity
 430 levels do give dt smaller than about 10 minutes which could indicate a directly-driven
 431 response in am . For low am levels these low- dt cases are extremely rare but for largest 1% of
 432 am values they are almost as common as the occurrences of the peak lag suggesting prior
 433 energy input resulting in a very large total energy stored in the near-Earth tail lobes (as for
 434 the $dt < 0$ cases). There are also some cases of very long lags for the top 1% of am data,
 435 suggesting that following very large energy input to the magnetosphere a series of substorms
 436 are required to return the system to lower stored energy levels. We here use $dt = 60$ min. as
 437 an overall average response lag but note that there is considerable variability in dt about this
 438 value which will lower correlations and increase noise in statistical studies.

439 We also make a study on hourly timescales ($\tau = 1$ hour). To do this we take means of the
 440 interplanetary data over one-hour intervals centered on times t_{sw} and compare these to am
 441 values that are interpolated from the observed 3-hourly am data to the times $(t_{sw}+dt)$, where
 442 dt is the derived optimum lag of 60 min., the interpolations being carried out using the
 443 Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) procedure. The reason these
 444 hourly data are instructive is because we find dt is of order this timescale. The am index
 445 responds primarily to the substorm current wedge during substorm expansion phases, which
 446 means that fluctuations in the power input to the magnetosphere on timescales below dt are of
 447 lesser importance because they are averaged out as energy is accumulated in the geomagnetic
 448 tail lobe field during the substorm growth phases, which is the largest contribution to the
 449 overall response lag dt . Hence by extending the analysis down to $\tau = 1$ hour we are covering
 450 the full range of timescales over which fluctuations in power input to the magnetosphere are
 451 not averaged out by the accumulation of magnetic energy in the tail lobes.

452 The left hand panels of Figure 6 are equivalent to Figure 4 for the 3-hourly averaged data
 453 obtained for this optimum dt and show the behaviour is quite similar to that for the daily data.
 454 The correlation is lower (as we would expect because of the variability in actual dt values)
 455 and the difference between with and without the solar wind Poynting flux is again very small.
 456 However the difference in the optimum α for these two cases is larger than for daily data. For
 457 hourly data (not shown) the results are very similar, but the scatter is greater and the
 458 correlation coefficients lower.

459 **4.3 Dependence on Averaging Timescales**

460 To fill the large gap in timescales between 1 day and 1 year, the right-hand panels in Figure 6
 461 repeat the same analysis for averages over Carrington synodic solar rotation periods of 27.26
 462 days. Table 1 summarises the results for the 4 averaging timescales analysed in Figures 4 and
 463 6 plus the one-hour data obtained by applying interpolation to the *am* observations. The
 464 optimum α required is almost constant if Poynting flux is included but increases with
 465 decreasing timescales if Poynting flux is omitted. The correlations are always very slightly
 466 higher if Poynting flux is included. The p -value that the correlations are not significantly
 467 different is high for annual and Carrington means, but low for daily, 3-hourly and hourly
 468 averages.

469 **5. The IMF orientation factor**

470 It was argued in Section 2 that the Poynting flux entry into the magnetosphere has a transfer
 471 function that depends on IMF orientation with the same $\sin^4(\theta/2)$ dependence for both
 472 energy derived from the kinetic energy flux of solar wind particles and from the Poynting
 473 flux in the solar wind. Where here test this using the method recommended by *Vasyliunas et*
 474 *al.* (1982), namely plotting the observed ratio $\{am/(G/G_o)\}$ against the observed values of
 475 $F_\theta = \sin^4(\theta/2)$, where G is given by equation (10) and is the total proposed coupling
 476 function, (allowing for both the kinetic energy flux and Poynting flux in the solar wind)
 477 without the IMF orientation factor, F_θ . G_o is the overall mean of G . The results are shown in
 478 Figure 7 for $\tau = 1$ day and Figure 8 for $\tau = 3$ hours. The study cannot be performed for $\tau =$
 479 27.26 days nor $\tau = 1$ year because the central limit theorem means that the width of the
 480 distribution of F_θ reduces with increased τ : Figure 2 of *Lockwood et al.* (2017) shows that

481 at $\tau > 1$ day a full range of F_θ values is not present (and even at $\tau = 1$ day samples of $F_\theta = 0$
 482 and, in particular, of $F_\theta = 1$ are rare). The narrowing of the distribution is such that F_θ can
 483 even be considered to have a single, constant, value to within an accuracy of $\pm 4.9\%$ (at the
 484 $1-\sigma$ level) for $\tau = 1$ year and to within $\pm 10.3\%$ for $\tau = 27$ days. Note that the equivalent
 485 plots to Figure 8 for magnetospheric power input, without the solar wind Poynting flux term,
 486 are shown in the supporting information file attached to the paper by *Lockwood et al.*
 487 [2019b]. The daily data (grey dots) in figure 7 and show considerable scatter but a linear
 488 trend, emphasised by the black dots that are means in non-overlapping bins of
 489 $F_\theta = \sin^4(\theta/2)$ that are 0.05 wide (the error bars are plus and minus one standard
 490 deviation). As discussed in section (2) this linearity shows that $\sin^4(\theta/2)$ remains an
 491 appropriate form for the IMF orientation term F_θ after we have allowed for solar wind
 492 Poynting flux. Figure 8 repeats this study for the 3-hourly data. It can be seen that the scatter
 493 in the individual data points is greater and the binned means do not agree quite as well with
 494 the linear fits as they do for the $\tau = 1$ day case. (The root mean square deviation of the binned
 495 means for $\tau = 3$ hours is $\Delta_{\text{rms}} = 0.071$ whereas for $\tau = 1$ day $\Delta_{\text{rms}} = 0.061$). Figure 9 shows
 496 how the $\sin^4(\theta/2)$ factor performs relative to other proposed alternatives. The observed ratio
 497 $\{am/(G/G_o)\}$ for $\tau = 1$ day in the 24 bins of IMF clock angle θ that are 7.5° wide are scaled
 498 to be between 0 and 1 and then plotted as a function of θ . As discussed by *Lockwood et al.*
 499 (2019b), the exponent n in $\sin^n(\theta/2)$ has been proposed to be between 2 and 6. In Figure 9
 500 we also test the frequently used function $U(\theta)\cos(\theta)$ where $U(\theta) = 1$ for southward IMF (θ
 501 $> 90^\circ$) and $U(\theta) = 0$ for northward IMF ($\theta \leq 90^\circ$). It can be seen that $n = 4$ gives the best
 502 fit, although the agreement is not quite as good as for the case when solar wind Poynting flux
 503 is omitted (see Figure S-12 of the Supporting Information file associated with *Lockwood et*
 504 *al.*, 2019b). A slightly better fit would be obtained using n of 3.8 but this is not statistically
 505 significant. The implications of the lack of a major effect of averaging timescale on the
 506 analysis of the optimum IMF orientation factor is discussed in the next section.

507 6. Discussion and Conclusions

508 We have used the near-continuous interplanetary data available for 1995-2017 (inclusive) to
 509 investigate if allowing for the solar wind Poynting flux significantly increases the correlation
 510 with geomagnetic activity. The use of near continuous solar wind data is very important
 511 because *Lockwood et al.* (2019a) have shown that the frequently-made assumption that the

512 effect of data gaps just “averages out” is invalid and causes serious error. In effect, the
513 presence of data gaps introduces noise into correlation studies which gives “overfitting”,
514 where the derived coupling function is fitted to the noise in the data and the overfitted
515 refinements to the coupling function have no predictive power as they are not robust. This is
516 important because adding the solar wind Poynting flux to the kinetic energy input adds an
517 additional unknown parameter that has to be fitted empirically and it is the use of too many
518 free fit variables that makes overfitting a particular problem.

519 It has been shown that a very slight improvement to the energy input to the magnetosphere
520 coupling function can be made by allowing for the Poynting flux that is in the solar wind and
521 adding it to the dominant solar wind energy flux (namely the kinetic energy density flux of
522 the particles). The improvement is small (raising the correlation with the *am* index from
523 0.908 to 0.924 for daily means and from 0.979 to 0.980 for annual means) and statistically
524 significant for data averaged on daily timescales or less but not significant for annual means
525 or means taken over Carrington Rotation intervals.

526 We derive empirically a value for the coupling fraction for solar wind Poynting flux, f_S , of
527 0.3 for annual means and so from the (ψ/f_S) variation for annual means shown in the right
528 column of Figure 2, we estimate that ψ varies between 0.016 at sunspot minimum and 0.035
529 at sunspot maximum, i.e. the contribution of solar wind Poynting flux to total power input
530 varies over the range 1.6-3.5% on these annual timescales. For daily values, the best estimate
531 of f_S is 0.68 which gives a larger range for the contribution of solar wind Poynting flux to
532 total power input of 0-47%. The distribution of daily ψ varies is lognormal in form with a
533 mode value of 0.060 and 10 and 90 percentile values of 0.030 and 0.119. Hence the most
534 common percentage of total power into the magnetosphere that arises from solar wind
535 Poynting flux on daily averaging timescales is 6% and of all daily values, 80% lie in the
536 range 3-12%.

537 Note that the frequently-used epsilon coupling function, ϵ , is based on the incorrect
538 assumption that the relevant energy flux in the solar wind is the Poynting flux and, although
539 this can be made consistent with the energy coupling function based on the dominant solar
540 wind kinetic energy flux, P_α , using an extreme value of the coupling exponent α , this does
541 not give as good agreement with geomagnetic indices as does the optimum value of α . This is
542 why ϵ performs considerably less well than P_α (and hence $P_{\alpha 1}$) on all averaging timescale
543 (see *Finch and Lockwood, 2007*).

544 Lastly, we noted in the last section that the optimum form of the IMF orientation factor F_θ is,
545 as expected, the same for the Poynting flux that is generated at the bow shock (or in the
546 magnetosheath) from the solar wind kinetic energy density and for the Poynting flux which
547 was present in the solar wind prior to it hitting the bow shock. In addition, the optimum form
548 of the IMF orientation factor at F_θ was found not change with timescale, although the noise
549 in the analysis is greater at low τ (largely caused by the increased importance of the variable
550 response lag of am) and at large τ the distribution of F_θ narrows to an almost constant value,
551 as shown in Figure 2 of *Lockwood et al.*, 2017). At high time resolution (1 minute) the
552 distribution of the optimum F_θ has an unexpected form with a great many samples in a narrow
553 spike at $F_\theta = 0$ (see explanation in Figure 9 of *Lockwood et al.*, 2019b). Figure 8 of
554 *Lockwood et al.* (2019b) and Figure 4 of *Lockwood et al.* (2019c) show that it is the
555 variability and distribution of F_θ which sets the distribution of power input to the
556 magnetosphere at high resolutions (1-minute) and that averaging causes these distributions to
557 evolve towards a log-normal form at $\tau = 1$ day which matches closely that in the am and ap
558 geomagnetic indices. For timescales τ up to the response lag $dt \sim 60$ min., the geomagnetic
559 response closely follows the average of the IMF orientation factor because during substorm
560 growth phases the effects the storage of energy integrate, and hence average out, the effects
561 of the rapid fluctuations in power input to the magnetosphere. Hence it is significant that
562 $F_\theta = \sin^4(\theta/2)$ factor works well right down to τ of 1 hour. At τ above 1 day, the F_θ factor
563 becomes increasing less important as it tends towards its quasi-constant value at $\tau = 1$ year.

564 **Acknowledgements.** The author is grateful to the staff of Space Physics Data Facility,
565 NASA/Goddard Space Flight Center, who prepared and made available the OMNI2 dataset used. The
566 data were downloaded from <http://omniweb.gsfc.nasa.gov/ow.html>. They are also grateful to the
567 staff of L'École et Observatoire des Sciences de la Terre (EOST), a joint of the University of
568 Strasbourg and the French National Center for Scientific Research (CNRS) and the International
569 Service of Geomagnetic Indices (ISGI) for making the am index data available from
570 http://isgi.unistra.fr/data_download.php. The work presented in this paper is supported by STFC
571 consolidated grant number ST/M000885/1, and by the SWIGS NERC Directed Highlight Topic Grant
572 number NE/P016928/1/.

573 **References**

574 Case, N. A., and J. A. Wild (2012) A statistical comparison of solar wind propagation delays
575 derived from multispacecraft techniques, *J. Geophys. Res.*, **117**, A02101,
576 doi:10.1029/2011JA016946.

577 Chandler, M.O., S.A. Fuselier, M. Lockwood and T.E. Moore (1999) Evidence of component
578 magnetic merging equatorward of the cusp, *J. Geophys. Res.*, **104** , 22623-22648, doi:
579 10.1029/1999JA900175

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

582 Cliver, E. W., Kamide, Y., and Ling, A. G. (2000) Mountains versus valleys: Semiannual
583 variation of geomagnetic activity, *J. Geophys. Res.*, **105**(A2), 2413– 2424,
584 doi:10.1029/1999JA900439.

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

588 Finch, I. D., & Lockwood, M. (2007). Solar wind-magnetosphere coupling functions on
589 timescales of 1 day to 1 year, *Annales Geophysicae*, **25** (2), 495–506. doi: 10.5194/angeo-25-
590 495-2007

591 Finch, I.D., M. Lockwood, A. P. Rouillard (2008) The effects of solar wind magnetosphere
592 coupling recorded at different geomagnetic latitudes: separation of directly-driven and
593 storage/release systems, *Geophys. Res. Lett.*, **35**, L21105, doi:10.1029/2008GL035399

594 Gonzalez, W. D., B.T. Tsurutani, A.L.C. Gonzalez, E.J. Smith, F. Tang, and S.-I. Akasofu
595 (1989) Solar wind-magnetosphere coupling during intense magnetic storms (1978-1979),
596 *Journal of Geophysical Research*, **94** (A7), 8835-8851. doi: 10.1029/JA094iA07p08835

597 Koskinen, H. E. J., and E. Tanskanen (2002) Magnetospheric energy budget and the epsilon
598 parameter, *J. Geophys. Res.*, **107** (A11), 1415. doi:10.1029/2002JA009283, 2002

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

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

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

608 Lockwood, M., M.J. Owens, L.A. Barnard S. Bentley, C.J. Scott, and C.E. Watt (2016) On
609 the Origins and Timescales of Geoeffective IMF, *Space Weather*, **14**, 406-432, doi:
610 10.1002/2016SW001375

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

614 Lockwood, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt and S. Bentley (2018a) Space
615 Climate and Space Weather over the past 400 years: 2. Proxy indicators of geomagnetic
616 storm and substorm occurrence, *J. Space Weather Space Clim.*, **8**, A12, doi:
617 10.1051/swsc/2017048

618 Lockwood, M., I.D. Finch, A. Chambodut, L.A. Barnard, M.J. Owens, and E. Clarke (2018b)
619 A homogeneous aa index: 2. hemispheric asymmetries and the equinoctial variation, *J. Space*
620 *Weather Space Clim.*, **8**, A58, doi: 10.1051/swsc/2018044

621 Lockwood, M. S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt, and O. Allanson,
622 (2019a). The development of a space climatology: 1. Solar wind magnetosphere coupling as a
623 function of timescale and the effect of data gaps, *Space Weather*, **17**. doi:
624 10.1029/2018SW001856

625 Lockwood, M., S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt, O. Allanson,
626 and M.P. Freeman (2019b). The development of a space climatology: 2. The distribution of
627 power input into the magnetosphere on a 3-hourly timescale. *Space Weather*, **17**. doi:
628 10.1029/2018SW002016

629 Lockwood, M., S. Bentley, M.J. Owens, L.A. Barnard, C.J. Scott, C.E. Watt, O. Allanson and
630 M.P. Freeman (2019c) The development of a space climatology: 3. The evolution of
631 distributions of space weather parameters with timescale, *Space Weather*, **17**, 180-209. doi:
632 10.1029/2018SW002017

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

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

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

640 Owens, M.J., M. Lockwood and L.A. Barnard (2018) Ion charge states and potential
641 geoeffectiveness: The role of coronal spectroscopy for space-weather forecasting, *Space*
642 *Weather*, **16**, doi: 10.1029/2018SW001855, (2018)

643 Palmroth, M., Pulkkinen, T. I., Janhunen, P., & Wu, C. C. (2003) Stormtime energy transfer
644 in global MHD simulation. *Journal of Geophysical Research: Space Physics*, **108**(A1),
645 doi.org/10.1029/2002JA009446

646 Russell, C. T., and McPherron, R. L. (1973) Semiannual variation of geomagnetic activity, *J.*
647 *Geophys. Res.*, **78**(1), 92– 108, doi:10.1029/JA078i001p00092.

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

652 Thébaud E, C.C. Finlay, C.D. Beggan, P. Alken, J. Aubert et al. (2015) International
653 Geomagnetic Reference Field: the 12th generation, *Earth Planet Space* **67**, 79. doi:
654 10.1186/s40623-015-0228-9.

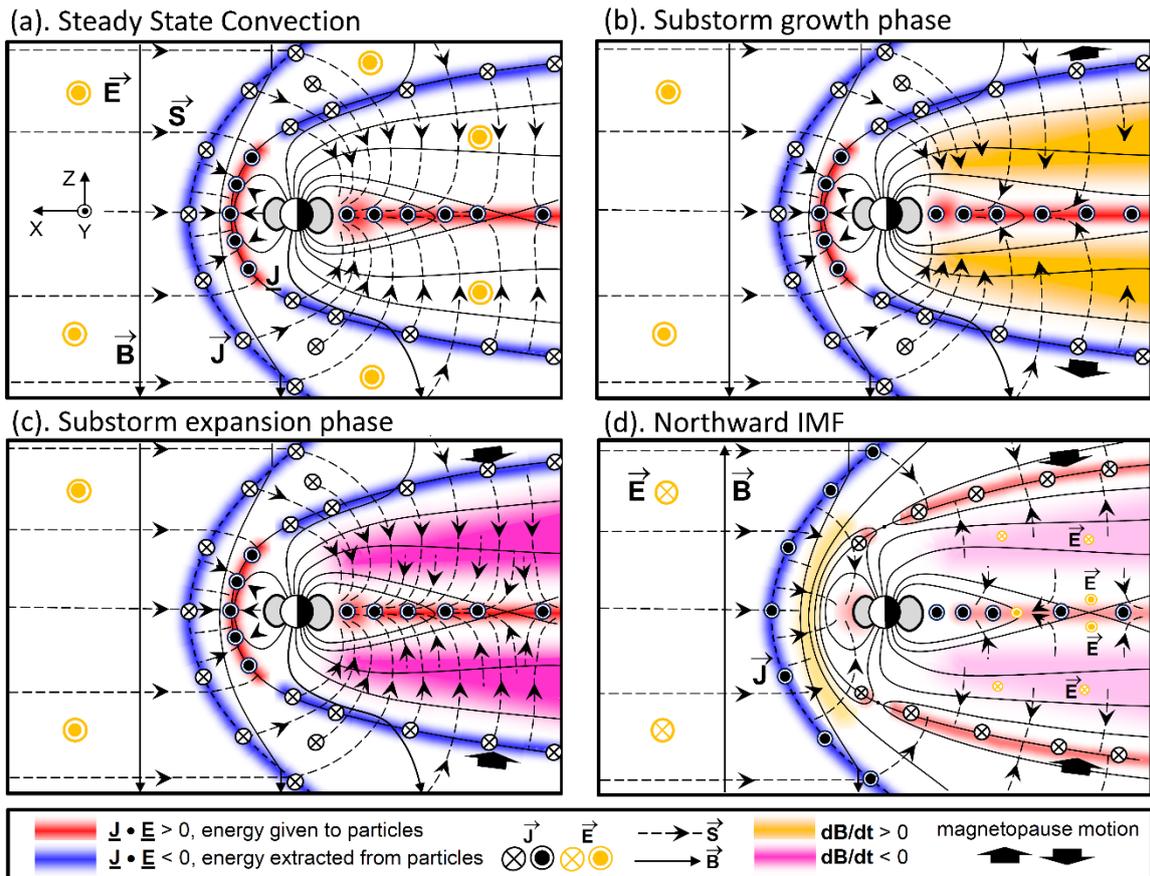
655 Vasyliunas, V. M., Kan, J. R., Siscoe, G. L., & Akasofu, S.-I. (1982). Scaling relations
656 governing magnetospheric energy transfer. *Planetary and Space Science*, 30(4), 359–365. doi:
657 10.1016/0032-0633(82)90041-1

658 Wilks, D.S. (1995) Statistical methods in the atmospheric sciences, International Geophysical
659 Series, Volume 59, ISBN-10: 0127519653. ISBN-13: 978-0127519654, Academic Press,
660 London.

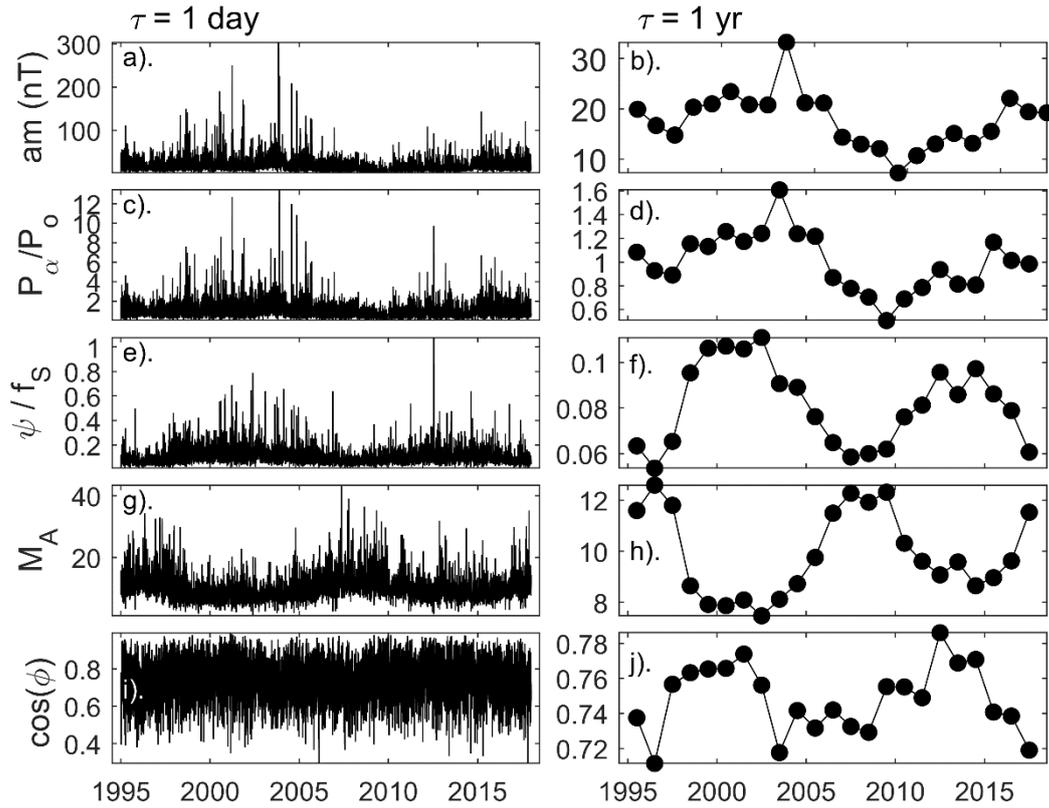
661 **Table 1.** Summary of correlations between estimated power input to the magnetosphere and
662 the *am* geomagnetic index for 1995-2017 (inclusive) and averaging timescales, τ , of 1 year, 1
663 (synodic) Carrington rotation (CR = 27.28 days), 1 day, 3 hours and 1 hour. (For $\tau = 1$ hour,
664 3-hourly *am* index data are interpolated using PCHIP interpolation and allowing for the
665 derived optimum 60-minute response lag of *am* with respect to the solar wind data). For
666 each τ , analysis is done with and without including solar wind Poynting flux. In each case,
667 the number of samples N , the best fit factor f_s (by definition f_s is zero if S is not included),
668 the best-fit coupling exponent α , and the peak correlation, r_p , are given. The last column
669 gives the p -value that the difference in the peak correlation without S is the same as that with
670 S . The number of the figure showing the data and the line colour used in that figure are given
671 for each case.

τ	with S ?	figure #	line colour	N	f_s	α	peak r	p -value of difference in peak r
1 year	N	4	green	23	0	0.34	0.978	0.65
	Y		mauve		0.30	0.32	0.979	
1 CR	N	6	green	309	0	0.37	0.9315	0.98
	Y		mauve		0.65	0.35	0.9316	
1 day	N	4	cyan	8401	0	0.42	0.908	0.001
	Y		black		0.68	0.36	0.924	
3 hours	N	6	cyan	67,208	0	0.43	0.842	0.03
	Y		black		0.69	0.39	0.843	
1 hour	N	-	-	161,627	0	0.49	0.7866	0.02
	Y		-		0.74	0.38	0.7886	

672

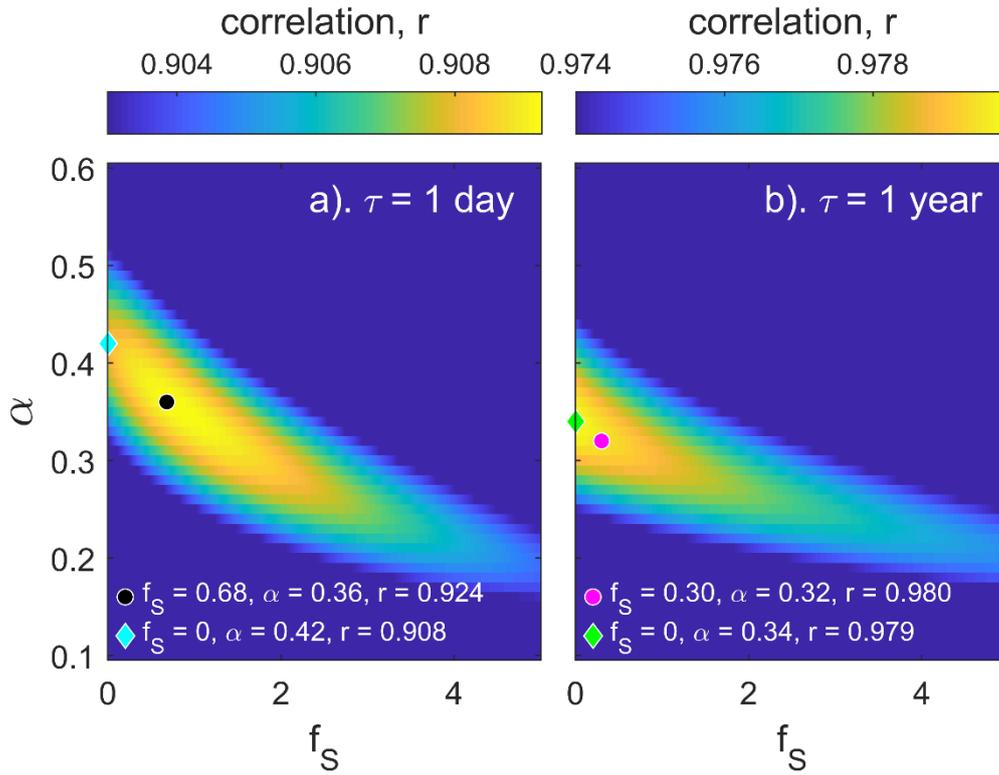


673 **Figure 1.** Schematic noon-midnight cross-sections of the magnetosphere for (a) steady-state
674 magnetospheric convection during southward IMF; (b) a substorm growth phase; (c) a
675 substorm expansion phase and (d) persistent northward IMF. Part (a) is after *Cowley (1991)*,
676 parts (b)-(d) after *Lockwood (2004)*. See text and key at the base of the figure for explanation
677 of coloured regions and symbols.

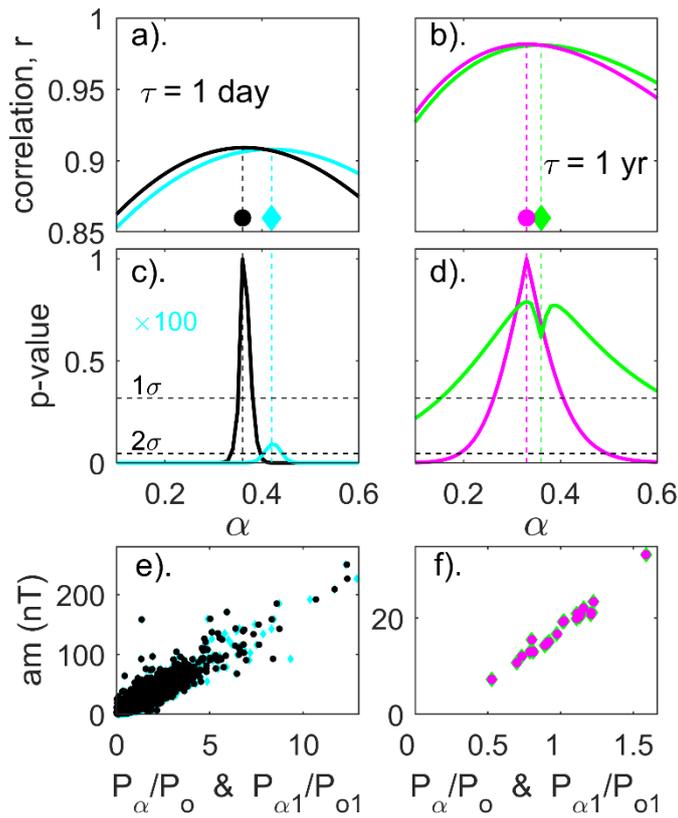


678

679 **Figure 2.** Daily (left) and annual (right) means of time-series data used in this paper. From
 680 top to bottom: (a and b) the am geomagnetic index; (c and d) the power input to the
 681 magnetosphere based on solar wind kinetic energy flux normalised to its overall mean, P_{α}/P_0 ;
 682 (e and f) a factor relating to solar wind Poynting flux entering the magnetosphere ψ/f_S (see
 683 text for details and equation (6) for definition); (g and h) the solar wind Alfvén Mach number
 684 M_A ; and (i and j) $\cos(\varphi) = (B_z^2 + B_y^2)^{1/2}/B$ where B_z and B_y are two orthogonal
 685 components of the IMF B that are perpendicular to the solar wind flow direction.

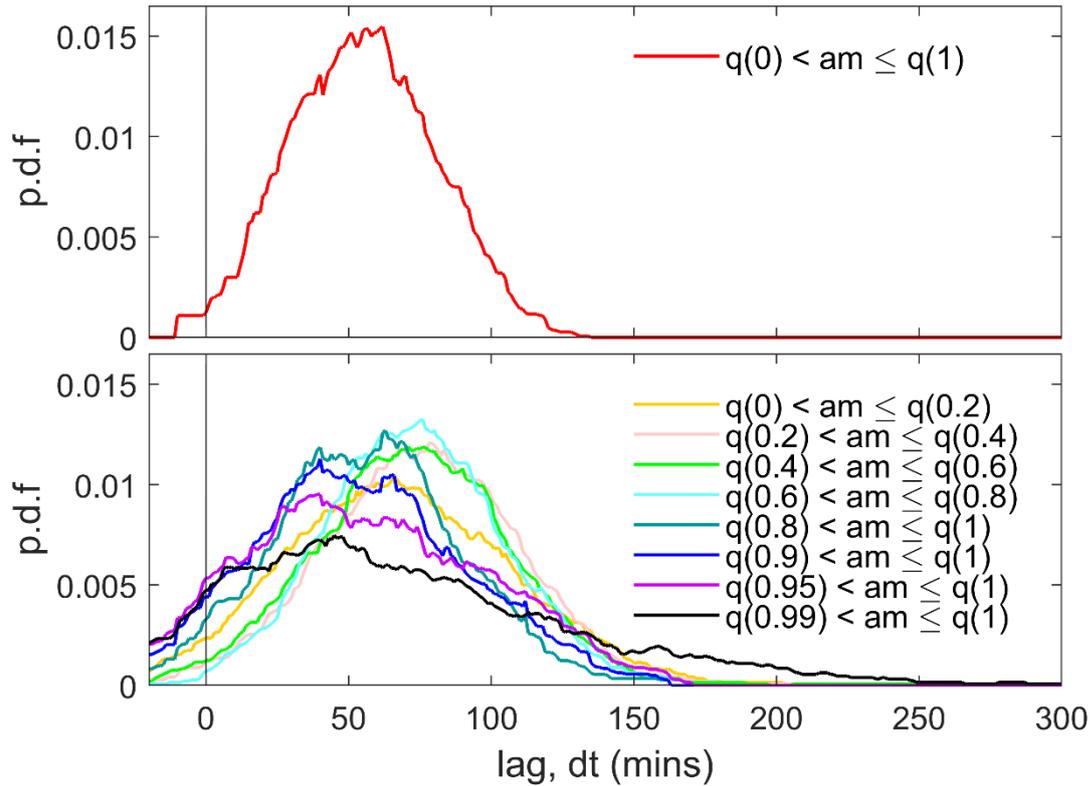


686 **Figure 3.** Linear correlation coefficients r between the am index and $P_{\alpha 1}$ (computed using
687 equations 7 and 8), colour contoured as a function of the fit parameters α (along the vertical
688 axis) and f_S (along the horizontal axis), for averaging timescale τ of (a) 1 day and (b) 1 year.
689 The black dot marks the peak correlation in (a) and the cyan diamond the peak for $f_S = 0$,
690 (for which $P_{\alpha 1} = P_{\alpha}$ as the solar wind Poynting flux is not included). The corresponding
691 points for annual mean data are shown by a cyan dot and a green diamond in (b). The values
692 of α and f_S at these points and the resulting correlation r are given at the bottom of each
693 panel.



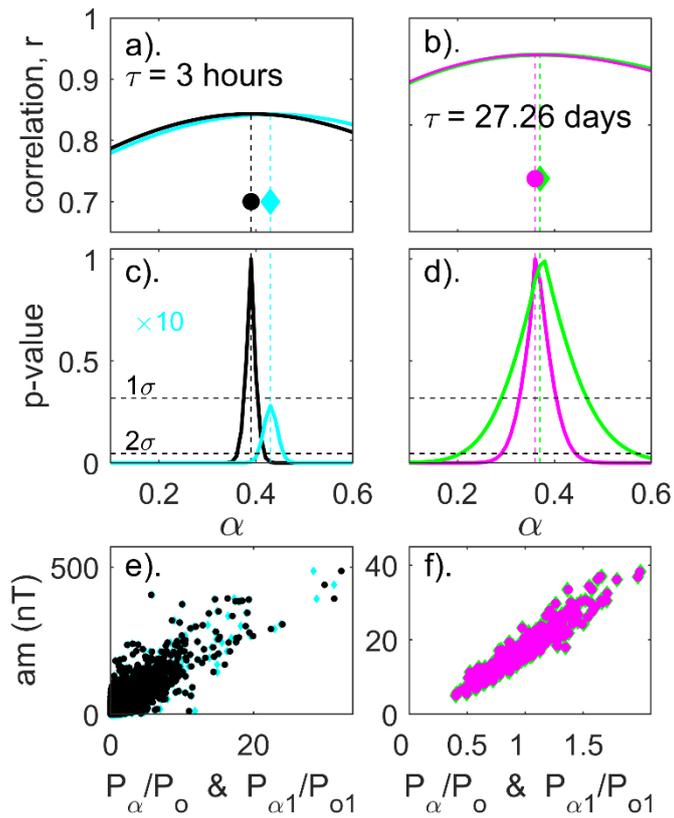
694

695 **Figure 4.** Comparison of the $r - \alpha$ correlograms for the optimum f_S and for $f_S = 0$. The left
696 hand panels are for 1-day means ($\tau = 1$ day) and the right hand panels are for annual means (τ
697 $= 1$ yr). Lines are coloured using the same colour scheme as the points in Figure 3. The top
698 panels show the linear correlation coefficient r as a function of the coupling exponent α : the
699 black line for $\tau = 1$ day is for the optimum f_S of 0.68, the cyan line is for $f_S = 0$; the mauve
700 line for $\tau = 1$ yr is for the optimum f_S of 0.30 and the green line is for $f_S = 0$. The vertical
701 dashed lines mark the peak correlations. The middle panels show the corresponding p -values
702 for the null hypothesis that the correlation is as good as for the peak value at the optimum f_S :
703 the 1σ and the 2σ levels are shown by horizontal dashed black lines. Note that the p value
704 for $\tau = 1$ day and $f_S = 0$ (the cyan line) has been multiplied by 100 and that even at the
705 optimum α , the correlation is significantly lower than the peak for $f_S = 0.68$ at more than the
706 3σ level. On the other hand, for $\tau = 1$ yr and $f_S = 0$ (the green line) the correlation is not
707 significantly lower at even the 1σ level at almost all values of α . The bottom panels shows
708 the scatter plots for the optimum α values for the optimum and zero f_S cases, using the same
709 colour scheme.



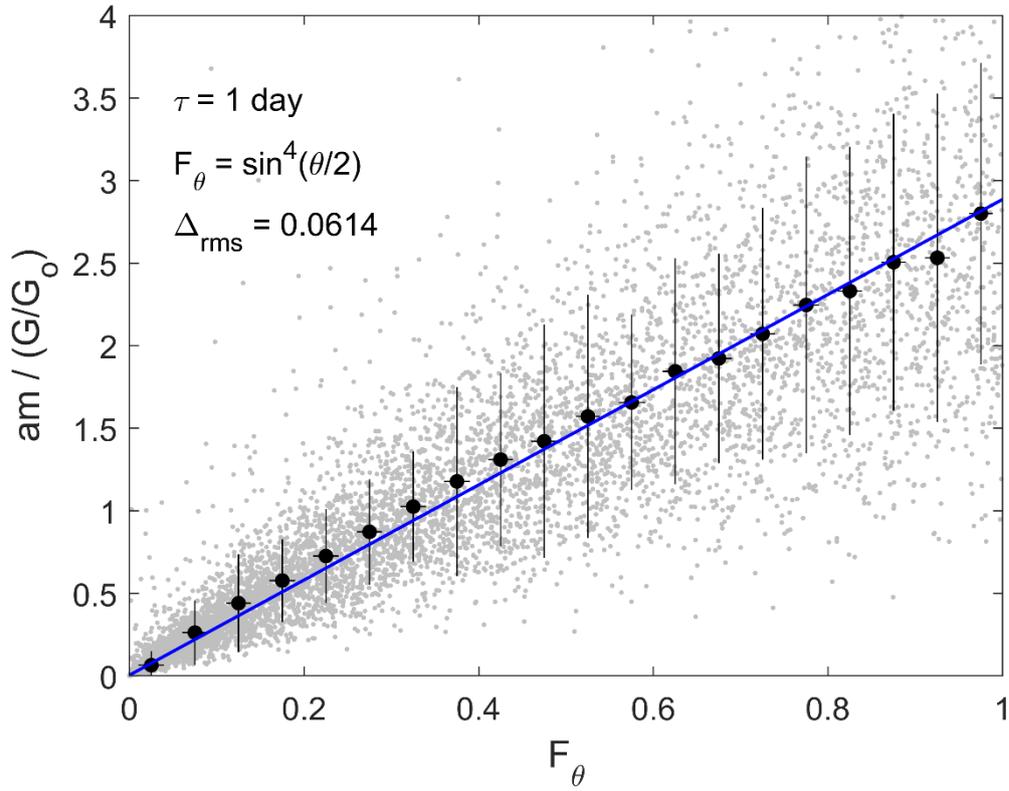
710

711 **Figure 5.** Analysis of response lags for the *am* index to power input to the magnetosphere
 712 for 1995-2017, inclusive. (Top). The 3-hourly *am* data have been sorted into 20 time-of year
 713 bins (*F*) and 8 UT bins. One-minute P_α data have then been averaged into 3-hourly intervals,
 714 centred on the mid-points of the *am* data intervals, minus a response lag *dt* that was varied
 715 between zero and 5 hours in steps of one minute. At each *dt* the correlation between $P_\alpha(t-dt)$
 716 and *am*(*t*) was evaluated and the red line shows the distribution of the *dt* giving peak
 717 correlation for the 160 *F*-UT bins. (Bottom) The same analysis but the *am* data have been
 718 further subdivided into 8 quantile ranges: the first 5 of these quantile ranges each contain
 719 20% of the data, for example the first is for *am* between zero and its 0.2 quantile, $q(0.2)$. The
 720 last three study the optimum lags for the largest *am* values, studying the top 10%, 5% and 1%
 721 of all *am* values.



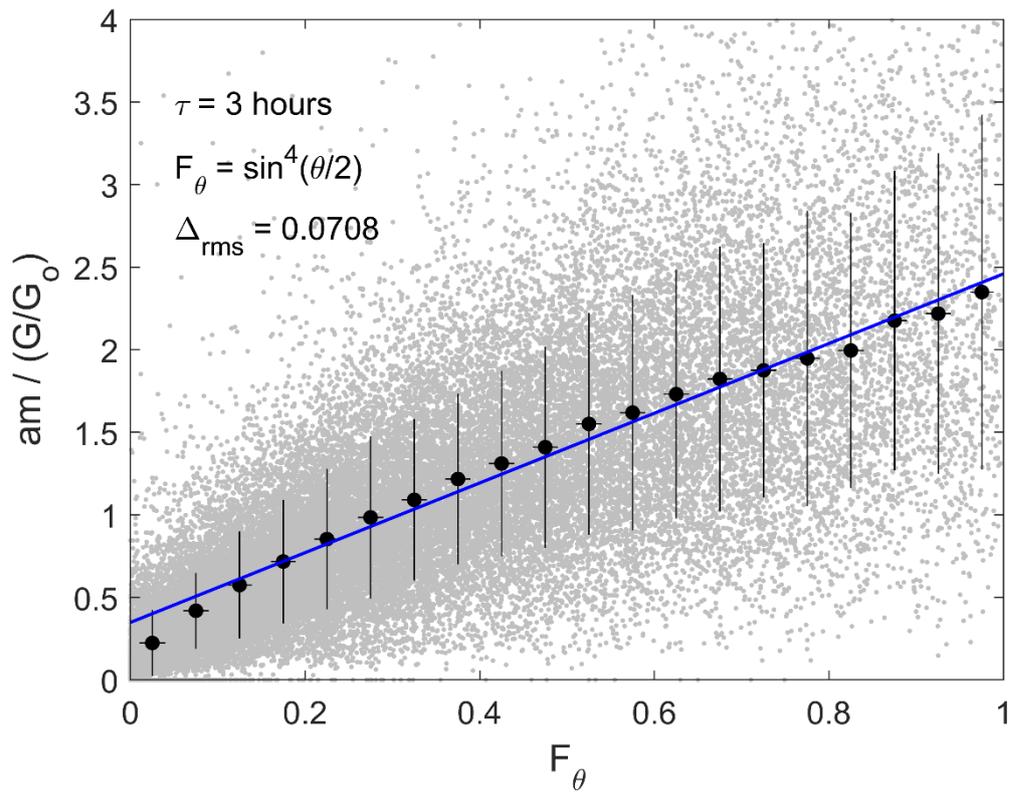
722

723 **Figure 6.** The same as Figure 4 but for (left hand panels) 3-hourly data and (right hand
 724 panels) averages over Carrington rotation intervals (27.26 days). For the 3-hourly timescale,
 725 the solar wind data are averaged over intervals a response delay of $dt = 60$ min before the
 726 three hours over which the am values are evaluated (see Figure 5).



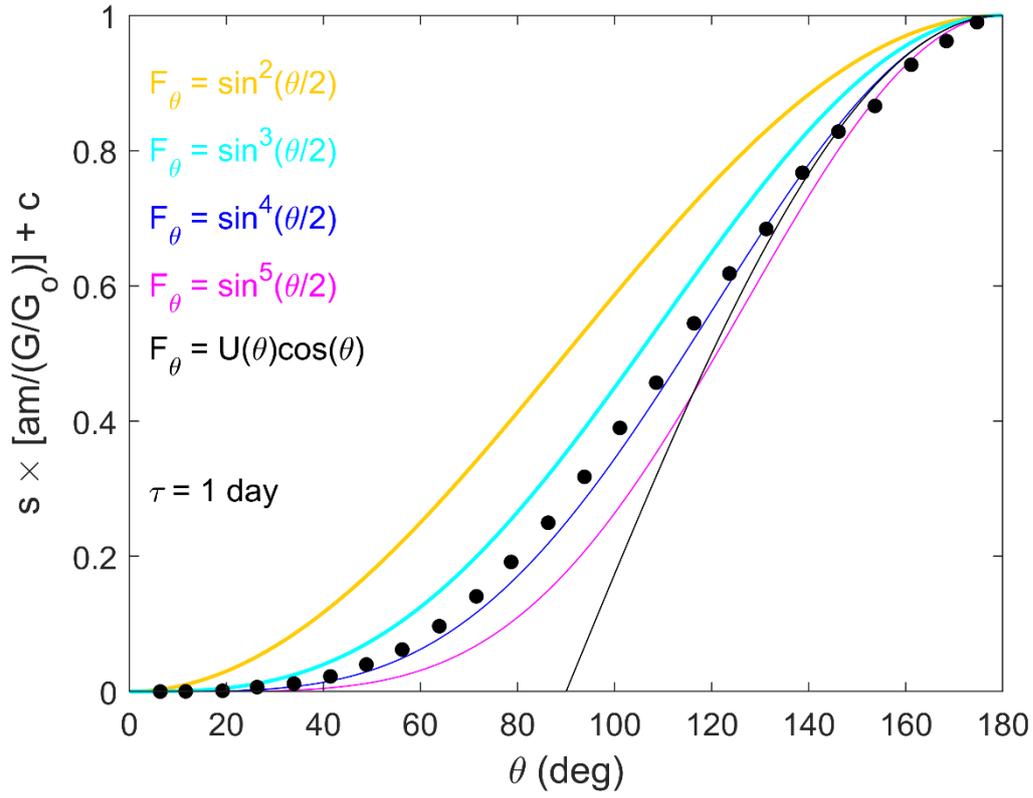
727

728 **Figure 7.** Test of the IMF orientation factor in the transfer function $F_\theta = \sin^4(\theta/2)$, where
 729 θ is the IMF clock angle, $\theta = \arctan(B_{YM}/B_{ZM})$, where B_{YM} and B_{ZM} are the Y and Z
 730 components of the IMF in the GSM frame of reference. These data are for averaging
 731 timescale $\tau = 1$ day. As shown by equations (11) and (12) of the text, the proportionality of
 732 the two confirms $\sin^4(\theta/2)$ is a good IMF orientation factor in the transfer function
 733 employed in $P_{\alpha 1}$. The grey dots are for daily means and the black dots are means for non-
 734 overlapping bins of F_θ of width 0.05 and both horizontal and vertical error bars are plus and
 735 minus one standard deviation. The blue line is the best-fit linear regression to the binned data.



736

737 **Figure 8.** The same as Figure 7 for an averaging timescale $\tau = 3$ hours and allowing for a 60-
 738 minute response lag of am to the solar wind power input data.



739 **Figure 9.** Comparison of the observed IMF orientation transfer factor with various forms
740 suggested and/or used in the literature. The black dots are the means of the observed θ and
741 the scaled ratio $\{am/(G/G_o)\}$ for $\tau = 1$ day and bins of θ that are 7.5° wide. The coloured
742 lines show 5 proposed analytic forms: (orange) $\sin^2(\theta/2)$; (cyan) $\sin^3(\theta/2)$; (blue)
743 $\sin^4(\theta/2)$; (mauve) $\sin^5(\theta/2)$ and $U(\theta)\cos(\theta)$ where $U(\theta) = 0$ for $\theta \leq 90^\circ$ and $U(\theta) = -1$
744 for $\theta > 90^\circ$.

745 Plain English Summary

746 Space weather is caused by energy extracted from the solar wind by the magnetosphere, the
747 volume of space surrounding the Earth that is dominated by Earth's magnetic field. That
748 energy arrives in two main forms in the solar wind: the kinetic energy of the particle flow and
749 an electromagnetic energy flux. The most successful predictors of space disturbances have
750 considered only the kinetic energy flux. The paper shows typically 10% of the power input
751 comes from the electromagnetic energy flux in the solar wind and allowing for this can make
752 a small, but significant, improvement to our ability to predict terrestrial space weather
753 disturbances.

754