The development of a space climatology: 2. the distribution of power input into the magnetosphere on a 3-hourly timescale

Article

Accepted Version

Final (Accepted Version) 21 October 2018

The development of a space climatology: 2. The distribution of power input into the magnetosphere on a 3-hourly timescale

Mike Lockwood[1], Sarah N. Bentley[1], Mathew J. Owens[1], Luke A. Barnard[1], Chris J. Scott[1], Clare E. Watt[1], Oliver Allanson[1] and Mervyn P. Freeman[2]

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

[2] British Antarctic Survey, High Cross, Madingley Road, Cambridge, CB3 0ET, UK.

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

Key Points:

- The normalized distribution of power input to the magnetosphere is set by IMF orientation variability via magnetopause reconnection rate
- 3-hourly normalized power input obeys a Weibull distribution with shape parameter k=1.0625 and scale parameter λ=1.0240 for all years
- Annual means can give the probability of space weather events in the largest 10% and 5% to within one-sigma errors of 10% and 12%, respectively

Abstract

Paper 1 in this series [*Lockwood et al.,* 2018b] showed that the power input into the magnetosphere P_α is an ideal coupling function for predicting geomagnetic "range" indices that are strongly dependent on the substorm current wedge and that the optimum coupling exponent α is 0.44 for all averaging timescales, τ, between 1 minute and 1 year. The present paper explores the implications of these results. It is shown that the form of the distribution of P_α at all averaging timescales τ is set by the IMF orientation factor via the nature of solar wind-magnetosphere coupling (due to magnetic reconnection in the dayside magnetopause) and that at

τ = 3hrs (the timescale of geomagnetic range indices) the normalized P_α (divided by its annual mean, i.e. $\langle P_\alpha \rangle_{\tau = 3\mathrm{hr}} / \langle P_\alpha \rangle_{\tau = 1\mathrm{yr}}$) follows a Weibull distribution with k of 1.0625 and λ of 1.0240. This applies to all years to a useful degree of accuracy. It is shown that exploiting the constancy of this distribution and using annual means to predict the full distribution gives the probability of space weather events in the largest 10% and 5% to within uncertainties of magnitude 10% and 12%, respectively, at the one-sigma level.

1. Introduction

In Paper I [*Lockwood et al.*, 2018b] in this series, it was established that the power input into the magnetosphere, P_α, computed from near-Earth interplanetary data using the physics-based formulation of *Vasyliunas et al.* [1982], is highly correlated with both the *am* geomagnetic index over a range of averaging timescales τ between a 3-hours and one year, with an optimum coupling function of $\alpha = 0.44$. In addition, the SME auroral index was used to show that this also applies down to τ = 1 min. (Note that allowance for response lag is required at these higher time resolutions to account for the effect of energy storage in the geomagnetic tail and its subsequent release during the substorm cycle). The averaging timescale employed is an important, but often overlooked, consideration in solar wind-magnetosphere coupling studies yet its effects on behaviour and conclusions can be considerable [*Finch and Lockwood*, 2007; *Badruddin and Aslam*, 2013]. In the current paper, we study the distribution of 3-hourly P_α values (τ = 3 hrs.) and investigate why it has the form that it does. The reasons for studying this distribution are associated with reconstructions of past space weather conditions (see sections 1.1 – 1.3 below), which exploit an important empirical result – namely that the annual distributions of values of various space weather parameters, X, averaged over an interval τ and divided by their annual mean, $\langle X \rangle_\tau / \langle X \rangle_{\tau=1\mathrm{yr}}$, are surprisingly constant over time [*Lockwood et al.*, 2017, 2018a]. This is an extremely valuable result, but one which would have greater predictive power (and in which we could have greater confidence) if we understood why it applies and what its limitations are. In Paper 3 of this series [*Lockwood et al*., 2018c], we study the evolution of the distribution of P_α with τ from the 3 hrs. studied here up to τ = 1 year. Together, these papers supply much of the understanding of the empirical result that we are searching for.

1.1 Space Climate: reconstructions of annual means of space weather parameters

Recent years have seen the development of reconstructions of past annual mean conditions in near-Earth space. These have been made from historic solar and geomagnetic observations, interpreted using understanding derived from modern measurements made by spacecraft and solar magnetographs. Initially these reconstructions employed single or multiple regression fits of co-incident data, but they have subsequently grown more complex and now also employ physical understanding and model simulations and have been checked using independent datasets, such as observed abundances of cosmogenic isotopes found in terrestrial reservoirs. The first attempt to reconstruct the interplanetary conditions of the past was made by *Feynman and Crooker* [1978] who used the geomagnetic *aa* index, which extends back to 1868. This index is based on the range of variation in the horizontal component of the geomagnetic field in 3-hour windows (as introduced by *Bartels* [1939]) and has, like all such "range" indices, an approximately square-law dependence on the speed of the solar wind impinging on Earth, V_{SW} [see *Lockwood*, 2013]. However, on annual timescales, *aa* also depends on the near-Earth IMF field strength, B, changes in which therefore also contribute to its long-term drift. *Feynman and Crooker* considered various combination scenarios of B and V_{sw}, including assuming that B was constant, in order to derive a long-term variation in V_{SW}. The first separation of these two factors was made by *Lockwood et al.* [1999] who used the relationship between the 27-day recurrence of *aa* and the annual mean V_{SW} to remove the dependence on V_{SW}. Rather than computing the near-Earth IMF B, *Lockwood et al.* evaluated the open solar flux (OSF, a global parameter, being the total magnetic flux leaving the top of the solar corona, whereas B is a local parameter as it only applies to the near-Earth heliosphere). In order to achieve this, these authors used the Ulysses result that the radial component of the heliospheric field is largely independent of heliographic latitude [*Smith and Balogh*, 1995; *Lockwood et al.,* 2004; *Owens et al.,* 2008]. *Solanki et al.* [2000] reproduced the OSF variation deduced by *Lockwood et al.* using the global OSF continuity equation, with sunspot number quantifying the global OSF production rate and with a constant fractional loss rate. This model has subsequently been developed, refined, and used many times [*Solanki et al*., 2002; *Schrijver et al.,* 2002; *Lean et al.*, 2002; *Wang and Sheeley*, 2002; *Mackay et al.,* 2002; *Mackay and Lockwood,* 2002; *Usoskin et al.*, 2002; *Lockwood*, 2003; *Wang et al.*, 2002; 2005; *Vieira and Solanki*, 2010, *Steinhilber et al.,* 2010; *Demetrescu et* *al.,* 2010; *Owens et al.*, 2011; *Owens and Lockwood*, 2012; *Goezler et al.,* 2013; *Lockwood and Owens*, 2014a;b ; *Wang and Sheeley*, 2013; *Karoff et al.*, 2015; *Rahmanifard et al*., 2017; *Asvestari et al.,* 2017]. The continuity model allows us to reconstruct the annual mean OSF variation using sunspot number as a proxy for the OSF emergence rate. Hence the OSF variation depends on the integral of the sunspot number and will only be influenced by relatively long-lived differences between the sunspot series employed. Over the interval for which we have reliable and homogeneous geomagnetic data (c. 1845 - present), almost identical results are obtained using the various sunspot number composites available, and all give good matches to the geomagnetic OSF reconstruction [*Lockwood et al.,* 2016a; *Owens et al.,* 2016a]. However, before 1845 the divergence of the various sunspot number reconstruction is greater and this does introduce changes to the derived OSF variation, particularly between the Maunder and Dalton minima (i.e. between about 1710 and 1790) [*Lockwood et al.,* 2016a; *Owens et al.,* 2016b].

The continuity model applies to OSF but has also been used to derive reconstructions of the near-Earth IMF, *B* [e.g. *Rahmanifard et al.,* 2017], which requires understanding of how OSF and *B* are related. It is often assumed, either explicitly or implicitly, that the two are linearly related [e.g., *Svalgaard and Cliver*, 2010]. In fact, proportionality is a much better assumption than linearity as it avoids the nonsensical possibility of a non-zero, near-Earth IMF *B* when its source, the OSF, is zero: assuming linearity yields a false "floor" minimum value to *B* (the intercept value). An assumption of proportionality was made in the analytic equations used in the first reconstruction of OSF by *Lockwood et al.* [1999] – however, this could be done only because the difference between the real OSF-*B* relation and the assumed proportional one was accounted for in the regressions that were then used to derive OSF from the data [*Lockwood and Owens*, 2011]. In general, there are two competing effects that make the OSF-*B* relationship more complex than either proportional or linear: for a given OSF, the near-Earth heliospheric magnetic field (hereafter called the interplanetary magnetic field, IMF) will decrease with increasing V_{SW} because of the unwinding of the Parker spiral. Secondly, as the mean V_{SW} increases its longitudinal structure also increases which enhances kinematic "folding" of open field lines, increasing *B* for a given OSF [*Lockwood et al.,* 2009a; b; *Lockwood and Owens*, 2009; *Owens et al., 2017*]. The resulting relationship of OSF and *B* has been studied by *Lockwood and Owens* [2011] and *Lockwood et al.* [2014a] and allows us to employ the continuity model, which can only apply to a global parameter such as OSF and not to a local one such as *B*, to model past variations of the near-Earth IMF *B* from sunspot numbers.

Svalgaard and Cliver [2005] developed their IDV geomagnetic index from Bartels' u-index [*Bartels*, 1932] and noted that it depended on *B*, with very little influence of V_{SW}. Indeed, several indices constructed from hourly mean geomagnetic data have this property, whereas range indices depend on both *B* and V_{SW} [*Lockwood*, 2013]. This is a very important result as it means that combinations of different indices can be used to derive both *B* and V_{SW}. The long-term variation of *B* that was derived by *Svalgaard and Cliver* [2005] was questioned by *Lockwood et al.* [2006] because their analysis employed non-robust regression procedures and also because it filled large data gaps in the observed IMF and solar wind speed time-series with interpolated values. (As demonstrated in Paper 1 [*Lockwood et al.*, 2018b], a much more reliable option is to mask out the geomagnetic data during data gaps when the interplanetary data are missing [*Finch and Lockwood*, 2007]). However, the insight provided by *Svalgaard and Cliver* is extremely valuable: *Rouillard et al.* [2007] used it in their reconstruction of both *B* and V_{SW}, and *Lockwood* [2014a] used 4 different pairings of different indices to derive both (as well as the OSF), with a full Monte-Carlo uncertainty analysis, back to 1845. Once the distinction between OSF and near-Earth IMF *B* is allowed for, there is a growing convergence between the different geomagnetic reconstructions of heliospheric parameters [*Lockwood and Owens*, 2011], and also with those from cosmogenic isotopes [*Asvestari and Usoskin*, 2016; *Asvestari et al*., 2017; *Owens et al.,* 2016b].

Svalgaard and Cliver [2010] extended the geomagnetic reconstructions back to 1835 using Bartels' work on diurnal variations. However, this results in a data series that is not homogeneous and *Lockwood et al.* [2014a] argue that geomagnetic reconstructions are only reliable for 1845 onwards. What is certain is that the start date cannot be before 1832, when Gauss introduced the first properly-calibrated magnetometer. To extend the series before the start of reliable geomagnetic data we have to employ the models based on sunspot number and the OSF continuity equation. These models can be run from the start of regular telescopic sunspot observations in 1612. *Lockwood and Owens* [2014a] extended the OSF modelling to compute the OSF in the both the streamer belt and in coronal holes and so computed the streamer belt width.

The results match well the streamer belt width derived from historic eclipse images [*Owens et al.*, 2017]. From this, and from modern magnetograph observations of the streamer belt width, *Lockwood and Owens* [2014b] made deductions about the annual solar wind speeds during the Maunder minimum. The reconstructed streamer belt width and OSF were used by *Owens et al.* [2017], in conjunction with 30 years' output from a data-constrained magnetohydrodynamic model of the solar corona based on magnetograph data, to reconstruct V_{SW}, B and solar wind number density, N_{SW} from sunspot observations. From these reconstructions, annual means of power input into the magnetosphere, P_{α}, have been computed by *Lockwood et al.* [2017].

1.2 The use of annual means in space climate reconstructions

There are a number of reasons why all of the reconstructions discussed in section 1.1 have been restricted to annual means. The first, but least compelling, reason is that the correlations exploited to make the reconstructions are higher for annual means than for data of higher time resolution. This is, at least in part, caused by the cancellation of observation noise in the annual means but there are also some systematic variations that are averaged out. For example, there is a seasonal variation in the ionospheric conductivities influencing any one geomagnetic observatory [*Wallis and Budzinski*, 1981; *Nagatsuma*, 2006; *Finch*, 2008; *Koyama et al*., 2014]. In the *aa* index, this effect is reduced by averaging data from two sites, one in each hemisphere, but better cancellation of seasonal effects is achieved by the *ap* index with its greater number of stations and the use of conversion tables that allow for season. However, there is still a remnant annual variation in *ap* because the sites are not distributed uniformly or equally in the two hemispheres [*Finch*, 2008] and the *am* index provides a much flatter time-of-day and time-of-year response pattern because of its more even geographical distribution of stations. Other systematic annual variations are introduced by the effects of Earth's dipole tilt and the variation of the Earth's heliographic latitude over the year (see *Lockwood et al*. [2016b], and references therein).

However, the fundamental limit that prevents the P_{α} reconstructions being of higher time resolution than annual is the importance of orientation of the near-Earth IMF in driving geomagnetic activity. This issue has been discussed by *Lockwood* [2013] and *Lockwood et al.* [2017b]. It is well known that on short timescales, because of the dominant role of magnetic reconnection, the coupled magnetosphere-ionosphere-thermosphere responds to the polarity and magnitude of the southward component of the IMF (in a suitable frame oriented with respect to the geomagnetic field axis, such as Geocentric Solar Magnetospheric, GSM). As discussed in Paper 1 [*Lockwood et al.,* 2018b], there are two time constants of response. The first is the directly-driven system which responds on a timescale of order a few minutes. The other response is the storage-unloading system, whereby the directly-driven flows store magnetic flux and energy in the magnetospheric tail which is released and deposited in the nightside auroral ionosphere and thermosphere via the substorm current wedge. This generates a second response after a delay of between about 30 and 60 min. The polarity of the southward field component rarely remains constant for more than about 1 hour [*Hapgood et al.*, 1981] and is always fluctuating under the influence of transient events such as coronal mass ejections, co-rotating interaction regions, smaller-scale stream-stream interactions, and turbulence (see review by *Lockwood et al.* [2016b] and references therein). There is very little historical evidence available that could be exploited further to improve reconstructions of timescales shorter than annual means. *Matthes et al.* [2017] have used the (extended) *aa* index to improve time resolution back to 1845 (with consideration of the known limitations of *aa*) and provide a set of plausible scenarios for the Dalton and Maunder minima which occurred before this date. Another potential source of daily information is auroral observations [*Legrand and Simon*, 1987; *Silverman*, 1992; *Kataoka et al.*, 2017]. However, there are severe complications introduced by: (1) the great differences between observing sites in the annual variations in hours of darkness and its effect on observation probability; (2) the effect of both secular drift in the Earth's field and of human migration on the numbers of people available to record sightings at latitudes where aurorae occur most frequently; (3) secular change in cloud cover at a given site; (4) the social factors that make recording of sightings fashionable and accurate; (6) subsequent loss of data through catastrophic events such as fires and wars; and (6) the increased use of street lighting in centers of population [*Lockwood and Barnard*, 2015]. Alternatively, and only after a great deal of further research, it may become possible to also use modelling of the solar corona, and its extension into the heliosphere, based on daily sunspot numbers; however, such applications remain in the future.

None of these possibilities are viable at the present time and so there is no source of historic information on IMF orientation at sub-annual times that can be applied back to the Maunder minimum. Hence the interplanetary time series, and their terrestrial space weather responses, cannot be reconstructed. The only solution is to average out the fluctuations in IMF orientation, such that only a dependence on the IMF magnitude, B, remains [*Lockwood*, 2013]. Averaging over sufficiently long intervals reduces the IMF orientation factor to an approximately constant factor. *Lockwood et al.* [2017] show that employing a single, overall average value for an IMF orientation factor in GSM causes only a 4% error in annual means (as opposed to 10% error for 27-day means and a 42% error for 1-day means).

1.3 Space Climatology: reconstructions of distributions of space weather parameters

From the discussion in Section 1.2, it is apparent that we cannot, for the time being at least, construct a time series of data at sub-annual resolution to study the space weather conditions far enough into the past to cover grand minimum conditions. However, this does not mean that we cannot construct a space weather climatology, giving the probability of events exceeding a certain size, by reconstructing the Probability Distribution Functions (PDFs) of space weather parameters. In this area, a surprising and powerful new empirical result has recently emerged: the annual distributions of many space weather indices for a given averaging timescale, τ, as a ratio of its annual mean, (i.e., the PDFs of $\langle X \rangle_{\tau} / \langle X \rangle_{1yr}$ for a generic space weather index X) are remarkably constant for a given τ. The distributions are quite close to lognormal at all τ but the variance decreases with increasing τ (i.e. the distribution becomes more Gaussian-like). *Lockwood et al.* [2017] showed this result held for daily means (τ = 1 day) during the space age of the power input into the magnetosphere, P_{α}, and of the *ap* geomagnetic index. This is despite the fact that the relative contributions to geomagnetic activity of recurrent disturbances such as Corotating Interaction Regions (CIRs) and random events (such us impacts by Coronal Mass Ejections) varied considerably during this interval [*Holappa et al.*, 2014]. *Lockwood et al.* [2018a] have shown that this result also holds for all of the full *ap* index data sequence (i.e., for 1932-2016) and all years of the *aa* index data (for 1868-2016).

Figure 1 stresses how ubiquitous this result is for space weather indices. Because the distributions maintain an almost constant shape, the number of events in each year above a given fixed threshold show a monotonic variation with the average value for that year [see *Lockwood et al.*, 2017]. Figure 1 is for an example τ of 1 day, showing scatter plots of $f[X>Xo]$, the fraction of days for a given year in which the daily mean of a parameter X exceeds its 95 percentile ($X > Xo$, where Xo is computed from the whole dataset), as a function of the annual mean of that parameter $<X>$. Figure 1(a) is for the *ap* index using all the available data (for 1932-2016); 1(b) is for the *Dst* index (1957-2016); (c) the *AE* index (1968-2016); (d) the *AU* index (1968-2016); (e) the *AL* index (1968-2016); and (f) the power input into the magnetosphere, P_α, computed from interplanetary data for a coupling exponent $\alpha = 0.44$ (1996-2016, although some years are not used as data gaps are too numerous and too long, see Paper 1 and *Lockwood et al.* [2017]). In each case, an increase in the average disturbance level (which means increasingly negative in the cases of *AL* and *Dst*) is associated with an increase in the fraction of days with disturbance in the top 5% of the overall distribution for that parameter. The scatter is greatest for *Dst*, but very small for *AL*, but this finding is of value to the climatology of a wide range of terrestrial space weather disturbance indices. The mauve lines in each panel of Figure 1 are third-order polynomial fits to these data points, constrained to pass through the origin (so that $f_{\rm fit}[X>Xo] = 0$ when $<X>_{\tau=1\rm yr} = 0$). The Table in Part 5 of the Supporting Information gives the coefficients for these fits for each index and also the values of $\Delta_{\rm rms}$, the root-mean-square (r.m.s.) of the fractional fit residuals. These confirm that the *AL* index has the lowest scatter. In fact the rank order by $\Delta_{\rm rms}$ is very revealing and shows a dependence on the latitudinal difference of the observing stations from the auroral oval. If we consider that the origin of this behaviour is the power input into the magnetosphere, the close adherence to the relationship by *AL* is consistent with this index being a good indicator of power released from the geomagnetic tail lobes as part of the storage/release behavior. If this is indeed the case, the fact that *AU* agrees slightly less well indicates that the power input to the magnetosphere is a slightly less good predictor of the directly-driven current system. The *AE* index is midway between *AL* and *AU* in its behaviour, being the difference of the two ($AE = AU - AL$ where *AL* is negative). The next closest agreement is the *Ap* index, which is a planetary index recorded at middle latitudes that is very sensitive to the substorm current wedge and so well correlated with *AL* (see Supporting Information file attached to Paper 1). The agreement for the *Dst* index is still good but not as good for the other geomagnetic indices. Ideally, if the relationships shown in Figure 1 all arose from the power input to the magnetosphere, then the relationship for P_α/P_o would be stronger than for all the geomagnetic indices. However, the scatter is greater for P_α/P_o than for any of the geomagnetic indices except *Dst*. We have repeated the analysis for G_α/G_o where $G_\alpha = P_\alpha/\sin^4(\theta/2)$, and hence is the power input without the IMF orientation factor, and G_o is the overall average of G_α. Note that whereas the geomagnetic indices have availability of essentially 100%, that of G_α is 96% and that of P_α is 86% for daily means (because, as described in Paper 1, we require just 9 samples in an hour to give an error below 5% for all the parameters used to compute G_α, whereas for the IMF orientation factor the same error requires 50 samples in an hour). Interestingly, Δ_{rms} is considerably smaller for G_α/G_o than for P_α/P_o and so much of the scatter for P_α/P_o is introduced by the IMF orientation factor. This may be associated with the limitations of the IMF orientation factor used, but it seems likely that data gaps also contributed to the additional scatter for P_α/P_o. What does seem to be clear is that the scatter gets increasingly larger for geomagnetic indices which are influenced by currents other than the nightside auroral electrojet because they employ stations that are further away from it.

We stress here that although the bulk (or "core") of the PDFs are usually best fitted by something like a lognormal distribution [e.g., *Riley and Love*, 2017], the extreme tail of the distribution is not generally well described by the core distribution and so the result will not, in general, hold for the number of the most extreme events [*Redner*, 1990]. In studies of extreme events using "Extreme Value Statistics" (EVS), a lognormal distribution has often been combined with a differently-shaped tail [e.g. *Vörös et al.*, 2015; *Riley and Love*, 2017]. Hence, although the use of this result can tell us about the occurrence of "large" events (in the top 5%), we should not expect it to hold well for the most extreme events. The relationship of large storms in the tail of the core distribution to extreme-event "superstorms" is discussed further in Paper 3 [*Lockwood et al.*, 2018c].

In the present paper, we exploit a number of findings that were presented in Paper 1 [*Lockwood et al.*, 2018a], namely: (1) that statistical studies of solar wind -magnetosphere coupling and coupling functions that employ data from before 1995 are unreliable and likely to be seriously in error because of the presence of more and longer gaps in the interplanetary data series; (2) the coupling exponent determining the power input to the magnetosphere, α, shows no significant variations with averaging timescale, τ, and the optimum value is $\alpha = 0.44$ at all τ studied (which was varied between 1 minute and 1 year); (3). annual values of power input to the magnetosphere P_α derived from combining annual means of the component interplanetary factors (the "average-then-combine" method) are not exactly the same as annual means of P_α that are computed at high time resolution and then averaged (the "combine-then-average" method); however, they are a usable approximation to within an error of about 5%; and (4) the uncertainty in the α estimate influences the magnitude of the average power into the magnetosphere , P_o , but has negligible effect on the waveform of the variation in P_α and hence on the ratios $P_\alpha/\langle P_\alpha\rangle_{\tau=1yr}$. The last point, (4), comes from further consideration of Figure 7 of Paper 1. Part (b) of that figure shows that the estimate of the average power into the magnetosphere, P_o, rises hyperbolically with α such that the maximum range of fitted α (0.40-0.48) causes a variation in P_o between 0.3×10^{19}W and 0.6×10^{19}W. However, part (d) of that figure shows the distributions of P_α/P_o are very similar for this range of α, all being lognormal in form. This is quantified in part (c) of the figure which plots the ratios of P_α/P_o to the values for the optimum $\alpha = 0.44$. This weak dependence of P_α/P_o on the precise values of α around the optimum value is also reflected in the flat-topped nature of the correlograms shown in Figures 4a and 5a of Paper 1. Thus, although the estimate of the absolute level of power input to the magnetosphere (averaging P_o for all data and $\langle P_\alpha\rangle_{\tau=1yr}$ tor annual means) depends strongly on the value of α, the waveform of the variation in P_α (that is tested by correlation studies) is only weakly dependent on α in the uncertainty range around the optimum value.

2. Analysis of the contributions to the magnetospheric Power input

The derivation of the equation for the power input to the magnetosphere (given in Paper 1), is reprised in the Supporting Information file attached to this paper for completeness. This file also includes a review of why the IMF magnitude is used (B) instead of the component transverse to the sun-Earth line (B_T) and a confirmation that the best IMF orientation factor is $\sin^4(\theta_{GSM}/2)$, using 20 years' data of both 1-minute and 3-hour resolution and with many fewer, and much shorter, data gaps.

The result that the annual distribution of the normalized power input into the magnetosphere $<P_\alpha>_\tau/<P_\alpha>_{\tau=1yr}$ has an approximately constant, quasi-lognormal form is a purely empirical one. Figure 2 gives an initial indication of why it applies, by looking at the annual distributions of $R = \log_{10}(<X>_{\tau=1min}/<X>_{\tau=1yr})$ where X is one of the parameters of near-Earth space that contributes to P_α. Equation (6) of Paper 1 (equation (S7) in the Supporting Information) shows that relevant parameters are: the mean ion mass of the solar wind, m_{sw}; its number density, N_{sw}; its speed, V_{sw}; the strength of the IMF frozen-in to the solar wind flow, B; the clock angle that the IMF makes with the north in the GSM frame of reference, θ_{GSM} (defined by $\theta_{GSM} = \arctan(|B_{yM}|/B_{zM})$, where B_{yM} and B_{zM} are the Y and Z components of the IMF in the GSM frame); Earth's magnetic moment, M_E; and a constant k_3. We here group terms according to their exponent in the expression for P_α. If the ratio $(<X>_\tau/<X>_{\tau=1yr})$ is lognormally distributed, R will be normally distributed about a mode and mean value of zero. Parts (a), (c), (e) and (g) of Figure 2 show the annual distributions of R for 1996-2017 (inclusive) for 1-minute averages (τ = 1min) where X is, respectively: the IMF, B; the solar wind mass density, $m_{SW}N_{SW}$; the solar wind speed, V_{SW}; and the IMF orientation factor, $\sin^4(\theta_{GSM}/2)$. In each case, the vertical axis gives $N/1000$, where N is the number of 1-minute averaged samples in bins of R that are 0.01 wide. There are 11.13 million 1-minute samples for which all parameters in P_α are available out of a possible total of 12.10 million for this interval, an availability of 92.1%. All the plots show similar distributions in the different years. Those for B, $m_{SW}N_{SW}$, and V_{SW}, in parts (a), (c) and (e) do indeed reveal near-Gaussian forms (on the logarithmic scale, R). They are not exactly Gaussian: that for V_{SW} is slightly asymmetric and the peaks for $m_{SW}N_{SW}$ tend to be slightly below the ideal value of zero (however, as noted below, the weighting of the $m_{SW}N_{SW}$ factor in P_α is small). The corresponding right hand plots (b), (d) and (f) show the variations in the variances of these distributions in R, σ_R for each year (normalized to their overall means for all years, i.e. $\sigma_R/<\sigma_R>$). By definition, the mean of each of these variations is unity, shown by the horizontal black line in each plot, and the surrounding grey areas show plus and minus one standard deviation about this mean. These show the variance is constant from year to year to within 6.9% (at the 1-sigma level) for B, 7.9% for $m_{SW}N_{SW}$, and 11.2% for V_{SW}.

The distribution is quite different for the $\sin^4(\theta_{GSM}/2)$ factor shown in Figure 2g. The annual distributions of R in Figure 2g show that $\sin^4(\theta_{GSM}/2)$ is far from lognormal in form (note the very large number of samples at $R = -1$, corresponding to $\sin^4(\theta_{GSM}/2) = 0$: the peak N is always for the extreme bin plotted at $R = -1$ (which is for $-\infty \leq R < -0.99$). Note that $R = -\infty$ and $R = -0.99$ correspond to $(<X>_{\tau=1min} / <X>_{\tau=1yr})$ of 0 and 0.1036: given that the average $\sin^4(\theta_{GSM}/2)$ for all years is 0.355 to within about 5% [*Lockwood et al.,* 2017], this bin covers a range of $\sin^4(\theta_{GSM}/2)$ of just 0 to ≈0.036 and yet 21% of 1-minute samples lie in this small range of $\sin^4(\theta_{GSM}/2)$ (which is for northward IMF with θ_{GSM} less than about 51.8°). N varies between 57530 and 68225 for this $\sin^4(\theta_{GSM}/2)$ bin, depending on the year. However, Figure 2h shows the year-to-year variability is low for $\sin^4(\theta_{GSM}/2)$, with $\sigma_R/<\sigma_R>$ being constant to within 3.2% at the 1-sigma level. To understand the implications for the $P_\alpha / <P_\alpha>_{\tau=1yr}$ distribution we note that from the equation for P_α (equation (6) of Paper 1 and (S7) of the Supporting Information):

$$\log_{10}(P_\alpha / <P_\alpha>_{\tau=1yr}) = \log_{10}(P_\alpha) - \log_{10}(<P_\alpha>_{\tau=1yr})$$

$$= \log_{10}(k_3) + a\log_{10}(B) + b\log_{10}(m_{SW}N_{SW}) + c\log_{10}(V_{SW}) + d\log_{10}(\sin^4(\theta_{GSM}/2)) - \log_{10}(<P_\alpha>_{\tau=1yr}) \quad (1)$$

Where for the best-fit α of 0.44 found in Paper 1

$$a = 2\alpha = 0.88 \quad (2)$$

$$b = (2/3 - \alpha) = 0.227 \quad (3)$$

$$c = (7/3 - \alpha) = 1.893 \quad (4)$$

$$\text{and } d = 1 \quad (5)$$

Figure 9b of paper 1 shows that, to a good approximation (error ≈ ±5%), annual "average-then-combine" values of P_α are equal to the "combine-then-average" values, hence

$$<P_\alpha>_{\tau=1\text{yr}} \approx k_3\, (<B>_{\tau=1\text{yr}})^{2\alpha}\ (<m_{SW}N_{SW}>_{\tau=1\text{yr}})^{(2/3-\alpha)}\ (<V_{SW}>_{\tau=1\text{yr}})^{(7/3-\alpha)} <\sin^4(\theta_{GSM}/2)>_{\tau=1\text{yr}} \qquad (6)$$

Substituting for $\log_{10}(<P_\alpha>_{\tau=1\text{yr}})$ in (1) using (6) gives

$$\log_{10}(P_\alpha /<P_\alpha>_{\tau=1\text{yr}}) - a\log_{10}(B/<B>_{\tau=1\text{yr}}) + b\log_{10}(m_{sw}N_{sw}/<m_{sw}N_{sw}>_{\tau=1\text{yr}}) +$$
$$c\log_{10}(V_{SW}/<V_{SW}>_{\tau=1\text{yr}}) + d\log_{10}(\sin^4(\theta_{GSM}/2) /<\sin^4(\theta_{GSM}/2>_{\tau=1\text{yr}}) \qquad (7)$$

Equation (7) shows that the distribution of $\log_{10}(P_\alpha/<P_\alpha>_{\tau=1\text{yr}})$ is the weighted sum of those shown in Figure 2. The combined contribution of the terms in B, $m_{sw}N_{sw}$, and V_{sw} remains close to Gaussian (on the logarithmic scale of R), dominated by the distribution of V_{SW}. However, the corresponding distribution for $\sin^4(\theta_{GSM}/2)$ is very far from Gaussian. From equations (2) - (5) this last term has a weighting of $d/(a+b+c+d) = 1/4$. Thus the dependence of P_α on $\sin^4(\theta_{GSM}/2)$ perturbs the distribution of $P_\alpha /<P_\alpha>_{\tau=1\text{yr}}$ from the quasi-lognormal form that it would otherwise have had. However, the right hand panels of Figure 2 explain the small year-to-year variation in the shape of the distribution of $P_\alpha /<P_\alpha>_{\tau=1\text{yr}}$ because each parameter has a quite constant standard deviation of its R variation, i.e. the standard deviation of X is approximately proportional to the mean. It should be remembered that Figure 2 is for 1-minute averaged data, and it becomes important to understand the effect of the averaging timescale, τ. It is not, in itself of great importance or application in this paper that some of the parameters in P_α are quasi-lognormally distributed at high time resolution; however, it does make their evolution with τ more understandable. This is because on averaging over a larger τ, the Gaussian distributions in the logarithmic R parameter remain Gaussian and become narrower because of the central limit theorem [*Heyde*, 2006; *Fischer*, 2011]. As a result, the distributions of the X parameters remain lognormal but evolve in shape, becoming less asymmetric. However, shown by Figure 2 of *Lockwood et al.* [2017], the highly non-Gaussian distribution of $\sin^4(\theta_{GSM}/2)$, shown here in figure 2g, varies in a complex way as the averaging timescale, τ, is increased.

In order to analyze the behavior of the distribution of power input into the magnetosphere P_α with averaging timescale τ, we here break equation (6) in Paper 1 [*Lockwood et al.*, 2018b] into five terms

$$P_\alpha = (k_3 M_E^{2/3})\, F_B\, F_V\, F_N\, F_\theta \qquad (8)$$

where $F_B = B^{2\alpha}$, (9)

$$F_V = V_{sw}^{(7/3-\alpha)} \;, \qquad (10)$$

$$F_N = (m_{sw}\, N_{sw})^{(2/3-\alpha)} \;, \qquad (11)$$

and $F_\theta = \sin^4(\theta_{GSM}/2)$ (12)

We now analyze the variation of annual means of these terms and their distributions around those means. In each case, we take the distribution of 3-hourly means (τ = 3hr., the resolution of range geomagnetic indices, which is longer than the average substorm cycle duration so we are integrating over substorm cycles) as a ratio of the annual mean value. This lets us look at the contributions of the various terms, not only to the variation in annual means of P_α, but also to the distributions of $<P_\alpha>_{\tau=3hr} / <P_\alpha>_{1yr}$.

2.1 The effect of the IMF

Figure 3 analyses the behavior of the term in P_α that depends on the IMF magnitude B, F_B (equation 9). Paper 1 shows that 0.44±0.02 for τ = 3hrs gives the optimum agreement with the *am* index [*Lockwood et al.*, 2018b], the best estimate of F_B reduces to $B^{0.88}$. In Figure 3a, the annual distributions of 3-hourly values of F_B (normalized for convenience to its overall mean for 1995-2017 $[F_B]_o$) are shown as vertical slices and as a function of year along the horizontal axis. We use the criterion for a valid 3-houly mean established in Paper 1. The number N of the 61126 valid 3-hourly means of $F_B/[F_B]_o$ obtained during 1995-2017 (a data availability of 91.0%) is color-coded in bins of $F_B/[F_B]_o$ that are 0.01 wide. The back line gives the mean values of these distributions and displays a clear solar cycle variation, with larger values at sunspot maximum (around 2002 and 2014), as expected. The distribution for all years is shown in Figure 3b by the gray histogram which shows N/N_{max} as a function of $F_B/[F_B]_o$, where N_{max} is the peak value of N. *Lockwood et al.* [2017; 2018a] have shown that the annual distributions of $<P_\alpha>_{\tau=3hr} / <P_\alpha>_{\tau=1yr}$ are remarkably constant from year to year and Figure 3c investigates the corresponding contribution of the IMF term by showing the distributions of $F_B/<F_B>_{\tau=1yr}$ (where the means F_B are also taken over τ = 3hrs: note that, hereafter, values given without the average symbols and a τ value subscript are 3 hourly values), in the same format as Figure 3a and the number N is again counted in bins 0.01 wide. The black line shows the annual means which, by virtue of the normalization, are always unity. The distributions for the different years are very similar and the logarithm of their variances v is close to constant, as shown in Figure 3e. The distribution for all years is shown by the gray histogram in Figure 3d, where N is again normalized to its peak value. As expected from Figure 2a, this distribution is well matched by the best-fit (using least squares) log-normal distribution shown in mauve which has unity mean and a variance v of 0.120. The r.m.s. deviation of the fitted lognormal from the observed N/N_{max} distribution is $\delta_{B,logn}$ = 2.4×10^{-2}: we use this parameter to compare the quality of this fit to others presented in the subsequent sub-sections. Note that the largest values are not as well fitted, as tends to be the case for all the fits to the core distribution presented in this paper, indicating the need to use extreme value statistics to add an appropriate tail to the distribution.

Figures 4-7 are equivalent to Figure 3 for the other factors in the equation (8). Note that the y axis scales are the same for each panel within each figure but are not the same for all figures. We noted that some of the parameters showed largest deviations in 2003, a solar maximum year in which the major series of "Halloween" storms occurred during the interval 19th October – 7th November. The energetic particles associated with these storms themselves caused some data gaps, but as a test we removed the whole interval and found no detectable differences in Figures 3-7. Hence even the largest storms do not perturb the distributions shown.

2.2 The effect of the solar wind speed

Figure 4 is the same as Figure 3 but for the term in P_α that depends on the solar wind speed, F_V (equation 10). As for the IMF term, taking the distribution of 3-hourly values and dividing by the annual mean gives a near lognormal distribution that is very similar from year to year. The solar cycle variation in F_V has almost been removed by this normalization; however, there is some residual effect of the dominance of recurrent fast streams in the declining phase of the solar cycle when Earth intersects long-lived, fast solar-wind streams [*Cliver et al.,* 1996; *Tsurutani et al.*, 2006] emanating from coronal holes that have expanded to low heliospheric latitudes [*Wang et al.*, 1996] and rotating with the equatorial photosphere approximately every 25 days. They slightly raise the distribution variances in the declining phases (seen in Figure 4e, and also as the increased difference between the mode and mean values in Figure 4c, particularly around 2008).

The normalized distribution of $F_V/<F_V>_{\tau=1yr}$ for τ = 3hrs and all years is shown in Figure 4d by the grey histogram which has been fitted with log-normal distribution with a mean of unity and variance $v = 0.127$ (mauve line). This is a similar, but slightly higher, variance than for the IMF factor F_B. The RMS deviation of the fitted lognormal from the observed N/N_{max} distribution is $\delta_{V,logn} = 3.6\times10^{-2}$ which is a 50% larger than that for F_B.

2.3 The effect of the solar wind mass density

Figure 5 is the same as Figure 3 but for the term in P_α that depends on the solar wind mass density, F_N (equation 11). Figure 4a shows that there is a very slight solar cycle variation in the distributions and mean of F_N, but none for $F_N/<F_N>_{\tau=1yr}$. Note that Figure 5a shows an increase in F_N for 2017, but the distribution of $F_N/<F_N>_{\tau=1yr}$ in Figure 5c is the same as for previous years. The overall distribution of $F_N/<F_N>_{\tau=1yr}$ (Figure 4d) is much narrower than that for either $F_B/<F_B>_{\tau=1yr}$ or $F_V/<F_V>_{\tau=1yr}$ and has here been fitted with a lognormal of mean unity and variance $v = 0.009$ (mauve line). For such a low variance-to-mean ratio, the lognormal distribution is very close to Gaussian. The RMS deviation of the fitted lognormal from the observed N/N_{max} distribution is $\delta_{N,logn} = 2.6\times10^{-2}$ which is almost the same as that for F_B.

2.4 The effect of the IMF orientation

Figure 6 is the same as Figure 3 but for the term in P_α that depends on the IMF orientation, F_θ (equation 12). The shape of the distributions of both $F_\theta/[F_\theta]_o$ and $F_\theta/<F_\theta>_{\tau=1\mathrm{yr}}$ for this τ of 3 hours is not well described by any of the standard parameterizations. Figure 2 of *Lockwood et al.* [2017] shows that this distribution evolves from having a singular and large peak at zero for τ = 5 min, into a lognormal form as τ increases to $\approx$ 6 hrs., which then falls in variance v as τ further increases, becoming close to Gaussian for $\tau > 1$ day and a low-variance Gaussian tending to a delta function at unity as τ approaches 1 year. Figure 6a shows that there is a very slight variation in the distributions and means of $F_\theta/[F_\theta]_o$ but it does not follow the solar cycle and has almost completely been suppressed in $F_\theta/<F_\theta>_{\tau=1\mathrm{yr}}$ [*Stamper et al.,* 1999; *Lockwood,* 2003; *Lockwood et al*., 2017].

2.5 The resulting distribution of P_α

Figure 7 is the same as Figure 3 but for the combination of these terms, P_α (equation 7). Given that the normalized factors in P_α ($F_B/<F_B>_{\tau=1\mathrm{yr}}$, $F_V/<F_V>_{\tau=1\mathrm{yr}}$, $F_N/<F_N>_{\tau=1\mathrm{yr}}$, and $F_\theta/<F_\theta>_{\tau=1\mathrm{yr}}$) all show very little year-to-year variation it is not surprising that neither does $P_\alpha/<P_\alpha>_{\tau=1\mathrm{yr}}$. The overall distribution shown in Figure 7d is quite close to a lognormal (the mauve line is the best fit with mean 1 and variance v = 1.788). Lognormal distributions arise when factors described by Gaussian or lognormal distributions are multiplied together. In this case, given that $F_B/<F_B>_{\tau=1\mathrm{yr}}$, $F_V/<F_V>_{\tau=1\mathrm{yr}}$, and $F_N/<F_N>_{\tau=1\mathrm{yr}}$ are described by three lognormal distributions (the last of which is of such low variance it is essentially Gaussian), so $F_BF_VF_N/<F_BF_VF_N>_{\tau=1\mathrm{yr}}$ is a (higher variance) lognormal.

However, the normalized IMF orientation factor $F_\theta/<F_\theta>_{\tau=1\mathrm{yr}}$ at τ = 3hrs does not follow a lognormal distribution and this has a major influence on the shape of the $P_\alpha/<P_\alpha>_{\tau=1\mathrm{yr}}$ distribution. The RMS deviation of the fitted lognormal from the observed N/N_{max} distribution of P_α is $\delta_{P,logn} = 6.5\times10^{-2}$ which is roughly three times larger than that for F_B and F_N and twice that for F_V. Visual inspection of Figure 7d shows that the reason why this fit to the P_α distribution is less good is that the lognormal distribution cannot match both the long tail of the observed distribution and the low mode value, which suggests a Weibull distribution. The best-fit Weibull distribution is described by a shape factor, k, of 1.0625, which with a scale factor, λ, of 1.0240 gives the required mean of unity, and is shown by the blue line in Figure 7d. For this fit, the RMS deviation from the observed N/N_{max} distribution for P_α is $\delta_{P,Wb} = 6.4\times10^{-4}$ which is 1% of that for the lognormal distribution.

Hence we have established that the power input into the magnetosphere, normalized to its annual mean value, does not change greatly from year to year because the same is true for each of the terms that multiply together to form it. The shape of the overall distribution of P_α (at τ = 3hrs) is better fitted with a Weibull form than a lognormal form because of the influence of the IMF orientation factor F_θ. In the next section we study why, for τ = 3 hrs., the P_α distribution has the form shown in Figures 7c and 7d. In comparing the relative widths of the factors in P_α, notice that the y-axis scales in Figures 4 -7 are different and have been chosen to show relative differences visually, yet also not supress any small scale features. Note also that the constancy of the P_α distribution is not absolute but is a usable approximation (accuracies that are discussed in section 4). For example, we note that in Figure 4 there is an anomalous feature in the distribution of F_V in 2003 and in Figure 6 there is an anomalous feature in F_θ in the same year. Figure 7 shows that this does percolate through to an anomaly (albeit of smaller magnitude) in the distribution of P_α for this year.

3. The origins of the magnetospheric power input distribution

Figure 8 studies the evolution of the distributions of $P_\alpha/\langle P_\alpha\rangle_{\tau=1yr}$ and of the factors $F_B/\langle F_B\rangle_{\tau=1yr}$, $F_V/\langle F_V\rangle_{\tau=1yr}$, $F_N/\langle F_N\rangle_{\tau=1yr}$, and $F_\theta/\langle F_\theta\rangle_{\tau=1yr}$ with averaging timescale τ between 1 minute and 3 hours. In each panel, the probability density function is color-coded as a function of the normalized parameter (vertical axis) and averaging timescale τ (horizontal axis). Panels (b), (d) and (e) (for, respectively, $F_B/\langle F_B\rangle_{\tau=1yr}$, $F_V/\langle F_V\rangle_{\tau=1yr}$, and $F_N/\langle F_N\rangle_{\tau=1yr}$) show that the distributions of normalized terms in F_B, F_V and F_N hardly change at all between τ = 1 min. and

τ = 3 hrs., and so the plots shown in Figures 3d, 4d and 5d apply, to a good degree of approximation, to all timescales below 3hrs (at least down to the 1 min limit studied here). On the other hand, Figure 8a shows that the distribution of the normalized power input P_α does change considerably over this range of τ, and Figure 8c shows that this change for P_α in large part mirrors that for the IMF orientation factor F_θ. At τ = 1min., the distribution is dominated by a very large number of zero and near-zero F_θ samples and, because F_θ appears as a multiplicative term in Equation (8), this generates a very large number of zero and near-zero P_α samples. For both $P_\alpha / \langle P_\alpha \rangle_{\tau=1yr}$ and $F_\theta / \langle F_\theta \rangle_{\tau=1yr}$, the distributions evolve in accordance with the central limit theorem [*Heyde*, 2006; *Fischer*, 2011], as discussed in Paper 3 [*Lockwood et al.*, 2018c].

Figure 2g shows that for τ = 1 min. there is a secondary peak in the occurrence of values of $R = \log_{10}(\langle F_\theta \rangle_{1min} / \langle F_\theta \rangle_{1yr})$ around R of 0.45 associated with IMF orientations close to southward (explained below by Figure 9 and associated text). This peak is smaller in magnitude but broader than the corresponding one for R near -1 because of the $\sin^4(\theta_{GSM}/2)$ function used for F_θ. This feature is off-scale in Figure 8c which plots $\langle F_\theta \rangle_\tau / \langle F_\theta \rangle_{1yr}$ (i.e. on a linear scale rather than the logarithmic scale of R) as a function of τ. Rather than expand the scale in all panels of Figure 8 and lose important detail, in Figure S14 of Part 4 of the Supporting Information we repeat Figures 8a and Figure 8c on a y-axis doubled length and scale which enables us to see this feature and track its evolution with τ. The feature is seen in S14(b) at $\langle F_\theta \rangle_\tau / \langle F_\theta \rangle_{1yr} \approx 2.8$ and τ = 1min. As the averaging timescale is increased it disperses and moves towards average values for the same reasons that the large peak at $\langle F_\theta \rangle_{1min} / \langle F_\theta \rangle_{1yr} \approx 0$ disperses and moves towards average values, namely intervals of prolonged strongly southward and northward IMF become rarer as τ increases.

Hence the key to understanding the distribution for P_α at τ = 3 hr. is understanding the distribution of $F_\theta = \sin^4(\theta_{GSM}/2)$ at τ = 1 min. This investigated by Figure 9 which shows the distributions of one-minute averages of various IMF parameters. There are 10207789 valid 1minute samples of the IMF and its components obtained in the years 1996-2016 (inclusive) – an availability of 92.4%. Figure 9a shows the distribution for the IMF B_Y component in the GSM frame, $[B_Y]_{GSM}$; Figure 9b for the IMF B_Z component, $[B_Z]_{GSM}$; and Figure 9c for the ratio,

$[B_Y]_{GSM}$ /$[B_Z]_{GSM}$. The arctangent of this ratio is IMF clock angle in the GSM frame, θ_{GSM} = arctan($|[B_Y]_{GSM}|/[B_Z]_{GSM}$), the distribution of which is shown in Figure 9d. Figure 9d shows that IMF pointing due north ($[B_Y]_{GSM} = 0$, $[B_Z]_{GSM} > 0$, $\theta_{GSM} = 0$) is as common as IMF pointing due south ($[B_Y]_{GSM} = 0$, $[B_Z]_{GSM} < 0$, $\theta_{GSM} = 180°$), but IMF in the GSM equatorial plane ($[B_Z]_{GSM} = 0$, $\theta_{GSM} = 90°$) is twice as common. Figure 9e demonstrates what happens when the clock angle is divided by 2 and convolved with a sine function in $\sin(\theta_{GSM}/2)$: the directly northward case gives $\sin(\theta_{GSM}/2) = 0$, the directly southward IMF gives $\sin(\theta_{GSM}/2) = 1$, and $[B_Z]_{GSM} = 0$ gives $\sin(\theta_{GSM}/2) \approx 0.71$. Note that the distribution becomes less smooth than the distribution of θ_{GSM}, which is the combined effect of binning the data into equal-width bins of $\sin(\theta_{GSM}/2)$ and of $\sin(\theta_{GSM}/2)$ being a non-linear function of θ_{GSM}. What is not intuitive is what has happened to the occurrence frequency of these values. The distribution in Figure 9e is dominated by the shape of the sine function, the slope of which approaches 1 when $\theta_{GSM}/2 \rightarrow 0$ and approaches 0 when $(\theta_{GSM}/2) \rightarrow 90°$. This means that bins of equal width in $\sin(\theta_{GSM}/2)$ cover a smaller range of θ_{GSM} at $\theta_{GSM}/2 \rightarrow 0$ (and so contain fewer samples), whereas they cover a larger range of θ_{GSM} at $\theta_{GSM}/2 \rightarrow 90°$ (and so contain a greater number of samples). This effect is convolved with the distribution of samples with θ_{GSM}. This greatly reduces the number of samples with $\sin(\theta_{GSM}/2)$ near 0 (the quasi-northward IMF case) and greatly enhances the number of samples with $\sin(\theta_{GSM}/2)$ near 1 (the quasi-southward IMF case). This can be seen in Figure 9e. Figure 9f presents the distribution of $\sin^4(\theta_{GSM}/2)$ values. It can be seen that the peak near $\sin(\theta_{GSM}/2) = 0$ has been greatly enhanced whereas that near $\sin(\theta_{GSM}/2) = 1$ has been greatly diminished. The reason is that raising to the 4th power moves values (which are all less than unity) towards zero. The lowest bin of the histogram shown in figure 9f (for $\sin^4(\theta_{GSM}/2) < 0.02$) contains 18.94% of all valid samples. For $\sin^2(\theta_{GSM}/2)$ the two peaks are of roughly the same magnitude (6.2% of the samples are at $\sin^2(\theta_{GSM}/2) < 0.02$), and for $\sin^{8/3}(\theta_{GSM}/2)$ (as used by *Newell et al.* [2007]) the $\sin(\theta_{GSM}/2) \approx 0$ peak is greater than the $\sin(\theta_{GSM}/2) \approx 1$ peak, as in Figure 9f although to a lesser extent (10.5% of the samples are at $\sin^{8/3}(\theta_{GSM}/2) < 0.02$). An insight into this distribution of $\sin^4(\theta_{GSM}/2)$ is to compare it to an alternative IMF orientation factor that is often used, namely B_S/B, where the southward field $B_S = -[B_Z]_{GSM}$ when $[B_Z]_{GSM} < 0$ and $B_S = 0$ when $[B_Z]_{GSM} \geq 0$. This so-called "half-wave rectified" function means that all $[B_Z]_{GSM} > 0$ samples become zero in

B_S/B, and Figure 9b shows that this is true for half of the samples. Hence the distribution of B_S/B has an even larger peak at $\sin(\theta_{GSM}/2) \rightarrow 0$ (51.1% of samples are at B_S/B <0.02). The distribution shown in Figure 9f is that shown by the vertical slice at the left-hand edge of Figure 9c. It gives the distribution of $<P_\alpha>_\tau/<P_\alpha>_{\tau=1yr}$ for τ = 1min a form which, because of the Central Limit Theorem, evolves into the neo-Weibull distribution for $<P_\alpha>_\tau/<P_\alpha>_{\tau=1yr}$ at τ = 3yr, as shown in Figure 8a. Because the distribution of $<P_\alpha>_{\tau=1min}/<P_\alpha>_{\tau=1yr}$ is set by that for $<F_\theta>_{\tau=1min}/<F_\theta>_{\tau=1yr}$ (with its dominant occurrence of zero or near-zero values) it is, to a large degree, the nature of solar-wind magnetosphere coupling that the coupling function has to capture, which predominantly defines the form of the power input distribution at τ = 1 min. As illustrated by Figures 8a and 8c, this also defines the form of the distributions at longer averaging timescales such as τ = 3 hours. Hence the shape of the distribution is set by the large variability of F_θ on short timescales and although variations in F_N, F_V, and F_B influence the mean value of P_α (and hence the PDF at every P_α value) they have very little effect on the shape of the distribution.

4. Uncertainties caused by assuming the distribution of normalized power input is constant

As mentioned previously, the result that the distribution of normalized power input into the magnetosphere is almost stationary is a very useful one. It has been used by *Lockwood et al.* [2017, 2018a] to predict the distributions of power input to the magnetosphere and of geomagnetic indices over the past 400 year from the annual means of solar wind parameters reconstructed by *Owens et al.* [2017]. The analysis carried out in the present paper gives us an opportunity to assess the accuracy of such applications of this result.

The blue lines in Figure 10a shows the PDFs, *d,* of $<P_\alpha>_\tau/<P_\alpha>_{\tau=1yr}$ for τ = 3yr for the 21 individual years of the 1996-2016 period. (Note that, by definition, PDFs are normalized, the integral of each curve along the y-axis being unity). The black line is the mean and the orange area is between the mean plus and minus one standard deviation. Figure 10b shows the deviations from the mean, expressed as a percentage, $\delta_d = 100(d-<d>)/<d>$ and in the same formats as Figure 10a. The horizontal lines show the limits of the upper 1% , 5%, 10% and 20%

of the cumulative distribution function (CDF, see Figure 11). The 1-σ error in the PDF is below 11% for the lower 80% of the P_α /$<P_\alpha>_{\tau=1yr}$ values (the error being ±11% for the 20% threshold), but rises to ±14.5% for the 10% threshold, ±28% for the 5% threshold and ±57% for the 1%.

However, for space weather applications we are not as interested in the probability of a given $<P_\alpha>_\tau$ value as we are in the probability of $<P_\alpha>_\tau$ exceeding a certain threshold: in other words we are more interested in the CDFs, *c* , than the PDFs, *d*. The CDFs are shown in Figure 11a, using the same format as Figure 10a and the errors in the mean CDF, $\delta_c = 100(c-<c>)/<c>$ are shown in Figure 11b. In this case, the 1-σ uncertainty in predicting an event in the top 20% of all events is ±8.5%; in the top 10% of all events is ±10%; in the top 5% of all events is ±12%; and in the top 1% is ±40%.

2. Conclusions

We have studied why the power input into the magnetosphere, P_α (averaged over intervals of duration τ = 3 hours), follows the distribution that is does by looking at the component terms. We use the optimum coupling function $\alpha = 0.44$ which was shown in Paper 1 [*Lockwood et al.*, 2018b] to apply at all timescales between 1 minute and 1 year for the geomagnetic index, with the most uniform response, *am*.

The solar wind mass density factor introduces the smallest variability into the P_α distribution (the variance/mean ratio for the distribution this factor being 0.009). The factors depending on the IMF magnitude and on the solar wind speed follow quasi lognormal distributions of similar shape (the variance/mean ratios being 0.120 and 0.127, respectively). These factors all contribute to the shape of the P_α distribution, but the dominant one is the IMF orientation factor. We have shown how this arises from the nature of the optimum coupling functions and the role magnetic reconnection in the dayside magnetopause (the reconnection voltage being strongly dependent on the orientation of the IMF vector). The distributions of the total mass density factor, the IMF magnitude factor and the solar wind speed factor hardly change between an averaging timescale of 1 minute and 3 hours, whereas the IMF orientation factor distribution changes rapidly. At τ =

1 minute the distribution of the IMF orientation factor has a very large peak at near-zero values (see Figure 9f), which arises from the fact that for almost exactly half of all time the IMF points northward in the GSM frame (see Figure 9b) and so P_α is low. This peak is smoothed out as the averaging timescale as τ is increased (in accordance with the central limit theorem). As a result, the distribution of power input into the magnetosphere at any τ is set by the distribution of the IMF orientation factor at very high time resolution.

Given this great importance of the IMF orientation factor, it is sensible to check that we are using the best functional form in our analysis. A great many papers have deployed coupling functions using the form $\sin^n(\theta_{GSM})$, were θ_{GSM} is the IMF clock angle in the GSM frame, but the optimum exponent, n, has been estimated to be anything between zero and 6. The first coupling functions that allowed for IMF orientation were often referred to as "half wave rectifier" functions because they were set to zero the 50% of the time that the IMF had a northward component (see Figure 9b) (a reference to the signal processing effect of software and devices that pass only one polarity of a parameter of the input signal into the output signal) [*Burton et al.,* 1975; *Murayama, et al.,* 1980]. *Bargatze et al*. [1986] point out that in terms of IMF orientation studies using half-wave rectified B_{ZM} are using a factor of the form $U(\theta_{GSM})\cos(\theta_{GSM})$ where θ_{GSM} is the IMF clock angle in the GSM frame and $U(\theta_{GSM}) = -1$ when $\theta_{GSM} \geq 90^\circ$ and $U(\theta_{GSM}) = 0$ when $\theta_{GSM} < 90^\circ$. Because it is continuous in slope, and because it allows for the fact that low-latitude (between the cusps) magnetopause reconnection is not switched off whenever the IMF is northward [*Chandler et al.,* 1999], the $\sin^n(\theta_{GSM})$ function has generally been seen as preferable, from MHD magnetospheric modelling *Hu et al.* [2009] and *Fedder et al.* [2012] and found $n \approx 1$, but statistical estimates from observations vary from $n = 2$ [*Kan and Lee*, 1979; *Doyle and Burke*,1983; *Lyatsky et al.*, 2007; *Milan et al*., 2008], $n = 2.67$ [*Newell et al.*, 2007], $n = 4$ [*Perreault and Akasofu*, 1978; *Wygant et al.*, 1983; *Scurry and Russell*,1991; *Stamper et al.,* 1999], $n = 4.5$ [*Milan et al*., 2012], to $n = 6$ [*Temerin and Li*, 2006; *Boynton et al.,* 2011]. The wide range in estimated n values may be because these studies employ different indicators of terrestrial disturbance but most studies employ interplanetary data with large data and many data gaps which, as shown by Paper 1, introduce considerable noise. The Supporting Information file contains an analysis of 20 years' data of 1-minute auroral *SML* index values (the equivalent of *AL* from the very extensive SuperMAG network of magnetometers) and interplanetary data with few data gaps that are dealt with rigorously, as detailed in Paper 1. The results clearly confirm that $\sin^4(\theta_{GSM})$ is indeed the best IMF orientation factor for use in P_α.

We have shown that the distribution of power input into the magnetosphere (normalized to its annual mean value, i.e. of $<P_\alpha>_{\tau=3hr}/<P_\alpha>_{\tau=1yr}$) on an averaging timescale of τ = 3 hrs., is a Weibull distribution with $k = 1.0625$ and $\lambda = 1.0240$ (which yields the required mean of unity). All the factors, when normalized to their annual mean value, show annual distributions which vary only slightly from year to year. Hence the multiplicative product of these factors, the power input to the magnetosphere, also behaves this way.

We have studied the uncertainties inherent in using the fact that the normalized power input (and hence the geomagnetic activity indices that correlate highly with it) has a distribution of almost constant shape and variance. For the number of events in the largest 10% the one-sigma error is 10% and for events in the largest 5% the one-sigma error is 12%. Hence the probabilities given in the space climatological study by *Lockwood et al.* [2018a] (which were in the largest 5% and based on reconstructed annual means) have an uncertainty of 12%, which has to be convolved with the uncertainty in the reconstructed mean value. Moving to the more extreme events, we show that the uncertainty in using the constant shape distribution rises to 40% for the top 1% of events. This stresses the unsuitability of this approach for the most extreme events and the fact that the extreme tail of the distribution may show a different form and that this tail can vary in ways different to the bulk of the distribution [e.g. *Vörös et al.*, 2015]. Studies of extreme events in the tail of the distribution will be discussed in later papers, but here we study the bulk of the distribution and stress that the results, although useful for defining the occurrence of "large and extreme events" (for example in in the top 5% of the overall occurrence distribution), cannot be extended to cover the most extreme events without the use of Extreme Value Statistics (EVS).

In paper 3 of this series [*Lockwood et al.*, 2018s] we will study the how the distributions of power input into the magnetosphere and of geomagnetic indices continue to evolve with averaging timescale τ between 3 hours and 1 year. The reason this is of interest to the development of a space weather climatology is because several studies have shown that many geomagnetic storms are the response to the time-integrated solar wind forcing over an extended period [*Echer et al.,* 2008; *Turner et al.*, 2009; *Lockwood et al.,* 2016; *Mourenas et al.,* 2018] and also the time-integration of the geomagnetic activity response is important for space weather phenomena such as GICs (Geomagnetically Induced Currents) in systems like power grids [*Gaunt and Coetzee*, 2007; *Ramírez-Niño et al.*, 2016] and the growth of energetic particles that can be damaging or disruptive to spacecraft electronics [*Mourenas et al.,* 2018]. Using solar wind power input P_α as a metric, integrated forcing over an interval of duration τ is $\tau \times \langle P_\alpha \rangle_\tau$. However, we note that $\tau \times \langle P_\alpha \rangle_\tau$ (or the time integral of another form of coupling function) is unlikely to be a fully adequate predictor because preconditioning or multiple events may be factors [see discussion by *Lockwood et al.,* 2016] as may impulsive events, such as sudden increases in solar wind speed [*Balan et al*., 2017]. Furthermore, it is not yet clear what timescale τ is most relevant to a given phenomenon. Figure 6 of *Wygant et al.* [1983] is significant because it shows that it can take of order 10 hours following a northward turning to return transpolar voltage to its baselevel values, which implies 4 or 5 substorms are required to reduce excess open flux (i.e. energy stored in the tail) even though the IMF is northward. Periods of northward IMF of 10 hrs. duration or more are rare [*Hapgood et al*., 1991] and so it is likely that southward IMF will drive renewed energy storage in the tail before the magnetosphere has returned to a quiet state. *Kamide et al.* [1977] showed that although substorms were more common when the IMF pointed southward, they do occur during northward IMF if the polar cap was large (indicating large open flux and hence high energy storage in the tail) and *Lee et al.* [2010] show that substorms during northward IMF driven by stored tail energy can be as strong as events during southward IMF. The ability of the tail to accumulate stored energy means that longer periods of solar wind forcing have the potential drive extremely large events, even if the forcing is intermittent and bursty on shorter timescales. *Lockwood et al.* [2016] estimate that the relevant τ may be as large as 4-5 days.

Acknowledgments and Data

The authors are grateful to the staff of Space Physics Data Facility, NASA/Goddard Space Flight Centre, who prepared and made available the OMNI2 dataset used. The data were downloaded from http://omniweb.gsfc.nasa.gov/ow.html. The work presented in this paper is supported by STFC consolidated grant number ST/M000885/1 and the work of ML and MJO is also supported by the SWIGS NERC Directed Highlight Topic Grant number NE/P016928/1/ and OA by NERC Grant NE/P017274/1. SB is supported by an NERC PhD studentship.

References

Asvestari, E. I.G. Usoskin, G.A. Kovaltsov, M.J. Owens, N.A. Krivova, S. Rubinetti and C. Taricco (2017) Assessment of different sunspot number series using the cosmogenic isotope 44Ti in meteorites, *Mon. Not. R. Astron. Soc*., **467** (2), 1608-1613, doi: 10.1093/mnras/stx190

Asvestari, E., and I. G. Usoskin (2016) An empirical model of heliospheric cosmic ray modulation on long-term time scale, *J. Space Weather Space Clim*., **6**, A15, doi: 10.1051/swsc/2016011

Badruddin, and O.P.M. Aslam (2013) Study of the solar wind-magnetosphere coupling on different time scales, *Planet. Space Sci.,* **85**, 123-141, doi: 10.1016/j.pss.2013.06.006

Balan, N., Y. Ebihara, R. Skoug, K. Shiokawa, I. S. Batista, S. Tulasi Ram, Y. Omura, T. Nakamura, and M.-C. Fok (2017) A scheme for forecasting severe space weather. *Journal of Geophysical Research: Space Physics*, **122**, 2824-2835, doi: 10.1002/2016JA023853

Balogh, A., T.J. Beek, R.J. Forsyth, P.C. Hedgecock, R.J. Marquedant, E.J. Smith, D.J. Southwood, and B.T. Tsurutani, The magnetic field investigation on the Ulysses mission-Instrumentation and preliminary scientific results, Astron. and Astrophys Supplement Series, 92 (2), 221-223, 1992, ISSN: 0365-0138

Bargatze, L.F.B., R. L. McPherron, D. N. Baker (1986) Solar Wind-Magnetosphere Energy Input Functions, in *Solar Wind-Magnetosphere Coupling*. Eds, Y. Kamide, J.A. Slavin, Terra Scientific Publishing Company, Tokyo, Japan, and D. Reidel Publishing Company, Dordrecht, Holland, ISBN: 90-277-2303-6, pp. 101-109, doi: 10.1007/978-94-009-4722-1_7.

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

Boynton, R. J., M. A. Balikhin, S. A. Billings, H. L. Wei, and N. Ganushkina (2011),Using the NARMAX OLS-ERR algorithm to obtain the most influential coupling functions that affect the evolution of the magnetosphere, *J. Geophys. Res*., **116**, A05218, doi:10.1029/2010JA015505.

Burton, R. K., R. L. McPherron, and C. T. Russell (1975) An empirical relationship between interplanetary conditions and Dst, J. *Geophys. Res*., **80** (31), 4204–4214, doi: 10.1029/JA080i031p04204

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

Cliver, E.W., V. Boriakoff, and K. H. Bounar (1996) The 22-year cycle of geomagnetic activity, *J. Geophys. Res*., **101**, 27091–27109, doi: 10.1029/96ja02037

Demetrescu, C., V. Dobrica, and G. Maris (2010) On the long-term variability of the heliosphere-magnetosphere environment, *Adv. Space Res.*, **46** (10), 1299-1312, doi: 10.1016/j.asr.2010.06.032

Doyle, M. A., and W. J. Burke (1983) S3-2 measurements of the polar cap potential, *J. Geophys. Res*., **88** (11), 9125-9133, doi: 10.1029/ja088ia11p09125

Echer, E., W. D. Gonzalez, and B. T. Tsurutani (2008), Interplanetary conditions leading to superintense geomagnetic storms (Dst > 250 nT) during solar cycle 23, Geophys. Res. Lett., 35, L06S03, doi:10.1029/2007GL031755.

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

Fedder, J.A., C.M. Mobarry, J.G. Lyon (2012) Reconnection voltage as a function of IMF clock angle, *Geophys. Res. Lett*., **18** (6), 1047-1050, doi:10.1029/90GL02722

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

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

Finch, I.D., and M. Lockwood (2007), Solar wind-magnetosphere coupling functions on timescales of 1 day to 1 year, *Annales Geophys*., **25**, 495-506, doi: 10.5194/angeo-25495-2007

Finlay C.C., S. Maus, C.D. Beggan, T.N. Bondar, A. Chambodut, T.A. Chernova, A. Chulliat, V.P. Golovkov, B. Hamilton, M. Hamoudi, R. Holme, G. Hulot, W. Kuang, B. Langlais, V. Lesur, F.J. Lowes, H. Lühr, S. Macmillan, M. Mandea, S. McLean, C. Manoj, M. Menvielle, I. Michaelis, N. Olsen, J. Rauberg, M. Rother, T.J. Sabaka, A. Tangborn, L. Tøffner-Clausen, E. Thébault, A.W.P. Thomson, I. Wardinski, Z. Wei, T.I. Zvereva (2010) International Geomagnetic Reference Field: the eleventh generation, *Geophys. J. Int*., **183**, 1216–1230, doi: 10.1111/j.1365-246X.2010.04804.x.

Fischer, H. (2011) A History of the Central Limit Theorem, from Classical to Modern Probability Theory, book in series "*Sources and Studies in the History of Mathematics and Physical Sciences*", Springer, New York, doi:10.1007/978-0-387-87857-7

Gaunt, C.T. and G. Coetzee (2007) Transformer failures in regions incorrectly considered to have low GIC-risk, *in 2007 IEEE Lausanne Power Tech,* Lausanne, Switzerland, 1-5 July 2007, pp 807-812, doi: 10.1109/PCT.2007.4538419

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

Heyde, C. (2006) Central Limit Theorem, in *Encyclopedia of Actuarial Science*, 1, John Wiley and Sons, doi: 10.1002/9780470012505.tac019

Holappa, L., K. Mursula, and T. Asikainen (2014) A new method to estimate annual solar wind parameters and contributions of different solar wind structures to geomagnetic activity, *J. Geophys. Res. Space Physics*, **119**, 9407–9418, doi: 10.1002/2014JA020599

Hu, Y.Q., Z. Peng, C. Wang, J.R. Kan (2009) Magnetic merging line and reconnection voltage versus IMF clock angle: Results from global MHD simulations, *J. Geophys. Res.,* **114,** (8), A08220, doi:10.1029/2009JA014118

Kan, J. R. and Lee, L. C. (1979) Energy coupling function and solar wind-magnetosphere dynamo, *Geophys. Res. Lett.,* **6**, 577-580. doi:10.1029/GL006i007p00577

Karoff, C., F. Inceoglu, M. F. Knudsen, J. Olsen and A. Fogtmann-Schulz, The lost sunspot cycle: New support from ^{10}Be measurements, *Astron. and Astrophys*., **575**, A77, doi: 10.1051/0004-6361/201424927

Kamide, Y., P. D. Perreault, S-I. Akasofu, and J. D. Winningham (1977) Dependence of substorm occurrence probability on the interplanetary magnetic field and on the size of the auroral oval, *Journal of Geophysical Research* 82, (35), 5521-5528, doi: 10.1029/ja082i035p05521

Kataoka, R., H. Isobe, H. Hayakawa, H. Tamazawa, A.D. Kawamura, H. Miyahara, K. Iwahashi, K. Yamamoto, M. Takei, T. Terashima, H. Suzuki, Y. Fujiwara and T. Nakamura (2017), Historical space weather monitoring of prolonged aurora activities in Japan and in China, *Space Weather*, **15**, 392–402, doi:10.1002/2016SW001493.

Koyama, Y., A. Shinbori, Y. Tanaka, T. Hori, M. Nosé, and S. Oimatsu (2014) An Interactive Data Language software package to calculate ionospheric conductivity by using numerical models, *Computer Phys. Comms*.,**185** (12), 3398-3405, doi: 10.1016/j.cpc.2014.08.011

Lean, J., Y.-M. Wang, and N.R. Sheeley Jr. (2002) The effect of increasing solar activity on the Sun's total and open magnetic flux during multiple cycles: Implications for solar forcing of climate, *Geophys. Res. Lett*., **29** (24), 2224, doi: 10.1029/2002GL015880

Lee, D.-Y., K.-C. Choi, S. Ohtani, J. H. Lee, K. C. Kim, K. S. Park, and K.-H. Kim (2010) Can intense substorms occur under northward IMF conditions?, *J. Geophys. Res.,* **115**, A01211, doi:10.1029/2009JA014480.

Legrand, J.-P. and P.A. Simon (1987) Two hundred years of auroral activity (1780-1979), *Annales Geophys*., **A5**, 161-167.

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

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

Lockwood, M. and L. Barnard (2015) An arch in the UK: a new catalogue of auroral observations made in the British Isles and Ireland, *Astron. and Geophys*., **56** (4), 4.254.30, doi: 10.1093/astrogeo/atv132

Lockwood, M. and M. Owens (2009) The accuracy of using the Ulysses result of the spatial invariance of the radial heliospheric field to compute the open solar flux, *Ap. J.*, **701** (2), 964-973, doi: 10.1088/0004-637X/701/2/964, 2009

Lockwood, M. and M.J. Owens (2011) Centennial changes in the heliospheric magnetic field and open solar flux: the consensus view from geomagnetic data and cosmogenic isotopes and its implications, *J. Geophys. Res.*, **116**, A04109, doi: 10.1029/2010JA016220

Lockwood, M., and M.J. Owens (2014a) Centennial variations in sunspot number, open solar flux and streamer belt width: 3. Modelling, *J. Geophys. Res. Space Physics*, **119** (7), 5193–5209, doi: 10.1002/2014JA019973

Lockwood, M. and M.J. Owens (2014b) Implications of the recent low solar minimum for the solar wind during the Maunder minimum*, Astrophys. J. Lett*., 781:L7, doi: 10.1088/20418205/781/1/L7

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

Lockwood, R.B. Forsyth, A. Balogh, and D. J. McComas (2004) Open solar flux estimates from near-Earth measurements of the interplanetary magnetic field: comparison of the first two perihelion passes of the Ulysses spacecraft, *Annales Geophys*, **22**, 1395-1405, doi: 10.5194/angeo-22-1395-2004

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

Lockwood, M. Owens, and A.P. Rouillard (2009a) Excess Open Solar Magnetic Flux from Satellite Data: I. Analysis of the 3rd Perihelion Ulysses Pass, *J. Geophys. Res*., **114**, A11103, doi:10.1029/2009JA014449

Lockwood, M., M. Owens, and A. P. Rouillard (2009b) Excess open solar magnetic flux from satellite data: 2. A survey of kinematic effects, *J. Geophys. Res.,* 114, A11104, doi: 10.1029/2009JA014450.

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

Lockwood, M., M.J. Owens, E. L.A. Barnard, and I.G. Usoskin (2016a) An assessment of sunspot number data composites over 1845-2014, *Astrophys. J.*, **824**, 54, doi: 10.3847/0004-637X/824/1/54

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

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

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

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

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

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

Mackay, D.H. and M. Lockwood (2002) The Evolution of the Sun's Open magnetic Flux: II. Full solar cycle simulations, *Solar Phys.,* 209 (2), 287-309, doi: 10.1023/A:1021230604497

Mackay, D.H., E. R. Priest, and M. Lockwood (2002) The Evolution of the Sun's Open Magnetic Flux: I. A Single Bipole, *Solar Phys.,* **207** (2), 291-308, doi: 10.1023/A:1016249917230, 2002

Matthes, K., B. Funke, M.E.Andersson, L. Barnard, J. Beer, P. Charbonneau, M.A. Clilverd, T. Dudok de Wit, M. Haberreiter, A. Hendry, C.H. Jackman, M. Kretzschmar, T. Kruschke, M. Kunze, U. Langematz, D.R. Marsh, A. Maycock, S. Misios, C.J. Rodger, A.A. Scaife, A. Seppälä, M. Shangguan, M. Sinnhuber, K. Tourpali, I. Usoskin, M. van de Kamp, P.T. Verronen, and S.Versick (2017) Solar Forcing for CMIP6 (v3.1), *Geosci. Model Dev.*, 10, 2247-2302, doi: 10.5194/gmd-2016-91

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

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

Mourenas, D., A.V. Artemyev, and X.-J. Zhang (2018), Statistics of extreme time integrated geomagnetic activity, *Geophysical Research Letters*, **45**, 502-510, doi: 10.1002/2017GL076828

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

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

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

Owens, M. J., C.N. Arge, N.U. Crooker, N.A. Schwadron, and T.S. Horbury (2008), Estimating total heliospheric magnetic flux from single-point in situ measurements, *J. Geophys. Res.*, **113**, A12103, doi: 10.1029/2008JA013677

Owens, M. J., Lockwood, M., Riley, P., & Linker, J. (2017). Sunward strahl: A method to unambiguously determine open solar flux from in situ spacecraft measurements using suprathermal electron data. *Journal of Geophysical Research: Space Physics*, **122**, doi: 10.1002/2017JA024631

Owens, M.J. and M. Lockwood (2012) Cyclic loss of open solar flux since 1868: The link to heliospheric current sheet tilt and implications for the Maunder Minimum, *J. Geophys. Res.,* 117, A04102, doi:10.1029/2011JA017193.

Owens, M.J., E. Cliver, K. McCracken, J. Beer, L.A. Barnard, M. Lockwood, A.P. Rouillard, D. Passos, P. Riley, I.G. Usoskin, Y.-M. Wang (2016a) Near-Earth Heliospheric Magnetic Field Intensity since 1800. Part 1: Sunspot and Geomagnetic Reconstructions, *J. Geophys. Res*., **121** (7), 6048-6063, doi: 10.1002/2016JA022529

Owens, M.J., E. Cliver, K. McCracken, J. Beer, L.A. Barnard, M. Lockwood, A.P. Rouillard, D. Passos, P. Riley, I.G. Usoskin, Y.-M. Wang (2016b) Near-Earth Heliospheric Magnetic Field Intensity since 1800. Part 2: Cosmogenic Radionuclide Reconstructions, *J. Geophys. Res*., **121** (7), 6064-6074, doi: 10.1002/2016JA022550

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

Owens, M.J., N.U. Crooker, and M. Lockwood, How is open solar magnetic flux lost over the solar cycle?, J. Geophys. Res., 116, A04111, 2011, doi: 10.1029/2010JA016039, 2011

Perreault, P. and Akasofu, S. (1978), A study of geomagnetic storms. *Geophys. J. Roy. Astron. Soc.,* **54**, 547-573. doi:10.1111/j.1365-246X.1978.tb05494.x

Ramírez-Niño, J., C. Haro-Hernández, J. HéctorRodriguez-Rodriguez, R. Mijarez (2016) Core saturation effects of geomagnetic induced currents in power transformers, *Journal of Applied Research and Technology*, **14**, (2), 87-92, doi: 10.1016/j.jart.2016.04.003.

Rahmanifard F., N.A. Schwadron, C.W. Smith, K.G. McCracken, K.A. Duderstadt, N.Lugaz, and M.L. Goelzer (2017) Inferring the Heliospheric Magnetic Field Back through Maunder Minimum, *Astrophys. J.,* **837** (2), 165/1-165/14, doi: 10.3847/15384357/aa6191
Redner, S. (1990) Random multiplicative processes: an elementary tutorial, *Am. J. Phys*., **58** (3), 267-273, doi: 10.1119/1.16497

Riley, P., and J. J. Love (2017) Extreme geomagnetic storms: Probabilistic forecasts and their uncertainties, *Space Weather*, **15**, 53–64, doi: 10.1002/2016SW001470

Roelof, E.C. and D.G. Sibeck (1993) Magnetopause shape as a bivariate function of interplanetary magnetic field Bz and solar wind dynamic pressure. *Journal of Geophysical Research: Space Physics*, ***98*** (A12), pp.21421-21450, doi: 10.1029/93ja02362

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

Schrijver, C.J., M.L. DeRosa, and A.M. Title (2002) What is missing from our understanding of long-term solar and Heliospheric activity?, *Astrophys. J.*, **577**, 1006–1012, doi: 10.1086/342247

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

Silverman, S. M. (1992), Secular variation of the aurora for the past 500 years, *Rev. Geophys.*, **30** (4), 333–351, doi:10.1029/92RG01571.

Smith, E. J. and A. Balogh (1995) Ulysses observations of the radial magnetic field, *Geophys. Res. Lett*., **22**, 3317–3320, doi: 10.1029/95gl02826

Solanki, S. K., M. Schüssler, and M. Fligge (2000) Evolution of the Sun's large-scale magnetic field since the Maunder Minimum, *Nature*, **408**, 445–447, doi:10.1038/35044027.

Solanki, S.K., M. Schüssler, and M. Fligge (2002) Secular variation of the Sun's magnetic flux, *Astron. & Astrophys*., **383**, 706-712, doi: 10.1051/0004-6361:20011790

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

Steinhilber, F., J. A. Abreu, J. Beer, and K. G. McCracken (2010) Interplanetary magnetic field during the past 9300 years inferred from cosmogenic radionuclides, *J. Geophys. Res.,* **115**, A01104, doi: 10.1029/2009JA014193

Svalgaard L., and E.W. Cliver (2010) Heliospheric magnetic field 1835–2009. *J. Geophys. Res.,* **115** A09111, doi:10.1029/2009JA015069.

Svalgaard, L. and E.W. Cliver (2005) The IDV index: Its derivation and use in inferring longterm variations of the interplanetary magnetic field strength, *J. Geophys. Res*., **110**, A12103, 2005, doi: 10.1029/2005JA011203

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

Tsurutani, B.T., W.D. Gonzalez, A.L.C. Gonzalez, F.L. Guarnieri, N. Gopalswamy, M. Grande, Y. Kamide, Y. Kasahara, G. Lu, I. Mann, R. McPherron, F. Soraas, and V. Vasyliunas (2006), Corotating solar wind streams and recurrent geomagnetic activity: A review, *J. Geophys. Res*., **111**, A07S01, doi:10.1029/2005JA011273.

Turner, N. E., W. D. Cramer, S. K. Earles, and B. A. Emery (2009), Geoefficiency and energy partition ing in CIR-driven and CME-driven storms, *J. Atmos. Sol. Terr. Phys*., **71**, 1023–1031, doi:10.1016/j.jastp.2009.02.005

Usoskin, I. G., K. Mursula, S. K. Solanki, M. Schussler, and G. A. Kovaltsov (2002) A physical reconstruction of cosmic ray intensity since 1610, *J. Geophys. Res.,* **107** (A11), 1374, doi:10.1029/2002JA009343

Vasyliunas, V. M., J. R. Kan, G. L. Siscoe, and S.-I. Akasofu (1982) Scaling relations governing magnetospheric energy transfer, *Planet. Space Sci*., **30**, 359–365, doi: 10.1016/00320633(82)90041-1

Vieira, L.E.A., and S.K. Solanki, Evolution of the solar magnetic flux on time scales of years to millennia, Astron. Astrophys., 509, A100, 2010, doi: 10.1051/0004-6361/200913276.

Vörös, Z., M. Leitner, Y. Narita, G. Consolini, P. Kovács, A. Tóth, and J. Lichtenberger (2015) Probability density functions for the variable solar wind near the solar cycle minimum, *J. Geophys. Res. Space Physics*, **120**, 6152–6166, doi: 10.1002/2015JA021257.

Wallis, D. D., and E. E. Budzinski (1981), Empirical models of height integrated conductivities, *J. Geophys. Res*., **86** (A1), 125–137, doi:10.1029/JA086iA01p00125.

Wang Y.-M. and N. R. Sheeley, Jr. (2013) The solar wind and interplanetary field during very low amplitude sunspot cycles *Astrophys. J*., **764**, 90, doi:10.1088/0004-637X/764/1/90

Wang, Y.-M. and N.R. Sheeley Jr. (2002) Sunspot activity and the long-term variation of the Sun's open magnetic flux, *J. Geophys. Res.,* **107**, 1302, doi: 10.1029/2001JA000500

Wang, Y.-M., J. Lean, and N.R Sheeley Jr. (2002) Role of a variable meridional flow in the secular evolution of the sun's polar fields and open flux, *Astrophys. J.,* **577**, L53–L57, doi: 10.1086/344196

Wang, Y.-M., J. Lean, and N.R Sheeley Jr. (2005) Modeling the Sun's Magnetic Field and Irradiance since 1713, *Astrophys. J*., **625** (1), 522–538, doi: 10.1086/429689

Wang, Y.-M., S.H. Hawley, and N. R. Sheeley, Jr (1996) The magnetic nature of coronal holes, *Science*, **271**, 464–469, doi: 10.1126/science.271.5248.464

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

Figures

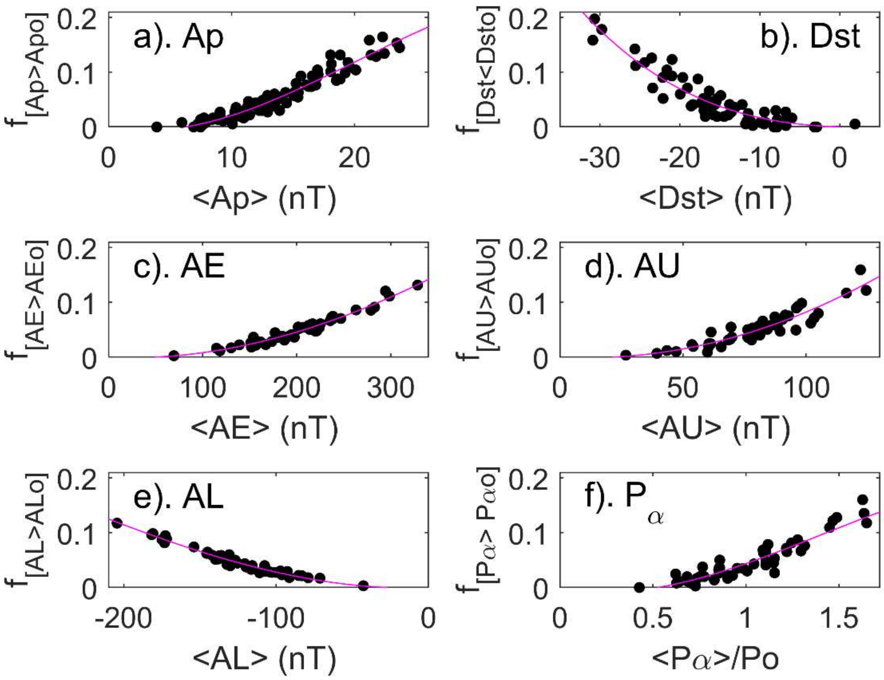


Figure 1. Scatter plots of $f[X>X\text{o}]$, the fraction of days in a given year for which the daily mean of a parameter X exceeds its 95 percentile Xo computed over the whole dataset, as a function of the annual mean of that parameter $<X>_{\tau=1\text{yr}}$. In each panel, the mauve line is a third-order polynomial fit to the data points, constrained to pass through the origin. (a) For the *Ap* index (*Ap* being daily means of *ap*, data available for 1932-2016), for which the 95 percentile is *Apo* = 38; (b) for the *Dst* index (data for 1957-2016), for which the 5-percentile *Dsto* = −53nT (note in this case, because *Dst* is increasingly negative as activity increases, *Dsto* is the 5-percentile and $f_{[\text{Dst}<\text{Dsto}]}$ is shown); (c) for the *AE* index (data for 1967-2016), for which the 95 percentile is *AEo* = 650nT; (d) for the *AU* index (data for 1967-2016), for which the 95 percentile is *AU*o = 228nT; (d) for the *AL* index (data for 1967-2016), for which the 5-percentile *ALo* = –444nT (note in this case, because *AL* is increasingly negative as activity increases, *ALo* is the 5-percentile and $f_{[\text{AL}<\text{ALo}]}$ is shown); and (e) the power input into the magnetosphere for a coupling exponent of α = 0.44, P_α (data for 1963-2016, although some years are omitted as data availability is too low – see Paper 1), for which the 95-percentile is $P_{\alpha o} = 2.73\ P_o$, where P_o is the mean P_α for all available data. In Paper 1 we derive an optimum value for P_o of $0.38{\times}10^{19}$ W (although note that, unlike P_α/P_o, this value is very sensitive on the derived coupling function, α) which yields an absolute estimate of the 95-percentile for the power input into the magnetosphere of $P_{\alpha o} = 1.04{\times}10^{19}$ W.

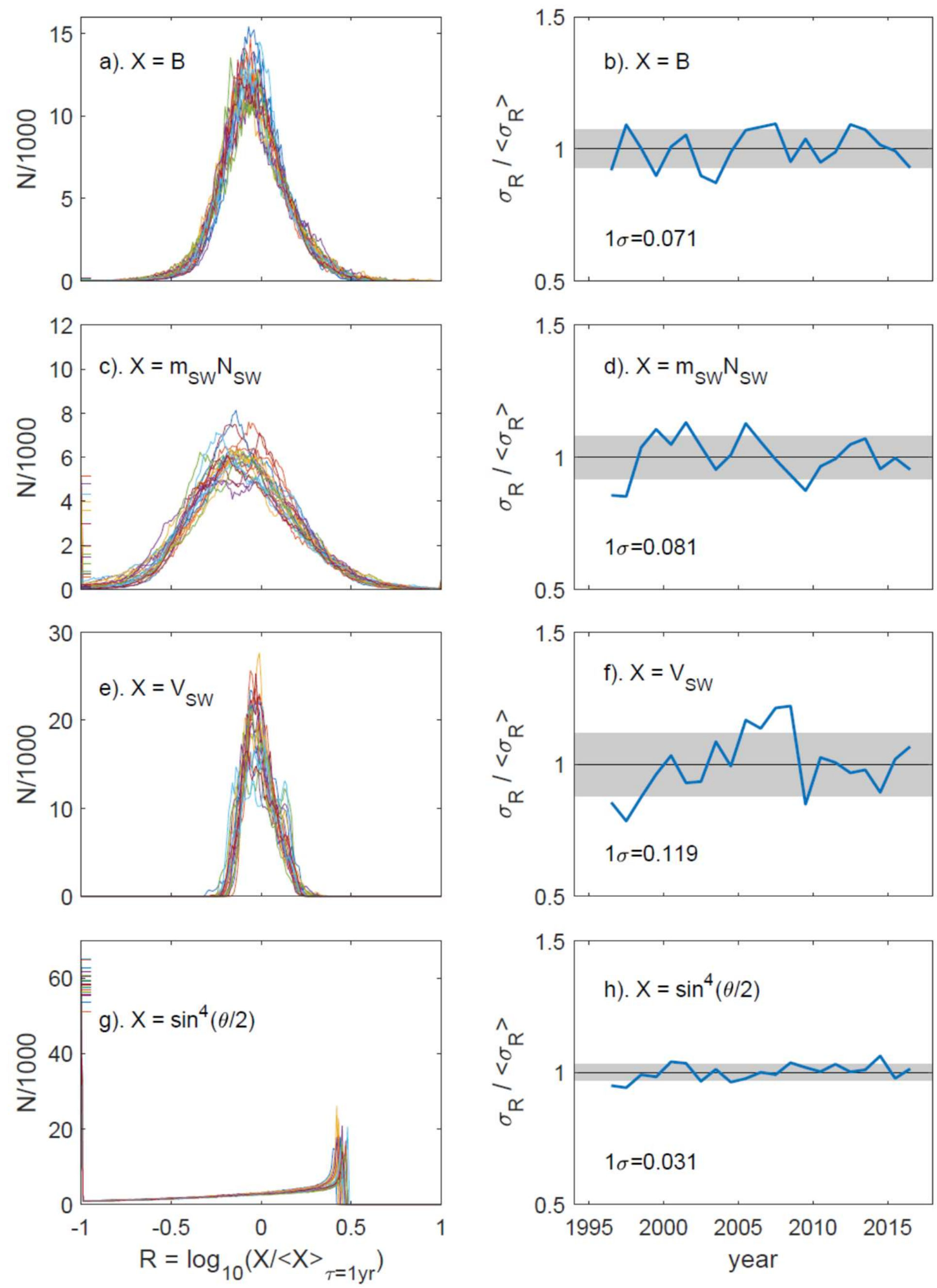


Figure 2. Analysis of annual distributions of the 11046240 1-minute averages of parameters contributing to the power input into the magnetosphere, P_α for 1996-2017 (inclusive). The left hand plots show 22 superposed annual distributions for individual years of $R = \log_{10}(X/\langle X\rangle_{\tau=1\mathrm{yr}})$ where X is (a) the IMF, B; (c) the solar wind

mass density, $m_{SW}N_{SW}$; (e) the solar wind speed, V_{SW}; and (g) the IMF orientation factor $\sin^4(\theta_{GSM}/2)$. $\langle X\rangle_{\tau=1yr}$ is the corresponding annual mean value in each case. Lognormal distributions in $X/\langle X\rangle_{\tau=1yr}$ would give Gaussian distributions in R, centered on zero. The vertical axis is $N/1000$, where N the number of 1-minute averaged samples in bins of R that are 0.01 wide. Note that in these left-hand plots the extreme bins are for $R \leq -0.99$ and $R \geq +0.99$ and the numbers of samples in these extreme bins are given for individual years by colored tick marks on the left and right (respectively) vertical axes of (a), (c), (e) and (g). There are negligibly few samples in the $R \geq 0.99$ bin for all four cases (7 for B, 1770 for $m_{SW}N_{SW}$, and none for either $\sin^4(\theta_{GSM}/2)$ or V_{SW}). The same is not always true for the $R \leq -0.99$ bin (for which, in total, there are 1597 samples for B, 36818 for $m_{SW}N_{SW}$ (0.3% of the total), 2351900 (21% of the total) for $\sin^4(\theta_{GSM}/2)$ and none for V_{SW}). In particular, the peak N in part (g) is always for this $R \leq -0.99$ bin and varies between 57530 and 68225, depending the year. Note that in many cases these coloured tick marks are indistinguishable from the x axis (at $N = 0$). The corresponding right hand plots (b), (d), (f) and (h) show the variations in the standard deviations of the distributions, σ_R for each year (normalized to their overall means for all years, i.e. $\sigma_R/\langle\sigma_R\rangle$). The horizontal black line in each plot gives the mean value (by definition unity), and the surrounding grey areas show plus and minus one standard deviation about this mean.

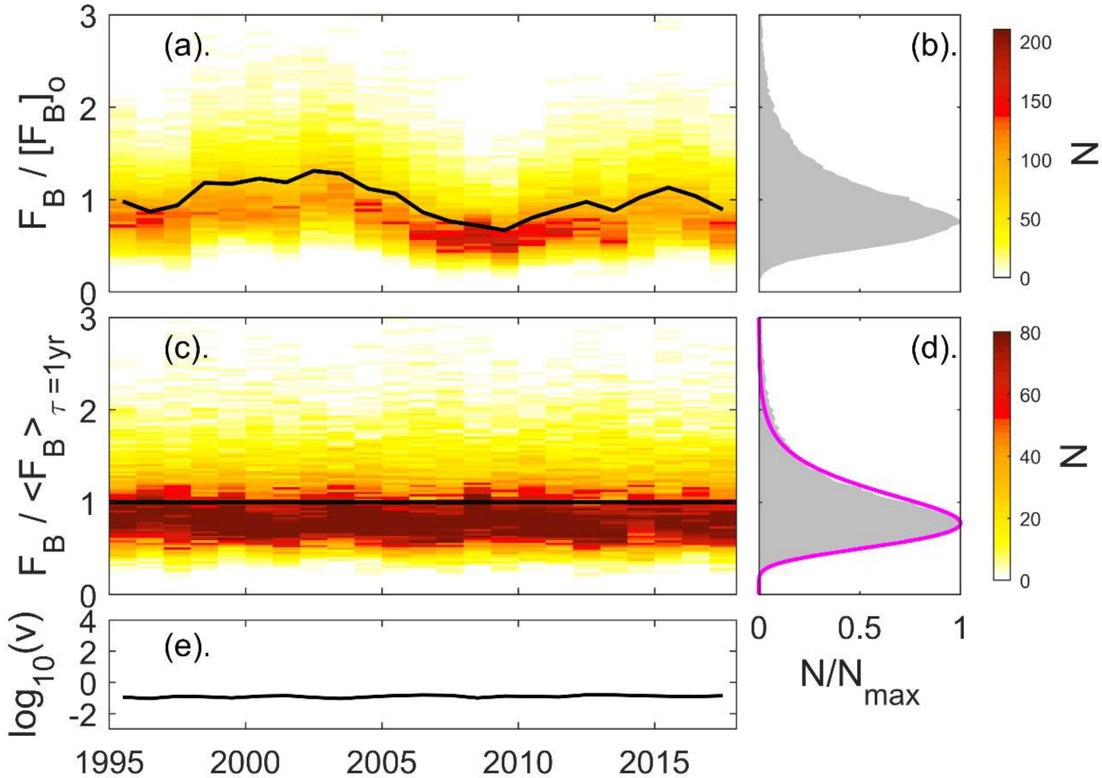
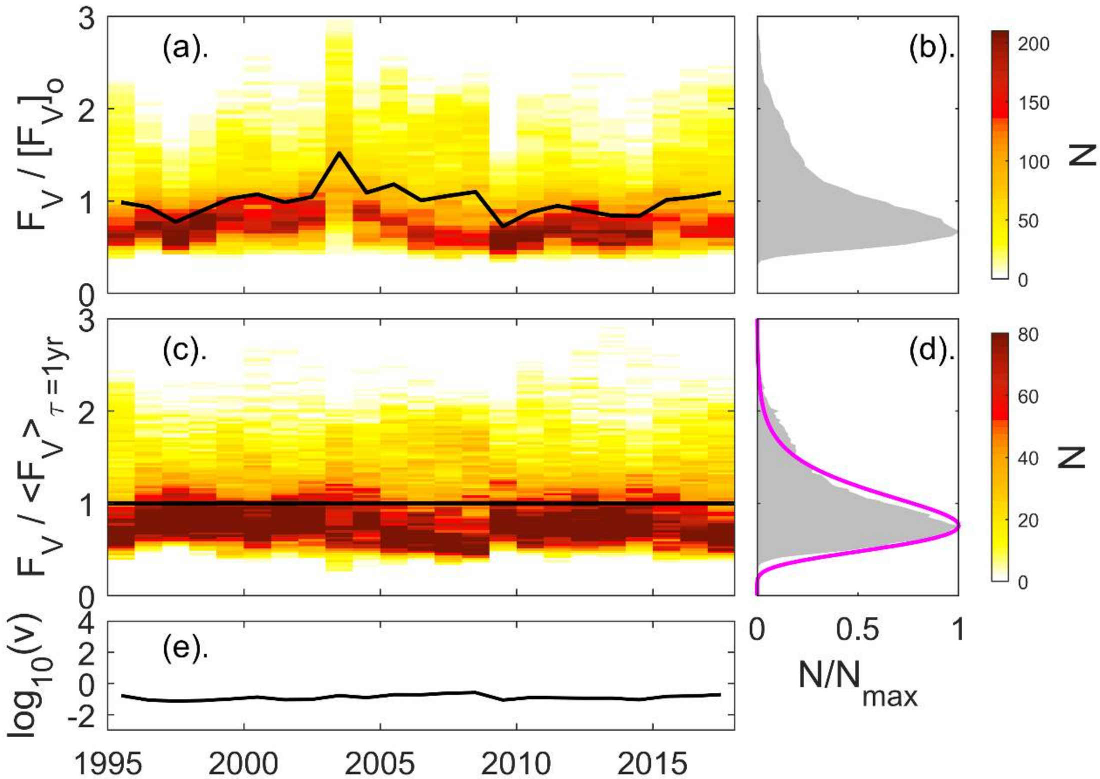


Figure 3. Analysis of the F_B term in P_α. (a) The annual distributions of 3-hourly values of $F_B/[F_B]_o$ (where $[F_B]_o$ is the mean of F_B for all the data from 1995-2017): the number of samples N in bins of $F_B/[F_B]_o$ that are 0.01 wide is color-contoured as function of year. The black line shows the annual mean values, plotted in the middle of the year. (b) The normalized distribution of $F_B/[F_B]_o$ for all years is shown as a grey histogram of N/N_{max}, where N_{max} is the peak value of N. (c) The annual probability density of 3-hourly values of $F_B/<F_B>_{\tau=1yr}$ (where $<F_B>_{\tau=1yr}$ is the annual mean of F_B for the year in question), color-contoured as function of year. The blackline shows the annual mean values which, by definition, are unity. (d) The normalized distribution of $F_B/<F_B>_{\tau=1yr}$ for all years is shown by the grey histogram which has been fitted with log-normal form with a mean of unity and a variance $v = 0.120$ (mauve line). (e) The logarithm of variance, v of the distributions.



826 **Figure 4**. Analysis of the F_V term in P_α in the same format as figure 3. The normalize distribution of $F_V/\langle F_V\rangle_{\tau=1yr}$ for τ = 3hrs and all years is shown in (d) by the grey histogram which has been fitted with log-normal distribution with a mean of unity and a variance v = 0.127 (mauve line).

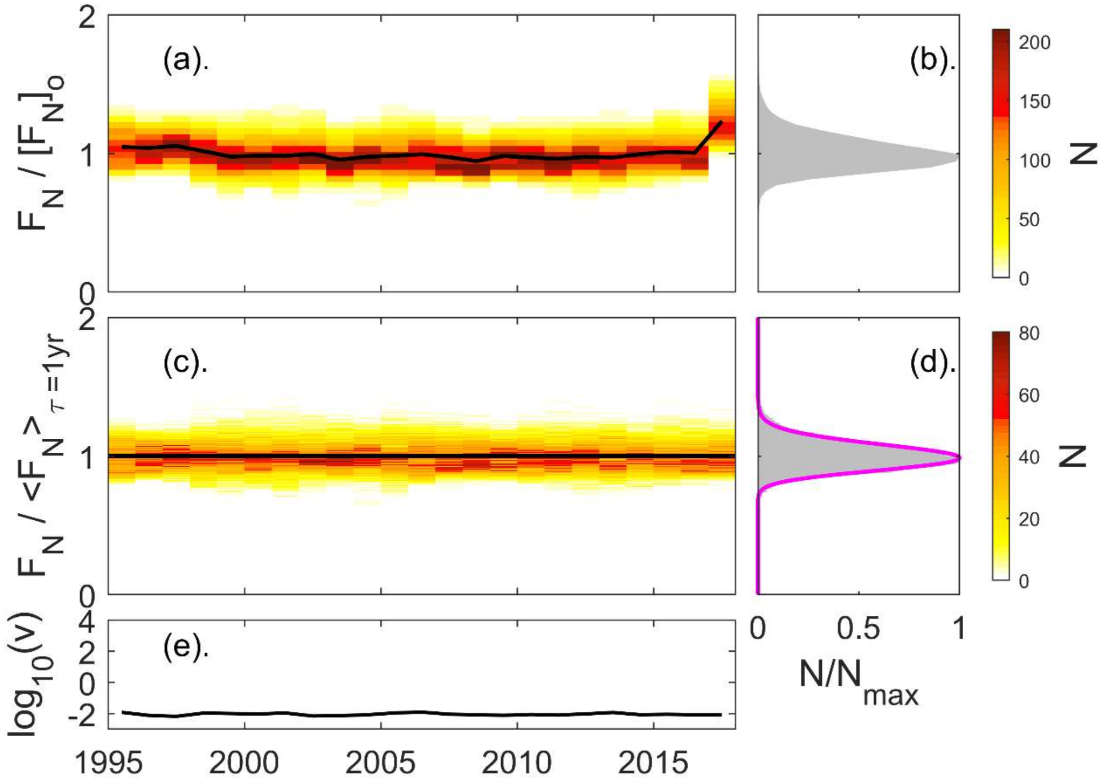


Figure 5. Analysis of the F_N term in P_α in the same format as figure 3. The normalized distribution of $F_N/<F_N>_{\tau=1yr}$ for τ = 3hrs and all years is shown in (d) by the grey histogram which has been fitted with log-normal distribution with a mean of unity and a variance $v = 0.009$ (mauve line).

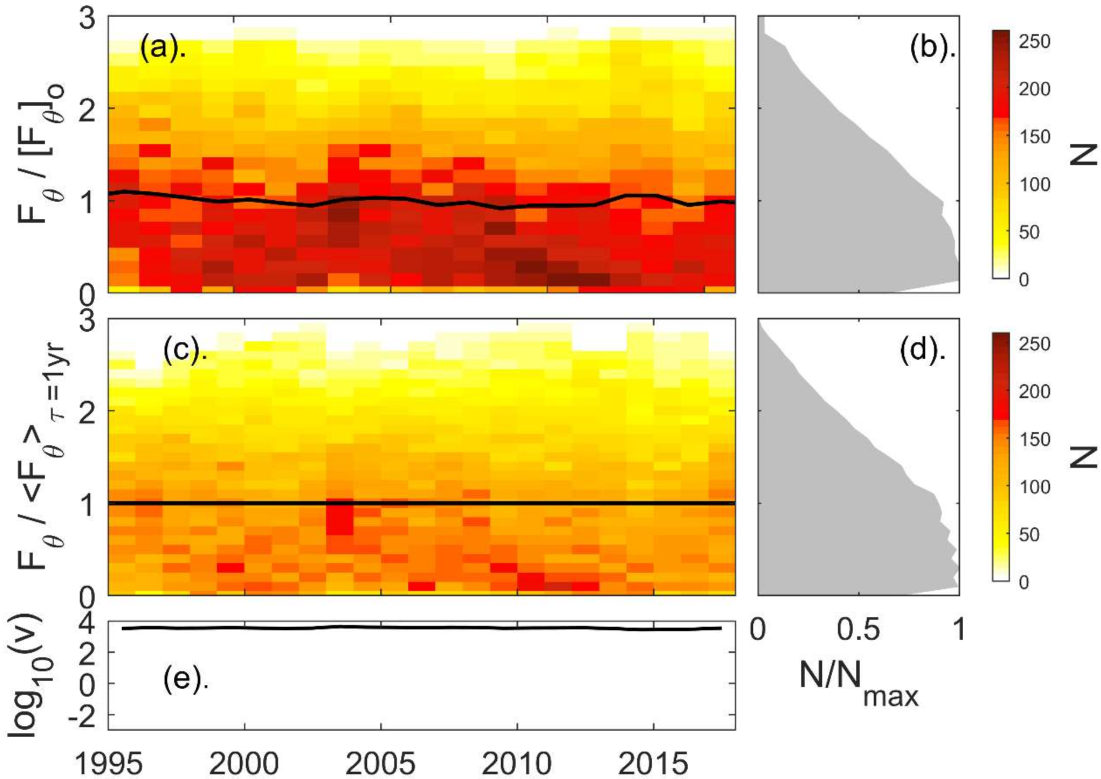


Figure 6. Analysis of the F_θ term in P_α in the same format as figure 3. The normalized distribution of $F_\theta/<F_\theta>_{\tau=1yr}$ for τ = 3hrs and all years is shown in (d) by the grey histogram which has not been fitted with a distribution as it does not match well any standard form. The mean of the annual variance values is $<v> = 3.542\times10^3$.

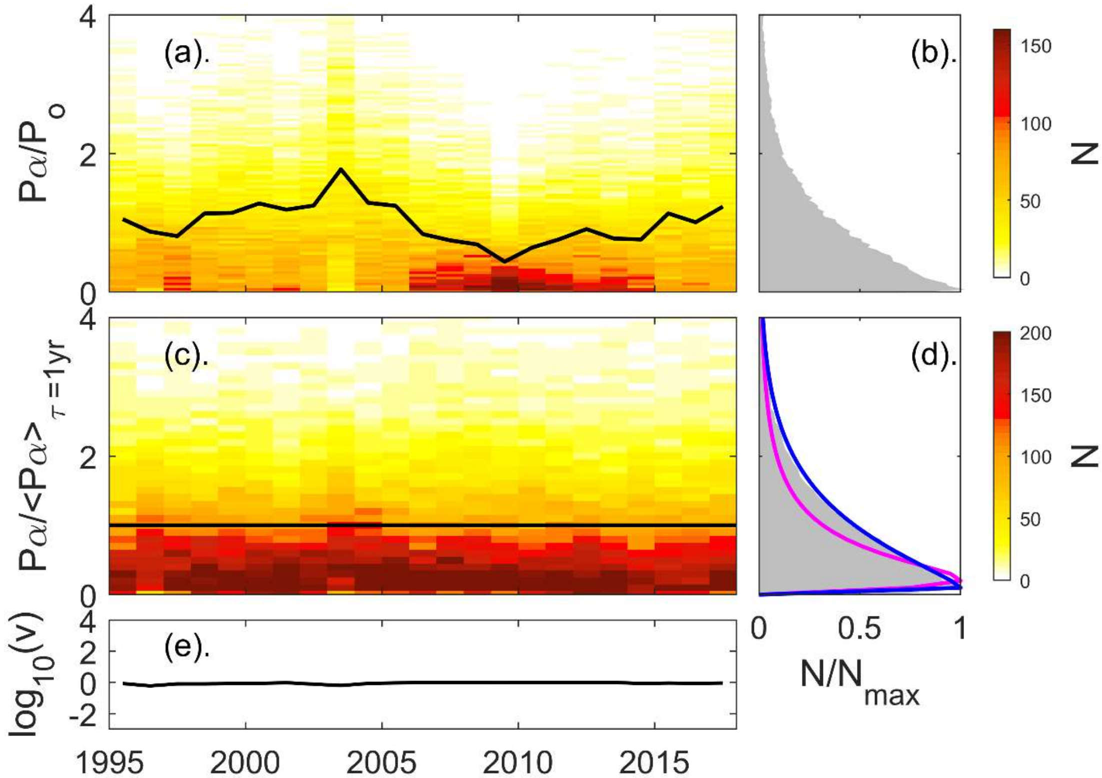


Figure 7. Analysis of P_α in the same format as figure 3. The normalized distribution of $P_\alpha/\langle P_\alpha\rangle_{\tau=1\mathrm{yr}}$ for τ = 3hrs and all years is shown in (d) by the grey histogram which has been fitted with log-normal distribution with a mean of unity and a variance $v = 1.788$ ($\mu = -0.5127$, mauve line) and a Weibull distribution with $k = 1.0625$ and $\lambda = 1.0240$ (blue line).

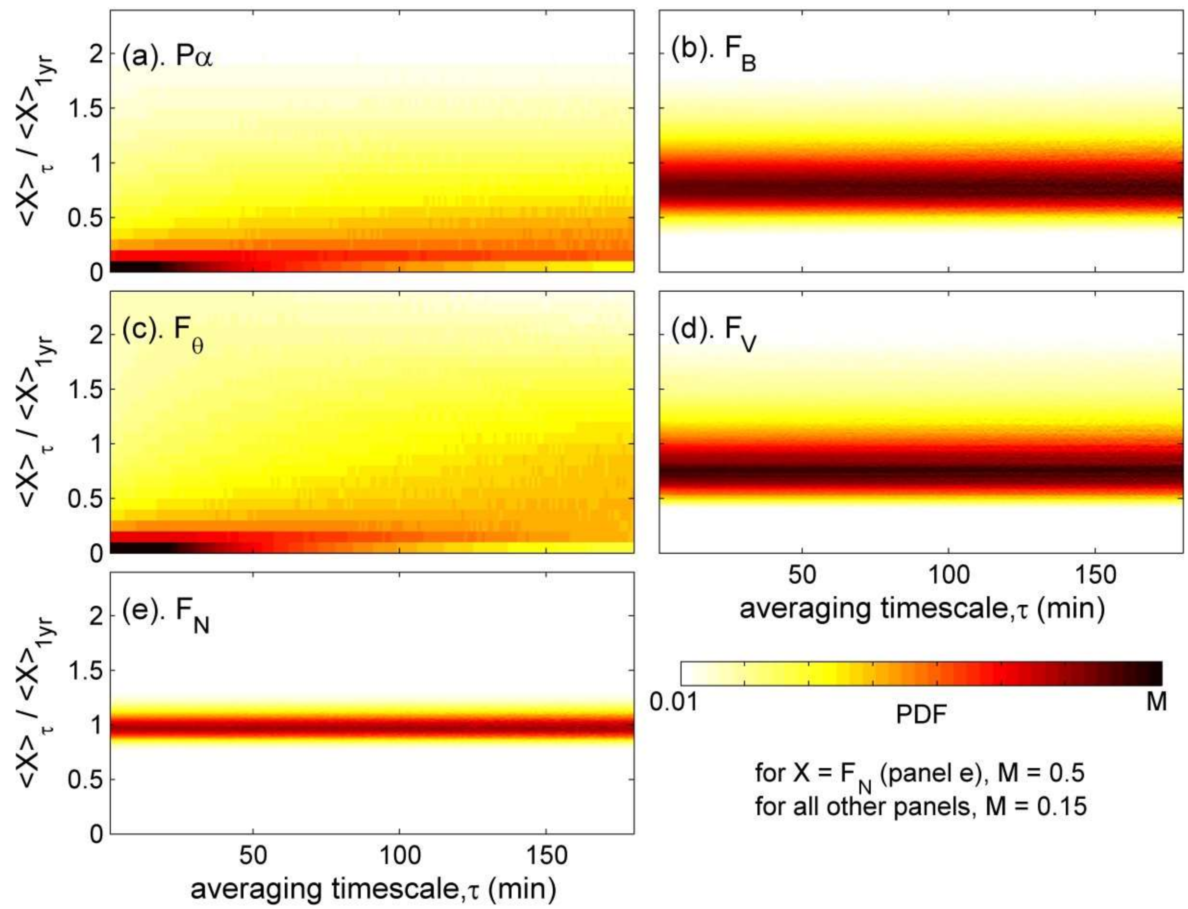


Figure 8. Analysis of the origin of the Weibull distribution of $<P_\alpha>_\tau/<P_{\alpha>\tau=1yr}$ for τ = 3hrs and all years, as shown in Figure 7d. In each panel, the PDF for a given τ is given as a vertical slice and τ varies along the horizontal axis between 1 min. and 3 hours. The panels are for: (a) $<P_\alpha>_\tau/<P_\alpha>_{\tau=1yr}$; (b) $<F_B>_\tau/<F_B>_{\tau=1yr}$; (c) $<F_\theta>_\tau/<F_\theta>_{\tau=1yr}$; (d) $<F_V>_\tau/<F_V>_{\tau=1yr}$; and (e) $<F_N>_\tau/<F_N>_{\tau=1yr}$.

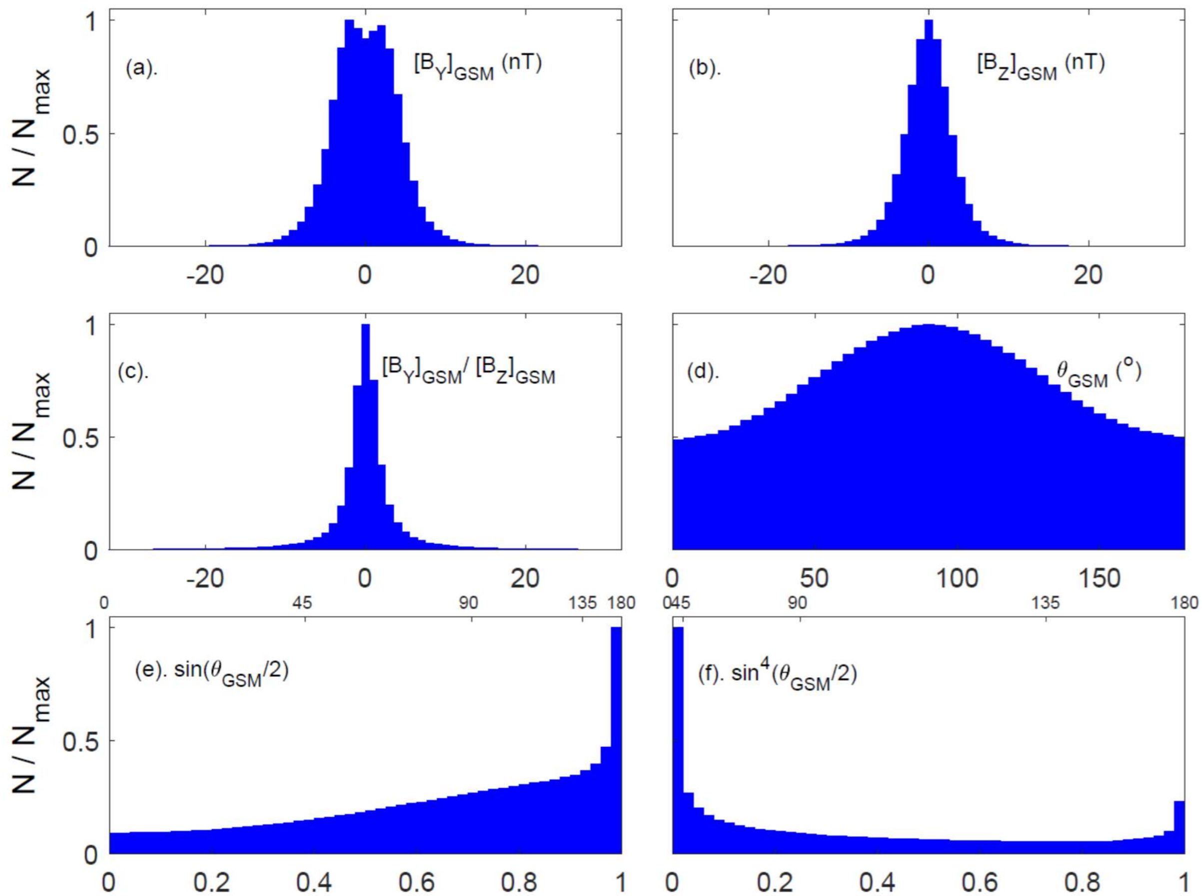


Figure 9. Analysis of the 10207789 valid 1-minute averages of the $\sin^4(\theta_{GSM}/2)$ term obtained in the years 1996-2016 (inclusive). The distribution of: (a) the IMF B_Y component in the GSM frame, $[B_Y]_{GSM}$; (b) the IMF B_Z component in the GSM frame, $[B_Z]_{GSM}$; (c) the ratio, $[B_Y]_{GSM}/[B_Z]_{GSM}$; (d) the IMF clock angle in the GSM frame, $\theta_{GSM} = \arctan(|[B_Y]_{GSM}| / [B_Z]_{GSM})$; (e) $\sin(\theta_{GSM}/2)$; and (f) $\sin^4(\theta_{GSM}/2)$. In each panel N is the number of samples per bin and N_{max} is the maximum value of N. In panels (e) and (f) the non-linear scales along the top (in small font) give the clock angle θ_{GSM} (in degrees) which corresponds to the lower scale, which is $\sin(\theta_{GSM}/2)$ in (e) and in $\sin^4(\theta_{GSM}/2)$ in (f).

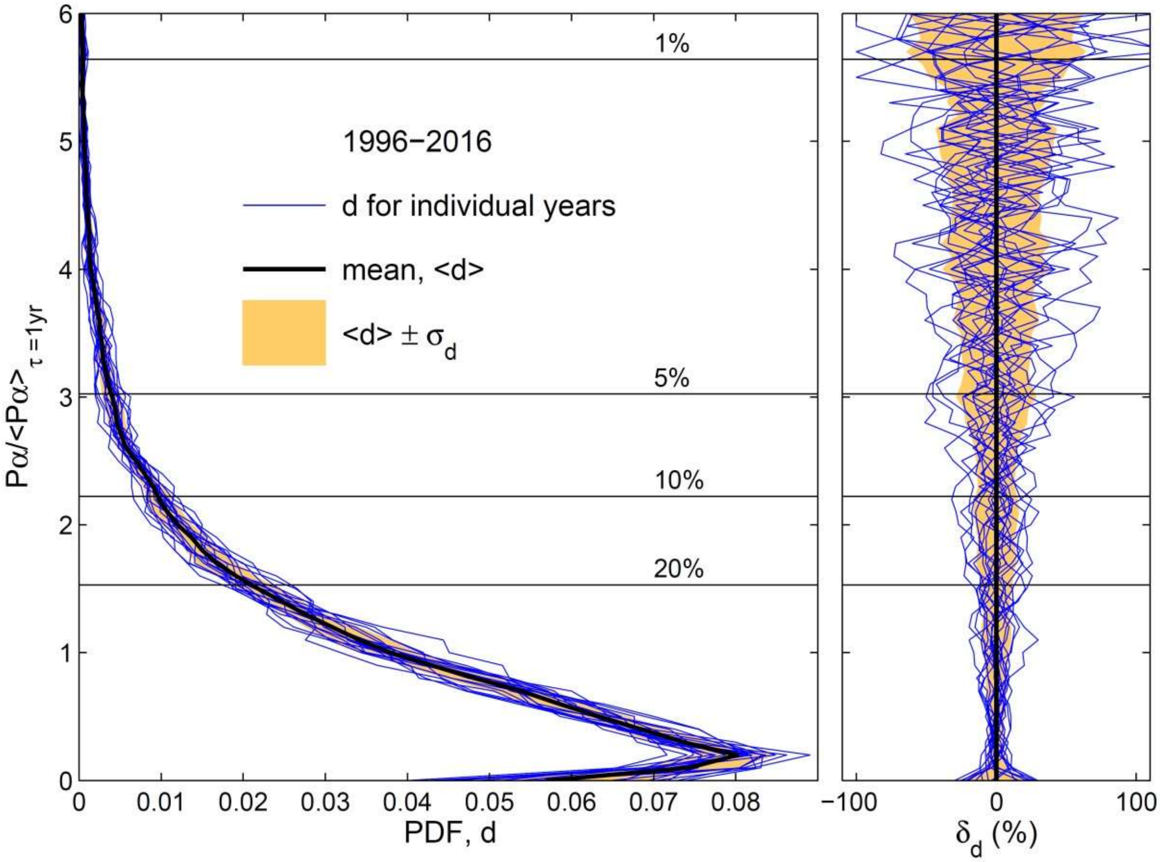


Figure 10. (a) Probability distribution function PDF, *d*, of $<P_\alpha>_\tau/<P_\alpha>_{\tau=1yr}$ for τ =3hrs. The black line is the overall distribution for all 21 years (as shown in figure 7d) and the blue lines are the values for individual years. The orange area is the mean of the annual values, plus and minus one standard deviation. Horizontal black lines are shown for the cumulative probability levels of 1%, 5%, 10% and 20%. (b). The percentage deviations of *d* from the mean, $\delta_d = 100(d-<d>)<d>$ in the same format.

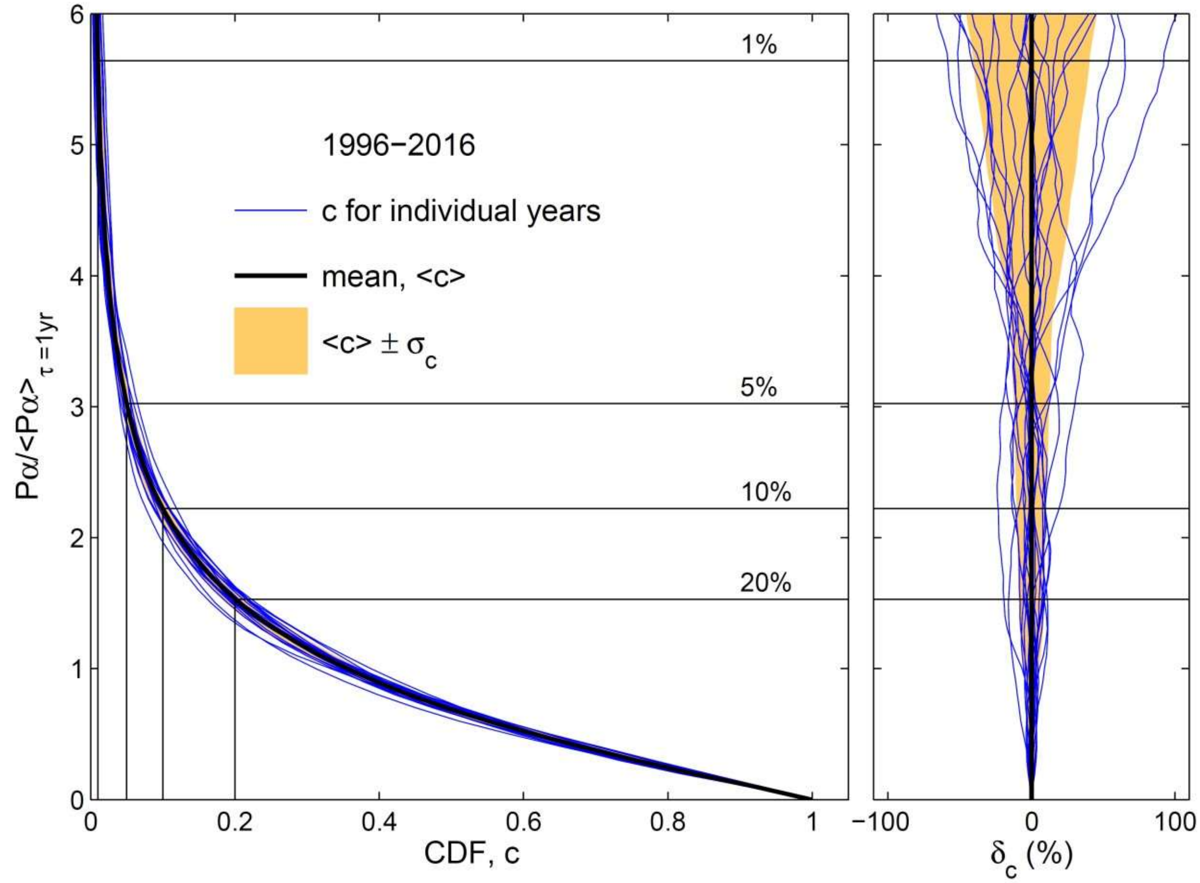


Figure 11. (a) Cumulative distribution functions CDF, *c*, of $\langle P_{\alpha} \rangle_{\tau} / \langle P_{\alpha} \rangle_{\tau=1yr}$ for τ =3hrs, corresponding to the PDFs in figure 10. The black line is the overall distribution for all 21 years and the blue lines are the values for individual years. The orange area is the mean of the annual values, plus and minus one standard deviation. Horizontal black lines are shown for the cumulative probability levels of 1%, 5%, 10% and 20%. (b). The percentage deviations of *c* from the mean, $\delta_c = 100(c - \langle c \rangle)\langle c \rangle$ in the same format.