

Original citation:

Tindale, Elizabeth and Chapman, Sandra C. (2017) Solar wind plasma parameter variability across solar cycles 23 and 24 : from turbulence to extremes. Journal of Geophysical Research : Space Physics.

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/92739>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 4.0 International license (CC BY 4.0) and may be reused according to the conditions of the license. For more details see: <http://creativecommons.org/licenses/by/4.0/>

A note on versions:

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk



RESEARCH ARTICLE

10.1002/2017JA024412

Solar Wind Plasma Parameter Variability Across Solar Cycles 23 and 24: From Turbulence to Extremes

Key Points:

- Statistical distribution of interplanetary magnetic field strength is multicomponent in fast and slow solar wind
- Core of distribution is close to lognormal; mean tracks solar cycle, insensitive to solar wind state and relative activity of each cycle
- Distribution of extremes has functional form invariant with solar cycle, its moments change with activity of each cycle

Supporting Information:

- Supporting Information S1

Correspondence to:

E. Tindale,
l.tindale@warwick.ac.uk

Citation:

Tindale, E., & Chapman, S. C. (2017). Solar wind plasma parameter variability across solar cycles 23 and 24: From turbulence to extremes. *Journal of Geophysical Research: Space Physics*, 122. <https://doi.org/10.1002/2017JA024412>

Received 30 MAY 2017

Accepted 18 SEP 2017

Accepted article online 21 SEP 2017

©2017. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

E. Tindale¹ and S. C. Chapman¹ ¹CFSA, Department of Physics, University of Warwick, Coventry, UK

Abstract Solar wind variability spans a wide range of amplitudes and timescales, from turbulent fluctuations to the 11 year solar cycle. We apply the data quantile-quantile (DQQ) method to NASA/Wind observations spanning solar cycles 23 and 24, to study how the uniqueness of each cycle maximum and minimum manifests in the changing statistical distribution of plasma parameters in fast and slow solar wind. The DQQ method allows us to discriminate between two distinct components of the distribution: the core region simply tracks the solar cycle in its moments but shows little sensitivity to solar wind state or the specific activity of each cycle. This would be consistent with an underlying in situ process such as turbulence driving the evolution of fluctuations up to an outer scale. In contrast, the tail component of the distribution is sensitive both to the differences between the maxima and minima of cycles 23 and 24, and the fast or slow state of the solar wind. The tail component varies over the solar cycle in such a way as to maintain a constant functional form, with only its moments varying with solar activity. Finally, after isolating the core region of the distribution, we test its lognormality over the solar cycle in each solar wind state and find the lognormal provides a more robust description of the statistics in slow wind than fast; however, in both states the goodness of fit is significantly reduced at solar maximum.

1. Introduction

The solar wind in the ecliptic exhibits two states: fast wind, which originates from the edges of long-lived, extended polar coronal holes (Hassler et al., 1999), and slow wind, thought to be generated in low-latitude, variable streamer belts and active regions (Schwenn, 2006). The fast wind is relatively smooth, whereas transient features including Viall puffs (Viall et al., 2010) and Sheeley blobs (Sheeley et al., 1997) are more characteristic of the slow wind (Rouillard et al., 2010). The 11 year solar cycle impacts upon the coronal sources of the fast and slow solar wind (Luhmann et al., 2002; Schwenn, 2010), so that on long timescales the solar wind reflects the variability of its solar origin. As each solar cycle is unique in its duration, its minimum, and its peak intensity, the impact on the solar wind is different for each cycle (Richardson et al., 2000; McComas et al., 2013). Turbulent fluctuations develop in both fast and slow wind (e.g., see textbooks, Tu & Marsch, 1995; Bruno & Carbone, 2013), resulting in further variability on scales from ~ 1 h down to below the ion gyroperiod (Kiyani et al., 2015). Turbulence evolving in situ in the solar wind flow may be expected to exhibit universal properties such as a power law power spectrum with a characteristic exponent (Osman et al., 2014) and non-Gaussian fluctuations (Sorriso-Valvo et al., 2001) that show scaling (Horbury et al., 2005; Chapman & Hnat, 2007). As the solar wind plasma is subject to solar cycle driving, the varying solar activity can be used to test the universality of these characteristics; for example, Chapman and Nicol (2009) compared the power spectra and scaling properties of the turbulence seen by Ulysses in fast polar flows and found that while the overall power in the fluctuations varied between the different minima of cycles 23 and 24, the scaling properties were stable against these changes in the driving, consistent with a universal turbulent process.

Extensive observations of the solar wind spanning several solar cycles allow the use of statistical methods to quantify the evolution of solar wind variability with solar cycle phase and between different cycles (Dmitriev et al., 2009; Veselovsky et al., 2010). Studying the full distribution of observed values of the interplanetary magnetic field (IMF), Burlaga and King (1979) found an approximately lognormal distribution; this was later corroborated by Slavin and Smith (1983). As the lognormal is the limit distribution for the observed values of a multiplicative process (Mitzenmacher, 2004), this finding is consistent with the turbulent nature of the solar wind plasma (Bruno & Carbone, 2013, and references therein). However, as shown by Feynman and Ruzmaikin (1994), the nonzero kurtosis of the IMF distribution is inconsistent with an exact lognormal.

Transient features of solar origin produce a heavier right-hand tail than that described by the conventional lognormal (Burlaga & Szabo, 1999), so a superposition of lognormals (Bruno et al., 2004) and more recently the log-kappa distribution (Vörös et al., 2015) have been explored as alternatives. Slavin and Smith (1983) fitted a lognormal distribution to the IMF at the maximum and minimum of solar cycle 20 and found that though its position was shifted, the width of the distribution did not change. Burlaga and Ness (1998) and Burlaga (2001) used hour-averaged Voyager data to investigate the evolution of the lognormal parameters with solar cycle and distance from the Sun and again found the width of the distribution was constant within uncertainties. Here we will use the quantile-quantile methodology to discriminate the core turbulence in a model-independent manner and test its lognormality across two solar cycles.

Despite their rarity, it is the most extreme space weather events that have the largest impact on Earth. However, they are sparsely observed so it is challenging to test potential models, and the dynamic nature of solar activity constantly changes what is considered an “extreme” event (Riley & Love, 2016). The tail of the distribution may therefore be sensitive to the solar cycle phase, as well as to longer-term trends in space climate (Lockwood et al., 1999). Wheatland (2000) modeled the waiting time between solar flares as a Poisson process whose rate parameter varied with solar activity. Tsubouchi and Omura (2007) applied the same approach to the *Dst* geomagnetic index but found the model was only valid provided the activity did not exceed a limit. Above this limit, extreme events occurred too close together in time to satisfy the independence condition inherent in the Poisson process, even after declustering had been applied. In other studies, the impact of the solar cycle on extreme events is investigated by performing the same extreme value analysis (Coles, 2001) at different solar cycle phases. Hush et al. (2015) compared the size distribution of bursts in the time series of the *AE* geomagnetic index, where a burst is defined as a continuous period where the signal exceeds a threshold, between three solar maxima and minima. The size distribution of the largest events was consistently exponential; however, the fit parameters were highly sensitive to each solar cycle. Moloney and Davidsen (2014) found that the size and duration of bursts in the solar wind ϵ parameter were insensitive to the solar cycle after adjusting for a cycle-dependent “characteristic timescale.” Amicis et al. (2006) studied the waiting times between periods of southward IMF, otherwise thought of as bursts in $-B_z$, and found a two-component distribution: a power law region exhibiting long-range memory and an exponential cutoff where Alfvénic fluctuations dominate. The relative weights of these components varied with solar activity. As the data quantile-quantile (DQQ) methodology allows us to explore the full distribution, our study will also address the solar cycle variation of the extreme values of solar wind plasma parameters, without requiring assumptions concerning the functional form of the distribution tail.

The data quantile-quantile (DQQ) plot can be used as a model-independent method for tracking changes in the statistical distribution of a variable across its full range. We applied it (Tindale & Chapman, 2016) to space weather relevant parameters in observations spanning solar cycles 23 and 24 seen by NASA’s Wind spacecraft. This revealed that solar wind-magnetosphere coupling parameters have a multicomponent functional form that is invariant to the solar cycle, with moments tracking changing solar cycle variability. Here we use the DQQ methodology to perform the first detailed study of variability in fast and slow solar wind across the full distribution. As no assumptions concerning the functional form of the distribution are required for this technique, it can be used to track changes in the moments of each part of a multicomponent distribution, as well as changes to the breakpoints between each component’s range of influence. The Wind data set spans two solar cycles, and we will use it to study the time dependence of the distribution of solar wind speed, IMF magnitude, and the components of the IMF observed in the fast and slow solar wind states. The more extreme values (the “tails”) of the likelihood distributions of IMF magnitude and component are composed of two parts whose functional forms do not change, but whose moments are sensitive both to solar cycle phase and intensity at solar maximum and minimum; this multicomponent behavior is not fully resolvable in slow wind, though the tail of the distribution is still distinct from the turbulent core. In the core region, we find the mean of the distribution varies with solar cycle while the variance is roughly constant, independent of the fast or slow state of the solar wind. Finally, we test the lognormality of the distribution core over solar cycles 23 and 24 and find the model provides a reasonable description at solar minimum, but provides a poor description of the distribution at maximum due to the dominance of sharp, localized peaks in the distribution, particularly in fast wind.

We begin by describing the data set and detailing aspects of the DQQ methodology relevant to this analysis in section 2, including a demonstration of the method using synthetic data. In section 3 we apply the method to the distribution of solar wind plasma parameters: first, to the full data set to resolve statistically the behavior

of the more extreme values and subsequently splitting the data into the fast and slow solar wind states. We identify the turbulent core of the IMF magnitude distribution and discuss its dynamics in section 4, focussing first on the time dependence of its moments and second on the efficacy of the lognormal model in describing the distribution during different phases of the solar cycle. Finally, our conclusions are presented in section 5.

2. The Data Sets and the DQQ Method

2.1. Data Used in the Study

We analyze data from the NASA Wind spacecraft at times when it was in the upstream solar wind, accessed through the NASA/GSFC OMNI data set. Fifteen-second IMF data from the MFI instrument and 92 s plasma velocity data from SWE were combined to give a 1 min resolution time series prior to our access. The data are in the geocentric solar magnetospheric (GSM) coordinate system. We took 1 year of data around each of the minima and maxima of solar cycles 23 and 24, spanning periods 1 December 1995 to 30 November 1996 (cycle 23 minimum), 1 October 1999 to 30 September 2000 (cycle 23 maximum), 1 August 2007 to 31 July 2008 (cycle 24 minimum), and 1 November 2013 to 31 October 2014 (cycle 24 maximum). These dates were chosen based on the sunspot number, as in Hush et al. (2015).

Each year-long sample contains roughly 500,000 data points. Around 10% of the data is missing from each time series, with data gaps roughly evenly spread through the four samples, leaving around 450,000 points per year. The data are split into the fast and slow streams with a boundary at 450 km s^{-1} , chosen as the midpoint between the typical fast and slow wind speeds of $\sim 300 \text{ km s}^{-1}$ and $\sim 600 \text{ km s}^{-1}$, respectively, to reduce as far as possible the contamination of slow wind data points into the fast wind data set and vice versa. This leaves around 150,000 data points in each fast wind sample and 300,000 in each slow wind sample. While the sample size in fast wind is significantly smaller than slow wind, both are sufficiently large to allow robust statistics for the DQQ method.

2.2. Constructing the DQQ Plot

The data quantile-quantile (DQQ) plot is a model-independent method for comparing two empirically sampled statistical distributions (Braun & Murdoch, 2016). Here we take two samples of the same variable at different times and use the DQQ plot to investigate changes in the distribution over its full range. While not requiring any a priori assumptions of the functional form of the underlying distribution, the DQQ plot can be used to test how the functional form changes over time, or whether the samples are drawn from distributions with the same model but with time-dependent moments. Moreover, if the underlying distribution is multicomponent, the DQQ plot will reveal the (potentially time-dependent) transitions between regions of parameter space that contain the distinct components or populations of observations. One can in principle compare different samples of data by qualitative inspection of overplotted probability density functions (PDFs). As we demonstrate in Figure 2 this is a far less sensitive indicator of how two distributions differ than the DQQ method. A quantitative approach would be direct fitting of the functional form of each PDF with standard distribution functions. However, for a two-component distribution, this implies an optimization to estimate 4–6 fitting parameters (location and scale, and if needed, shape parameter for each distribution) trialing distribution functional forms. If the problem at hand is to compare samples of the observations at different times to see how the system is changing, then this can be achieved simply by plotting the data on a DQQ plot without any need to fit the PDFs.

The analysis requires us to identify two timescales. First, we build up each distribution from a finite sample of data; therefore, the *sampling timescale* must be long enough to capture the full range of the distribution. However, it must also be sufficiently short that we can assume within-sample weak stationarity: within the sampling time we cannot resolve any change in distribution. Over the longer *observation timescale* the distribution may change from one state to another; by drawing samples spaced over the observation timescale we can test for these changes in distribution using the DQQ plot. The sampling timescale needs to be sufficiently short to avoid resolving changes on the known physical timescale of the system, here the approximately 11 year solar cycle, so our maximum sampling timescale is 1 year around solar maximum or minimum. At other phases of the solar cycle, this would need to be shorter. A minimum sampling timescale is chosen to give sufficient samples to resolve aspects of interest in the distribution. This depends on whether we are focussing on rare, extreme events, where we look across the entire distribution with 1 year samples, or on whether we are focussing on the core of the distribution only, where 1 month samples are sufficient.

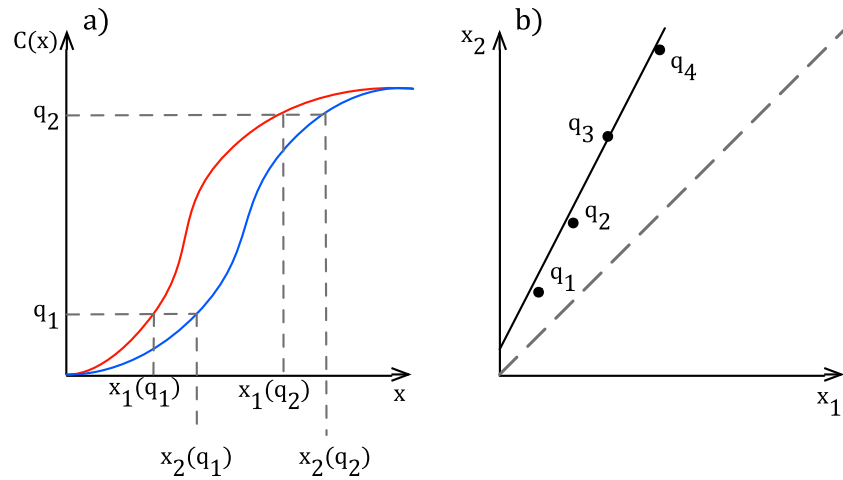


Figure 1. Diagrams showing the construction of the DQQ plot. (a) The empirical CDFs of samples x_1 (taken during period t_1) and x_2 (period t_2). Proportion q_1 of the set of observations is bounded by quantile $x_1(q_1)$ in sample x_1 and quantile $x_2(q_1)$ in sample x_2 , similar for q_2 . (b) The DQQ plot is produced by plotting $x_1(q)$ against $x_2(q)$ for all values of q between 0 and 1.

The DQQ procedure is summarized in Figure 1. Consider two samples of variable X , denoted by time series x_1 sampled during a period centered on t_1 and series x_2 sampled during a period centered on a later time t_2 . The cumulative distribution function (CDF) gives the probability of observing a value of X less than or equal to x , as a function of x (Figure 1a). It can be determined empirically by ordering the sampled data in increasing size and then plotting the data value against its position in the series normalized by the total number of observations. Note also that since the empirical CDF is normalized by the number of data points, after separating data into fast and slow solar wind, the DQQ plot traces changes in the conditional probability of the variable exceeding a certain value given that it is observed in one or the other solar wind state.

The CDF is a monotonic function and so can be inverted to give the quantile function: the value of $x(q)$ that is the upper bound of a given proportion q of the sample (Gilchrist, 2000). If the CDFs of time series x_1 and x_2 are given by $C_1(x_1)$ and $C_2(x_2)$, then the quantiles of the time series are

$$\begin{aligned} x_1(q) &= C_1^{-1}(q) \\ x_2(q) &= C_2^{-1}(q). \end{aligned} \quad (1)$$

As shown in Figure 1a, if the underlying distribution has changed between period t_1 and period t_2 (i.e., $C_1 \neq C_2$), the same proportion of data q will be bounded by different values of x in the two time series. If the quantile function of each sample is evaluated at the same set of q values from 0 to 1, we can compare each $(x_1(q), x_2(q))$ pair. The resulting DQQ plot (Figure 1b) has as its axes coordinates the empirical quantiles of X in each time series, and likelihood q as a parametric coordinate. Due to the step-like nature of the empirical CDF, interpolation between data values is required when calculating quantiles. Kernel density estimation is a popular tool for obtaining a smooth estimate of the empirical distribution (Bowman & Azzalini, 1997); here we will use it to obtain a continuous functional fit to the CDFs.

The shape of the DQQ plot depends on the relationship between the CDFs of the two samples. If both samples were drawn from the same distribution, the quantile functions would be the same; thus, all points on the DQQ plot would lie on the $y = x$ line. Indeed, this method is also used to compare data with empirical realizations of test distributions with the same moments; we refer to these fit-observation diagrams as QQ plots here. Provided the distribution can be uniquely specified in terms of scale and location parameters, any other straight line indicates

$$x_2(q) = \alpha x_1(q) + \beta, \quad (2)$$

so the CDFs are related by

$$C_2(x_2) = C_1(\alpha x_1 + \beta). \quad (3)$$

Therefore, the functional form of the distribution is the same in both cases; however, between t_1 and t_2 the parameters of the distribution changed. The change in the scale parameter is captured by the gradient of the DQQ plot, α , and the change in location parameter by the intercept, β . This simple linear transformation of variables allows the mapping of the first distribution onto the second, without requiring any knowledge as to the functional form of either distribution. As the CDF is the derivative of the probability density function (PDF), $P(x)$, equation (3) can be integrated to give an equivalent conversion of the PDF:

$$P(x_1)dx_1 = P\left(\frac{x_2 - \beta}{\alpha}\right) \frac{dx_2}{\alpha}, \quad (4)$$

so that the functional form of the PDF is also unchanged; however, an additional factor of α is required to conserve the normalization property.

If the DQQ plot is not linear (or piecewise linear in the case of a multicomponent distribution; Tindale & Chapman, 2016), this indicates the distribution functional form (if expressible in terms of location and scale parameters) has also changed between the samples. In this case, the DQQ plot can be used to discern whether the second functional form of the distribution is heavier or lighter tailed than the first (Wilk & Gnanadesikan, 1968). However, transformations (2) and (3) rely on the underlying distribution being expressible in standard form as $P(\bar{x}) = P((x - \beta)/\alpha)$. This is not always the case and an important exception is the lognormal distribution, which we discuss next.

2.3. DQQ Plot With Model Pseudodata

We will first demonstrate aspects of this method relevant to the analysis to follow, using pseudodata sets composed of numbers randomly drawn from test distributions. In Figure 2, we show three possible outcomes of DQQ analysis, with the empirically sampled PDF plots of the distributions to be compared in the left-hand column and the corresponding DQQ plot in the right-hand column. In the top row, both samples are drawn from the same test distribution: a Gumbel distribution with a location parameter $\mu = 1.5$ and scale parameter $\sigma = 0.75$. As discussed above, since the underlying distribution is the same in both cases, the DQQ trace lies along the $y = x$ line across its full range. In the middle row, the second sample is now drawn from a Gumbel distribution with different parameters, $\mu = 2.0$ and $\sigma = 1.0$. As the underlying functional form is unchanged, this results in a DQQ trace of a single straight line. The gradient and intercept of this line correspond to parameters α and β in equation (2); thus, they quantify the change in the location and scale parameters of the distribution between the two samples. Lastly, in the bottom row we compare the original Gumbel distribution to a Maxwellian with scale parameter $a = 1.2$. At the peak of the distribution these distributions track each other, and the tail is not sufficiently statistically resolved to distinguish between the two distributions within errors. However, the DQQ plot clearly shows that the tails differ; as the DQQ trace is not linear, the functional form of the two distributions must be different. The DQQ plot can therefore reveal whether two samples are drawn from two different distributions or from the same distribution with different moments, without requiring any a priori knowledge of the functional form of the distribution or fitting of its parameters.

The procedure was repeated using data drawn from the Gaussian, exponential and lognormal model distributions. In the first two cases, the PDFs can be expressed as a function of $(x - \mu)/\sigma$; thus, the linear transformation described above applies. However, for logarithmically distributed data, the natural logarithm of the variable X is distributed according to the Gaussian distribution; therefore, the probability of observing X in the range x to $x + dx$ is

$$P_{LN}(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right). \quad (5)$$

This is not of standard form $P(\bar{x}) = P((x - \beta)/\alpha)$, so even if two samples are both drawn from lognormal distributions they may no longer produce a straight line on the DQQ plot. However, as the natural logarithm of x is distributed according to the Gaussian, a transformation between two lognormal distributions will result in a straight line in natural log-log space. Therefore, the corresponding change of variables is

$$\ln(x_2) = \alpha \ln(x_1) + \beta. \quad (6)$$

We again demonstrate this using pseudodata in Figure 3. Now, sample x_1 is drawn from a standard lognormal distribution (location 0, scale 1), and samples x_2 and x_3 are drawn from lognormal distributions with location

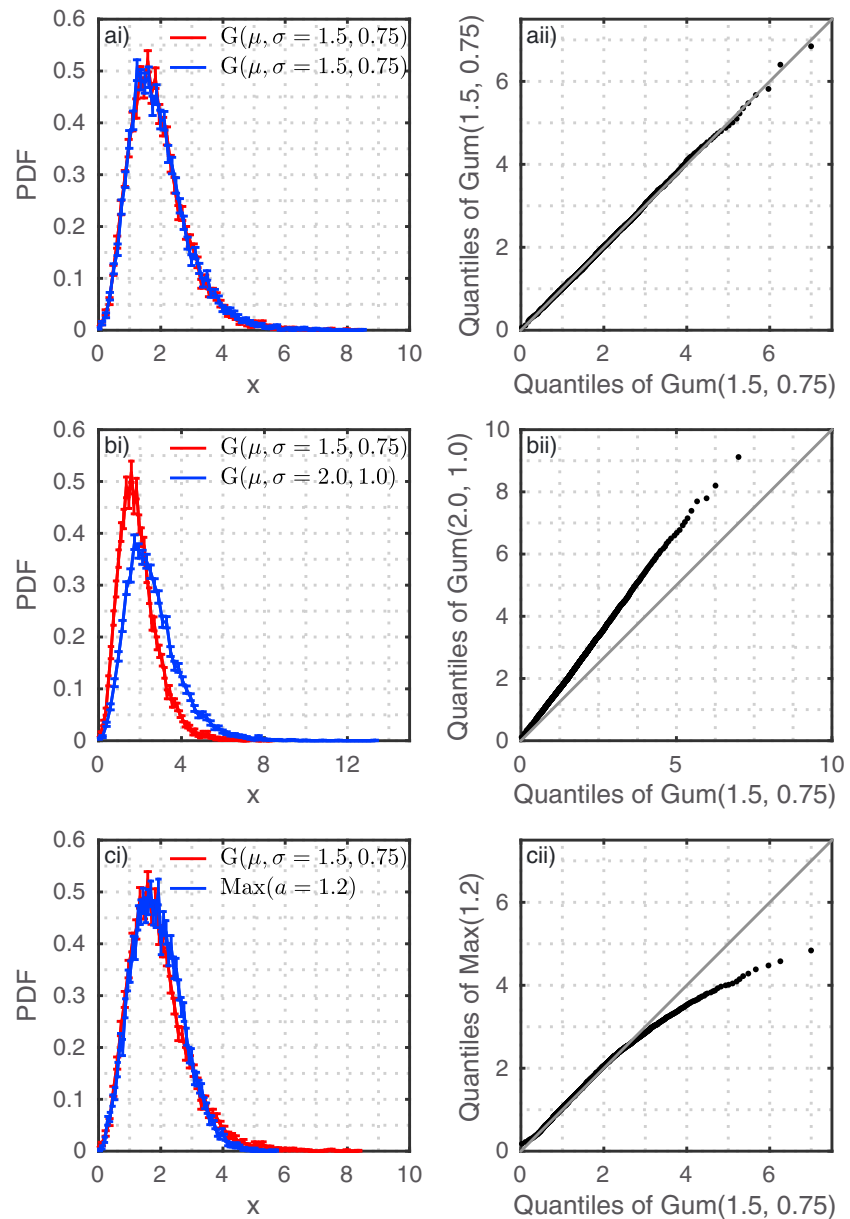


Figure 2. (Column i) The empirical probability density functions (PDFs) of pseudodata randomly drawn from two model distributions. (Column ii) The corresponding DQQ plot comparing the distributions in each case. Each PDF estimate is based on an unsmoothed normalized histogram, with errors defined as the square root of the count number. In row (a), both samples of pseudodata are generated from a Gumbel model distribution with a location parameter of 1.5 and scale parameter 0.75. In row (b), the second sample (blue) is drawn from a second Gumbel distribution with parameters $\mu = 2.0$ and $\sigma = 1.0$. Finally, in row (c), the original Gumbel distribution ($\mu = 1.5$ and $\sigma = 0.75$) is compared to a Maxwellian distribution with scale parameter $a = 1.2$. While the Gumbel and Maxwellian PDFs (Figure 2ci) appear indistinguishable within errors, the DQQ plot (Figure 2cii) exposes the difference between the distributions.

0 and scale 1.6 for x_2 , and location 1 and scale 1 for x_3 . In the left column of Figure 3, we see the DQQ plot with linear axes. It is immediately apparent that in the case of a change of scale (top row); a straight line will not capture the shape of the trace. When only the location parameter has varied (bottom row), the trace appears closer to linear, but its gradient and intercept do not reflect the change in moments. However, when we plot the same quantiles on natural logarithmic axes (right-hand column), a single straight line is recovered in both cases; in Figure 3aii the gradient is 1.6, as expected from this change in scale, while in Figure 3bii the intercept is 1, denoting the shift in mean.

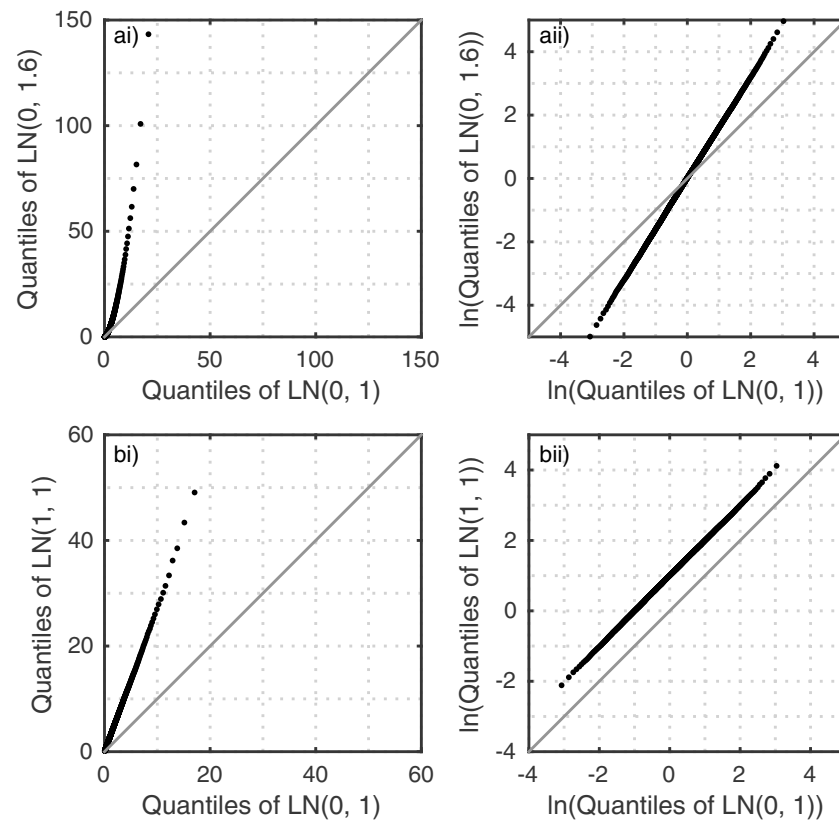


Figure 3. DQQ plot comparing lognormal distributions with different scales, with (left column) linear axes and (right column) natural log-log axes. In the upper panels we compare lognormal distributions with different scale parameters, while the lower panels show the comparison of lognormal distributions with different location parameters.

3. DQQ Plots for Plasma Parameters

3.1. Data Set Combining Both Fast and Slow Wind: Resolving the More Extreme Values

We now apply the DQQ method to study the evolution of the distribution of (a) solar wind speed v , (b) IMF magnitude B , and (c) components of the magnetic field. In the discussion of the turbulent core of the distribution that follows, we will for conciseness plot the IMF x component B_x only, as in the core the y and z component behavior was found to be similar. The DQQ plots in Figure 4 show how the distribution of the full range of each of these parameters changes between the minima of cycles 23 and 24 (left column), the consecutive maxima (middle column), and between the maximum and minimum of each cycle (right column), for cycle 23 in red and cycle 24 in blue. Here we use all data available in each year-long sample, comprising both fast and slow solar wind states.

All plots are composed of multiple linear regions. In most cases, more than 90% of the data fall into a single component, the “core” of the distribution. The remaining 10% of data form a distinct “tail” component, which itself splits into two parts in some cases. We label the quantile at which the tail splits from the core as \bar{q} , and any further breakpoints within the tail as q . In the case of the IMF components (here B_x), the variable is signed, so the distribution has both a positive and negative tail. With around 450,000 observations in each year-long sample, the tail component of the distribution of IMF magnitude B represents 50,000 observations that occurred during the year, of which 500 are in the uppermost distribution component above the $q \sim 0.999$ quantile. Each component undergoes a separate transformation between phases of the solar cycle, producing DQQ plots with a piecewise linear shape. The approximate linearity of each segment indicates that the functional form of the distribution is unchanged throughout these two cycles, while the displacement of each segment from the $y = x$ line signifies a solar cycle dependence in the moments of each component of the distribution.

The unique activity of each solar cycle is reflected in how the moments of the distribution change with time. The minimum of cycle 24 was deeper and more extended than those of the previous cycles observed

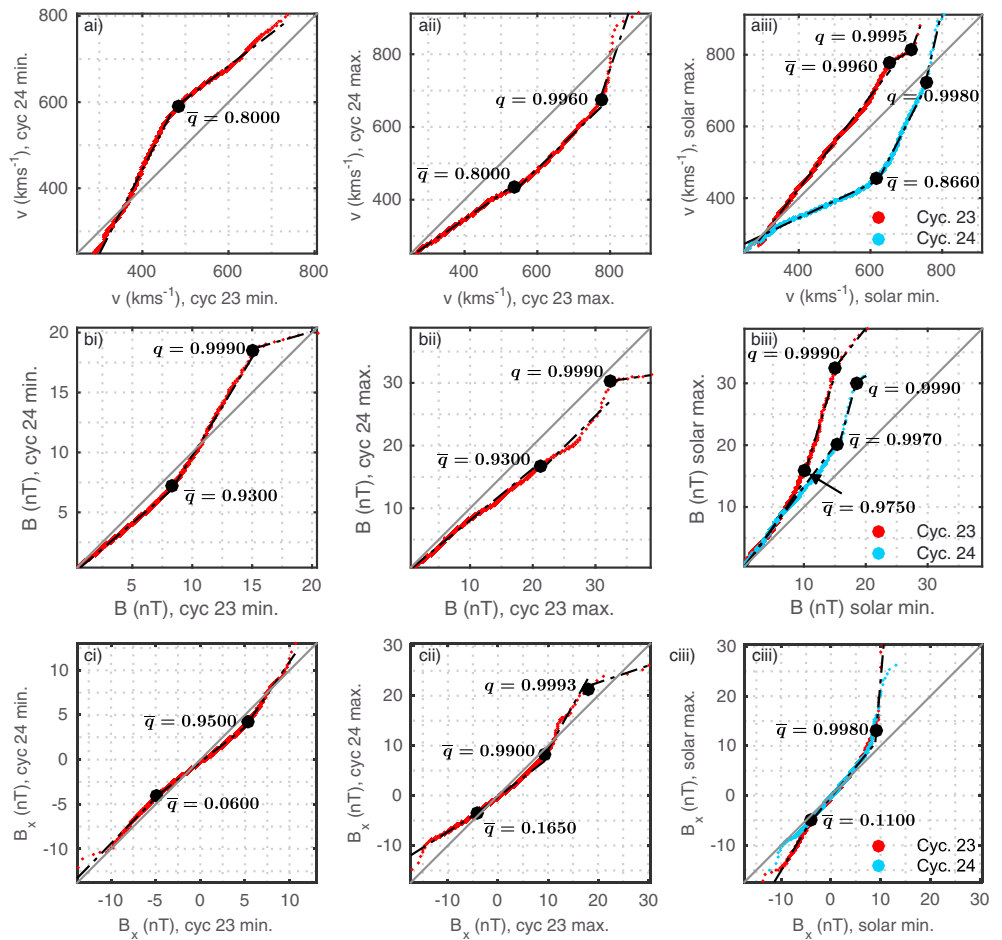


Figure 4. DQQ plots for (ai–aiii) plasma speed v , (bi–biii) interplanetary magnetic field magnitude B , and (ci–ciii) the x component of the IMF B_x . In column i we compare the minimum of cycle 23 to that of cycle 24; in column ii we compare the two solar maxima; and in column iii we compare the maximum and minimum within each cycle, for cycle 23 in red and cycle 24 in blue. Each DQQ plot is split into multiple linear components, with the range of each component defined by the highlighted quantiles; \bar{q} indicates the breakpoint between the core and tail of the distribution, while any subsequent breakpoints within the tail are denoted by q . As the IMF components are signed, their distribution has both a positive and negative tail; the breakpoints between the core and these two tail components are both denoted by \bar{q} .

since the space age. Its subsequent maximum also exhibited reduced activity, as shown by a lower sunspot number among other solar activity tracers (e.g., Tsurutani et al., 2011; McComas et al., 2013). Focussing on Figures 4bi–4biii, we see these differences between the cycles are reflected in the distribution of the magnetic field magnitude; Figure 4bii shows that the variance of all three components was larger at the maximum of cycle 23 than that of cycle 24, while in Figure 4biii we see a more marked change between the maximum and minimum of cycle 23 than 24, particularly in the extremal component. However, examining Figures 4ci–4ciii we see changes in distribution are less clear for the x component of the field; notably in Figure 4ciii we see for both cycles 23 and 24 the core of the distribution of B_x , located between the \bar{q} quantiles, did not change significantly between solar maximum and minimum.

Finally, inspecting the DQQ plots for solar wind speed in Figures 4ai–4aiii, we notice a natural split in the distribution at \bar{q} between 450 and 600 kms^{-1} . This is approximately the boundary between fast and slow solar wind, indicating that in distribution fast and slow wind transform differently over the solar cycles. The boundary between the fast and slow states also appears at different quantiles in each sample, revealing the variable probability of an observation occurring in fast or slow solar wind, depending on the solar cycle phase. This could suggest the impact of the solar cycle manifests differently in the two winds, so we will now use

the velocity measurements to create separate samples for the fast and slow wind states and repeat the DQQ procedure on each.

3.2. Separation of Fast and Slow Solar Wind

The DQQ plots for (a) v , (b) B and (c) B_x are shown in Figure 5, for fast (left column) and slow (right column) wind separately. In each case we compare the $q = 0.0001$ to $q = 0.9999$ quantiles of the distribution of each parameter at solar maximum to those at solar minimum, for cycles 23 (red) and 24 (blue). The highlighted points now show the 0.9 and 0.99 quantiles in all cases, i.e., the values that exceed 90% and 99% of each sample. To split the data into the fast and slow states, we choose a constant threshold of 450 km s^{-1} and define fast wind as any single data point whose speed measurement exceeds this, giving 150,000 (300,000) points in each year-long fast (slow) wind sample. Of these, roughly 15,000 (30,000) exceed the 0.9 quantile, and 1,500 (3,000) exceed the 0.99 quantile.

Beginning with the fast wind (Figure 5, left column), we see the same qualitative behavior in B and B_x as seen for the full solar wind data set in Figure 4. For the IMF magnitude this consists of a three part distribution; the $q_{0.9}$ quantile again acts as the upper limit for the core of the distribution, while the $q_{0.99}$ quantile splits the extremal tail component of the distribution into two regions with different time dynamics. The two-part tail dominates the plot on the linear scale due to the heavy-tailed nature of the distribution. The grey region highlights IMF magnitude values $B < 8 \text{ nT}$, which are within the core component of the distribution. This upper limit in B is determined using QQ plots with logarithmic axes of data versus lognormally distributed random variables (Figure 7) as discussed in section 4. The DQQ plots for B_x in Figure 4 exhibit a similar departure from the $y = x$ line at $|B_x| = 8 \text{ nT}$. In this region, the DQQ traces for each solar cycle lie on top of each other, so the different levels of activity at the maximum and minimum of cycles 23 and 24 do not impact upon how the moments of the core component vary over the solar cycle. This is also true for the distribution of IMF components; in Figures 5ci and 5cii the grey region highlights values $-8 \text{ nT} < B_x < 8 \text{ nT}$, wherein both traces again lie on top of each other, and here also along the $y = x$ line. The core component in the distribution of B_x therefore shows little sensitivity to the detailed differences across two solar cycles. At values outside of the $|B_x| < 8 \text{ nT}$ region, the DQQ traces deviate from each other and the $y = x$ line, showing the distinct time dynamics of the tail component that is sensitive to changes both within a solar cycle and between cycles 23 and 24.

Importantly, in fast wind the $q > q_{0.9}$ tail component of the IMF magnitude distribution ($B > 8 \text{ nT}$ values) can be approximated by a linear piecewise fit, suggesting that the functional form of the distribution tail does not change with solar cycle phase and simply tracks it with changing moments. The changing scale parameter of the tail component is shown by the steep gradient of each trace: at the maximum of cycle 23, the 0.9 and 0.99 quantiles of B were 12.2 nT and 23.9 nT, respectively, compared to the median value of 5.97 nT; whereas at the minimum of cycle 23, the corresponding values were 7.66 nT and 11.8 nT relative to a median of 5.02 nT. These values also demonstrate the unique activity of each cycle; while the highest plotted quantile at the maximum of cycle 23 was 6.7 times higher than the median; at the maximum of cycle 24 this ratio was 5.5. Importantly, this shows that the $\bar{q} < q < 0.99$ component transforms in the same way within cycles 23 and 24, whereas the transformation of the extremal $q > 0.99$ component is sensitive to the unique activity of each cycle.

In contrast, we cannot clearly identify the three-component structure in the distribution of the slow solar wind observations (Figure 5, right column), although there remains a distinct core component for both the magnitude and components of the IMF. The grey boxes again highlight the $B < 8 \text{ nT}$ and $-8 \text{ nT} < B_x < 8 \text{ nT}$ values, denoting core regions of the B and B_x distributions (as evident from Figure 7 in the case of B). The core of the distribution of IMF magnitude in slow wind has moments (location and scale parameters) that change between solar maximum and minimum, but these changes are the same between the respective maxima and minima of cycles 23 and 24. In contrast the core of the IMF component distribution (B_x shown), which in slow wind, shows no variation either within or between solar cycles. As in fast wind, the extremal tails of both the IMF magnitude and component distributions in slow solar wind are sensitive to varying solar conditions. The distribution of the IMF magnitude has a smaller scale parameter in slow wind than fast; during the maximum of cycle 23 the 0.9 and 0.99 quantiles were 9.97 nT and 15.9 nT, 1.6 and 2.5 times the median value, respectively. The scale of the distribution tail also varies less over the solar cycle for slow compared to fast wind, as shown by the smaller offset of the extremal quantiles from the $y = x$ line.

When plotted on linear axes, the DQQ plot naturally highlights the behavior of the extremal tail of the distribution. However, the majority of data samples are drawn from the core region of the distribution, and it

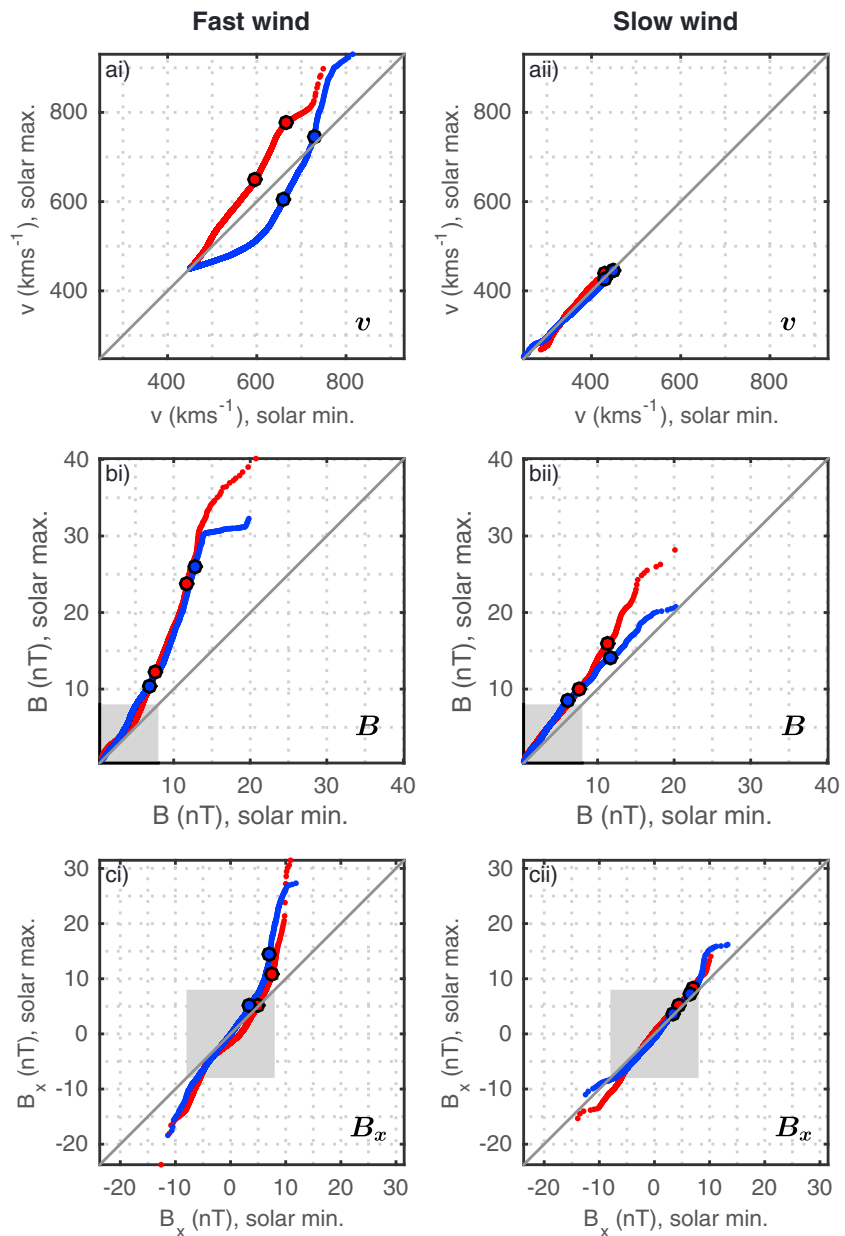


Figure 5. DQQ plots with linear axes for plasma parameters (ai and aii) solar wind speed v , (bi and bii) magnetic field magnitude B , and (ci and cii) IMF x component B_x , split into fast (left column) and slow (right column) wind. In each case we compare the distribution at solar maximum to that at solar minimum, for cycles 23 (red) and 24 (blue). In all cases the circles are the 0.9 and 0.99 quantiles. The grey shaded regions denote values $B < 8$ nT and -8 nT $< B_x < 8$ nT in Figures 5bi and 5bii and 5ci and 5cii), respectively.

displays little variation between the distinct maxima and minima of solar cycles 23 and 24. Above this scale (~ 8 nT) the IMF distribution is sensitive both to changes within the solar cycle and between cycles 23 and 24. These characteristics can be considered in the light of results (e.g., Horbury et al., 1995; Marsch & Tu, 1997) that suggest that active turbulence operates in both fast and slow solar wind. Alfvénic turbulence has been seen in the IMF components, while, due to the slightly compressive nature of the solar wind plasma, the signature of compressive turbulence has been noted in the IMF magnitude (Goldstein et al., 1995; Hnat et al., 2005, 2011). In Figure 6 we show the same quantiles of B seen in Figures 5bi and 5bii, but now on natural logarithmic axes. In particular for slow wind, and also to some extent for fast, we see straight lines up to approximately the $\bar{q} \sim 0.9$ quantile, indicative of a lognormal-to-lognormal transformation. The lognormal is the limit

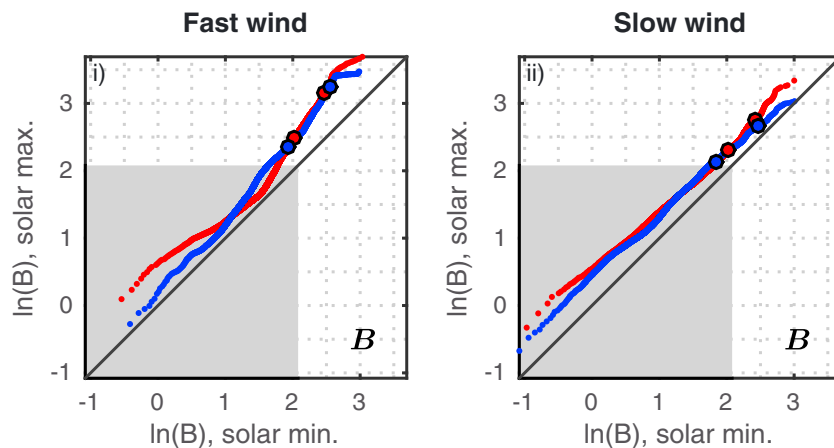


Figure 6. DQQ plots with natural logarithmic axes for the magnetic field magnitude B , comparing the distribution in (i) fast and (ii) slow wind between the maxima and minima of cycles 23 (red) and 24 (blue). The grey shaded regions denote $B \leq 8$ nT values.

distribution of a multiplicative process, which again has been associated with turbulence (Burlaga & Ness, 1998; Veselovsky et al., 2010; Vörös et al., 2015). The deviation of these lines from $y = x$ shows that the moments of this core turbulence component vary with solar cycle phase, though not with the unique activities of cycles 23 and 24.

In summary, we see a two-part distribution for the distribution of the IMF magnitude and its components. Values $B < 8$ nT and $|B_x| < 8$ nT are drawn from a core component whose moments vary between solar maximum and minimum, but are insensitive to the unique activity of each cycle as well as the fast or slow solar wind state. As shown by the logarithmic DQQ plot (Figure 6), the form of this core region is approximately lognormal. In both solar cycles the core component is more distinct in slow wind distributions, consistent with the longer period of development of turbulent fluctuations in slow wind before their detection at 1 AU. The largest observed values of the IMF form a separate extremal component in both solar wind states; in fast wind this extremal component has two parts, both of which retain their functional form over the solar cycle but have moments which change within and between the different cycles. In slow wind the transformation of the extremal component is less structured but is still separate to the transformation of the core of the distribution.

4. Solar Cycle Variation of the Distribution Core

We now investigate in more detail how well the lognormal model describes the core of the B distribution over the solar cycle. As shown by the deviation of the logarithmic DQQ plots from straight lines at high quantiles (Figure 6), we expect the data to be lognormally distributed only in a restricted core range. To determine this range, we first split each year-long sample into 12 month-long samples, maintaining separate data sets for fast and slow solar wind. We then discard any month-long samples with fewer than 5,000 observations; these will have insufficient data to support an accurate estimate of the empirical distribution. A test lognormal distribution is constructed for each of the remaining monthly samples, with parameters estimated by taking the natural logarithm of the data and determining its mean and variance. In Figure 7, the empirical quantiles of each month-long sample of B are compared to the quantiles of the test lognormal using QQ plots with natural log-log axes. The 12 month-long QQ traces from each year are superimposed on the same axes; row (a) shows the 12 months surrounding the minimum of cycle 23, row (b) shows the cycle 23 maximum, row (c) is the cycle 24 minimum and row (d) shows the cycle 24 maximum. As above, the left and right columns results for fast and slow wind, respectively.

In Figure 7 we can see a central region where the trace from each month collapses onto the $y = x$ line for all solar cycle phases. This reveals the range within which the data are lognormally distributed, which may correspond to the range within which the process is statistically dominated by turbulent fluctuations. At higher values the plots diverge due to the solar activity-sensitive tail component. This divergence can be seen

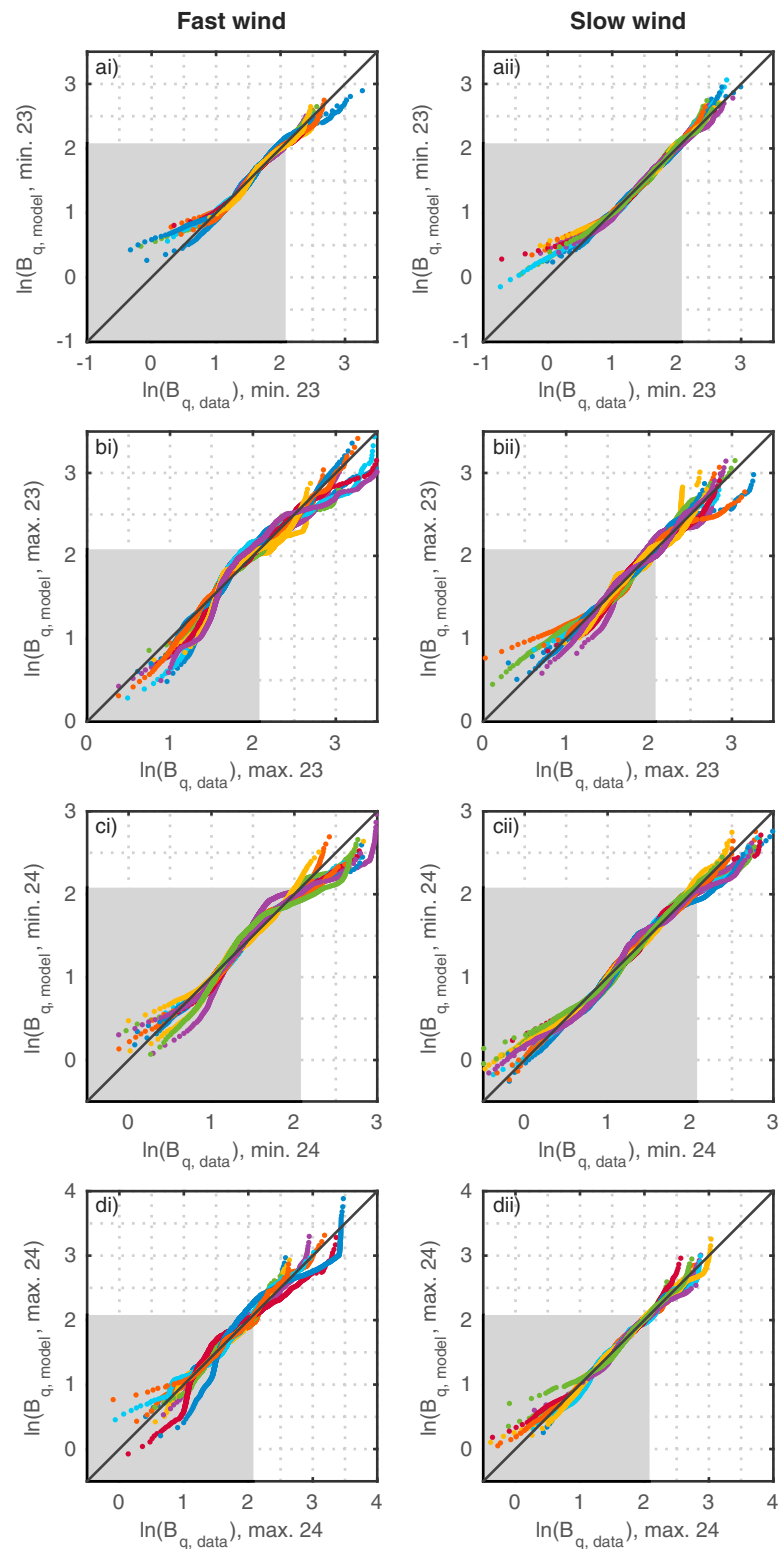


Figure 7. Fit-observation diagrams comparing the empirical quantiles of the IMF magnitude against the quantiles of empirical samples from a lognormal model with the same moments. In row (a) we split the year of data surrounding the minimum of cycle 23 into 12 months and superimpose the fit-observation diagrams for each month, for fast (left column) and slow (right column) wind. Rows (b) to (d) show the same plots for each of the years surrounding the maximum of cycle 23, and the minimum and maximum of cycle 24. The grey shaded boxes contain values $B \leq 8$ nT.

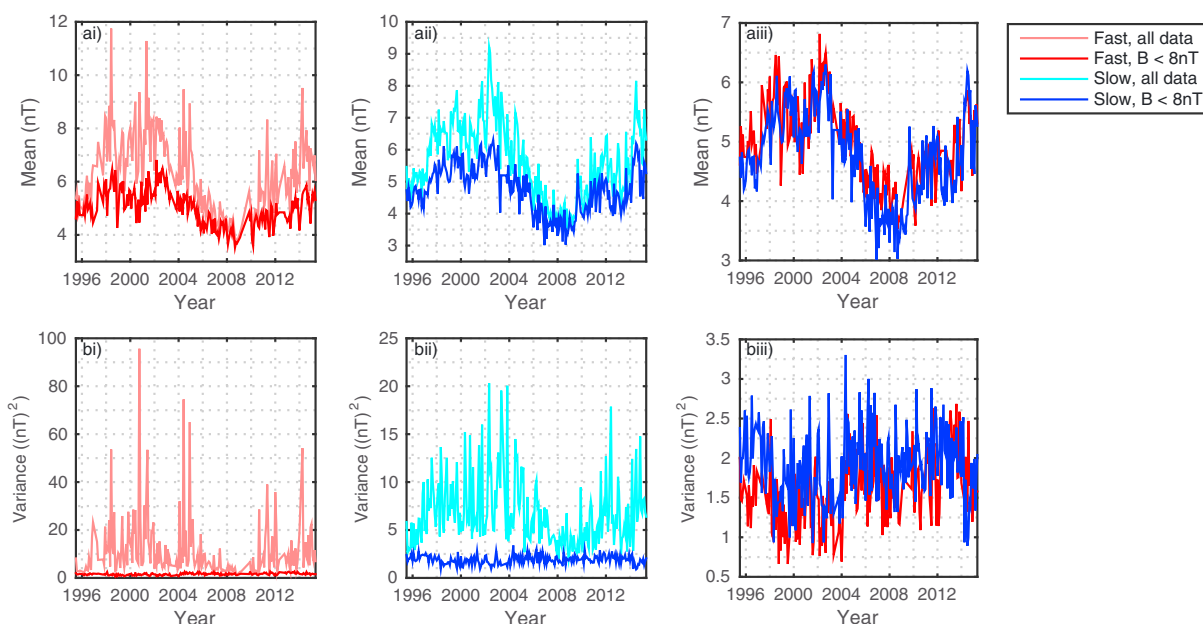


Figure 8. The time evolution of the mean (row a) and variance (row b) of month-long samples of IMF magnitude. Red lines denote properties of the fast wind data, while blue lines show slow wind properties. Light tones denote the moments of the full data set, whereas for darker shades only values in the core region (i.e., $B < 8$ nT) are considered.

around 8 nT, which we showed earlier as the upper boundary of the turbulent-dominated range. We also note the plots begin to diverge at low quantiles, at values around 3 nT; this may reflect uncertainties in calibration (Farrell et al., 1995; Gogoberidze et al., 2012). Since it is not possible to accurately determine the lower boundary of the turbulent range, as a default we retain all data up to 8 nT when discussing the core component.

Figure 8 shows the time evolution of the sample mean (top row) and variance (bottom row) of each month-long sample, with the moments of the full data set denoted by paler colors and those of data within the range $B < 8$ nT denoted by darker colors. The parameters are shown for fast and slow wind in red and blue, respectively. Inspecting first the top row of Figure 8, we see the expected solar cycle variation of the mean. When we consider the full distribution (i.e., paler colors), we see the variation in the calculated mean is greater in fast wind than slow, due to the sensitivity of the distribution to extreme events. When these are excluded by using only data $B < 8$ nT, the means of the turbulent core in fast and slow wind are similar and together track the solar cycle.

The variance of the full data set is also dominated by extreme values; at solar maximum the variance of the full fast wind data set can be 10 times higher than that of the restricted fast wind set. This ratio is lower in the slow wind, with a factor of 5 between the full and $B < 8$ nT data set variances. Again, when we use only

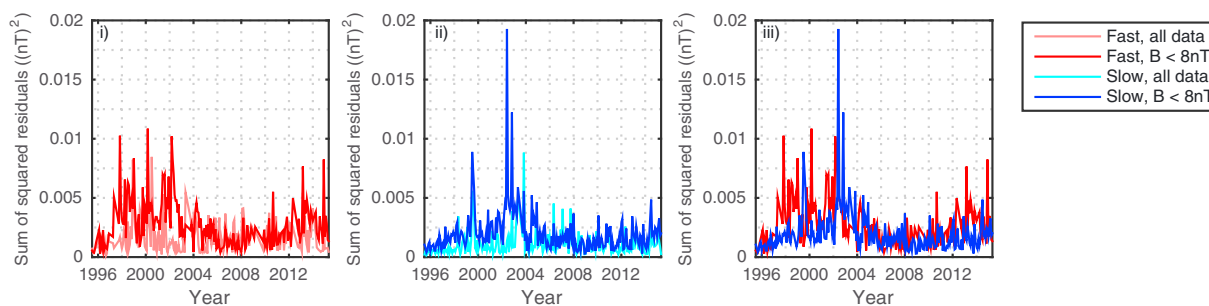


Figure 9. The normalized sum of squared residuals between the empirical PDF of month-long samples of the data and samples generated from a lognormal distribution with the same moments. Normalization is to the number of bins used to construct the empirical PDF. The format of the plots is the same as for Figure 8.

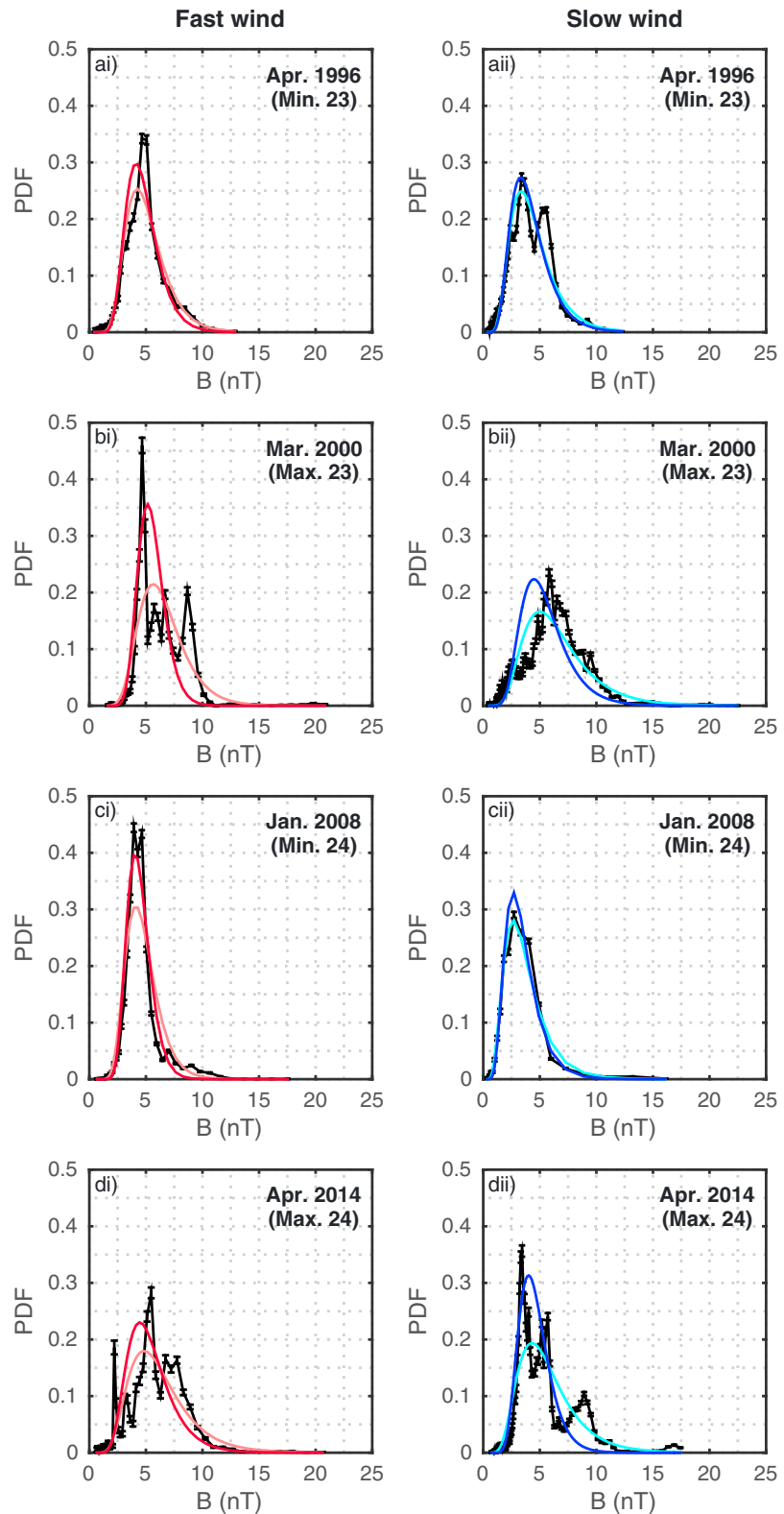


Figure 10. The empirical PDF of IMF field magnitude in (column i) fast and (column ii) slow wind, calculated from monthly data samples gathered in (a) April 1996, (b) March 2000, (c) January 2008, and (d) April 2014. The PDF is given by the normalized histogram with logarithmically spaced bins; thus, its error is the square root of the number of samples in each bin. Superimposed are the lognormal models whose moments are the same as those of the data, with the same color format as Figures 8 and 9.

the $B < 8$ nT values, the variance of the fast and slow wind become more similar, though variance in slow wind appears to consistently exceed that in fast wind by roughly 0.5 (nT)^2 . However, we now see that the intrinsic spread in the variance of the turbulent core dominates over any underlying solar cycle variation.

Having studied the evolution of the moments of the data, we now test to what extent the data are described by the lognormal model across the solar cycles. In Figure 9, we plot the evolution of the normalized sum of squared residuals between the empirically sampled lognormal model (as in Figure 7) obtained by using the moments and the empirical PDF calculated from each monthly sample, in the same format as Figure 8. The empirical PDFs are estimated by a normalized histogram with logarithmically spaced bins, with an error estimated as the square root of the count number. The number of bins was allowed to vary so as to be high as possible while maintaining a relative error below 5% for all bins below the 8 nT boundary for the turbulence-dominated range. The sum of square residuals (SSR) for the full data set, shown by the paler colors in Figure 9, is normalized by the number of bins used to sample the PDF; the normalized SSR of the restricted fit is normalized by the number of bins centered below 8 nT so that these are directly comparable. In Figure 10 are four sample distributions, from (a) April 1996, (b) March 2000, (c) January 2008, and (d) April 2014 (corresponding to the minima and maxima of cycles 23 and 24), for fast (left column) and slow (right column) wind. The lognormal model distributions for the IMF data are overlaid, using the same color scheme as Figures 8 and 9.

In Figure 9, the consistency of the lognormal model with the full distribution (pale lines) is reasonably constant over these two solar cycles for both solar wind states, fluctuating around 0.0015 (nT)^2 for fast and 0.0010 (nT)^2 for slow wind. The amount of variability around these values is increased at solar maximum. When we estimate the model parameters using only the < 8 nT core turbulence component, we see a much clearer solar cycle variation, particularly in fast wind. At solar minimum, the lognormal corresponds more closely to the $B < 8$ nT observations than the full data set, with a relatively smaller SSR of 0.0010 (nT)^2 in fast and 0.0008 (nT)^2 in slow wind. However, at solar maximum the lognormal function to $B < 8$ nT data performs less well, with a significantly larger normalized SSR around 0.004 (nT)^2 in both cases, suggesting at solar maximum the data are less well described by the lognormal model. The maximum normalized SSR values in slow wind appear later than those in fast; however, this may simply be due to restricted resolution on the plot; as there were few slow wind data points in these months, the number of bins used to calculate the empirical PDF had to be reduced to ensure statistical accuracy. This also accounts for the three anomalous peaks during 1999 and 2002 seen in the $B < 8$ nT model for slow wind.

Finally, in Figure 10 we overlay PDFs of these lognormal models with the observed PDFs. This may provide some insight into why the lognormal may not be a suitable description of the data at solar maximum. At solar minimum, the distribution tends to be dominated by a single peak; by excluding extreme events from the data set the < 8 nT model is better able to capture the height and shape of this peak. However, at solar maximum the shape of the distribution is highly irregular, with multiple smaller features as opposed to a single peak. The uncertainties on the empirical PDF show that these features are intrinsic to the data and are not simply an artifact of the PDF estimation method. We provide further examples in the supporting information (SI) (Figures S1 to S4). These show that for some month intervals around solar maximum the PDF has a two-peak structure, which has been interpreted as “ejecta” plasma (Klein & Burlaga, 1982; Xu & Borovsky, 2004). However, it is not a persistent feature and, as discussed in the SI, the DQQ plot would effectively distinguish it should contribute significantly to the distribution of the year-long data samples.

5. Conclusions

We have used data quantile-quantile (DQQ) plots to investigate how the difference between and across solar cycles is manifested in the statistical distribution of plasma parameters in fast and slow solar wind. As the DQQ method is model independent, we can trace changes to the distribution across its full range with no assumptions regarding its functional form. Should the underlying distribution be multicomponent, the DQQ method also allows us to identify transition points between the regions where each distribution component is statistically dominant.

Applying the DQQ method to observations of the interplanetary magnetic field (IMF) taken in the upstream solar wind by the Wind spacecraft, we are able to identify a two-part functional form for the distribution of the IMF magnitude and its components. The majority of the data falls into the core component; in this region the mean tracks the solar cycle while the variance fluctuates around a constant value, with the same variation

seen in both solar cycles and both solar wind states. The largest $\sim 10\%$ of observations form a distinct tail component, which varies in a distinguishable way between the distinct activity levels of the minimum and maximum of each solar cycle and also behaves differently in fast and slow solar wind. We find that the tail region becomes statistically dominant over the core for values above $B = 8$ nT ($|B_x| = 8$ nT).

The core region of the IMF distribution is not sensitive to the distinct activity levels of the maximum and minimum of each cycle. This is consistent with an underlying in situ process modifying the evolution of small fluctuations in this region. Using the DQQ plot with logarithmic axes, we find an approximately lognormal form for the core region; as the limit distribution of a multiplicative process, this is consistent with turbulence underlying the statistics of fluctuations in this region. However, when we test the lognormality of the core region over the solar cycle, we find that while the lognormal is a more robust description of the data in slow wind than fast across the solar cycle, in both solar wind states the lognormal most closely describes the data at solar minimum. At solar maximum, particularly in fast wind, the distribution exhibits multiple sharp peaks, and only in an interpolated sense follows a lognormal.

The extreme tail of the distribution is not captured by the lognormal model at any solar cycle phase, and excluding it from the data set used to create the lognormal model results in an improved correspondence to the core region at solar minimum. Unlike the turbulent core, the extremal component is sensitive to the differences in solar wind conditions in fast and slow solar wind seen at each distinct solar maximum and minimum. On DQQ plots with linear axes, the tails of the distributions of B and its components can be decomposed into piecewise linear fits. This means the distribution of extreme values tracks solar cycle changes in a manner consistent with an unchanging underlying functional form that simply shifts in its mean and variance. If this behavior is characteristic of all solar cycle changes, then it offers potential predictability of solar wind statistical “climate.”

Acknowledgments

We acknowledge NASA/GSFC for the provision of OMNI data through the Space Physics Data Facility's OMNIWeb service. The data were accessed at http://omniweb.gsfc.nasa.gov/ftpbrowser/wind_min_merge.html. This work was funded by funded by the U.K. Science and Technology Facilities Council (grant ST/N504506/1) and the U.S. Air Force Office of Scientific Research (grant FA9550-17-1-0054).

References

- Amicis, D. R., Bruno, R., Bavassano, B., Carbone, V., & Sorriso-Valvo, L. (2006). On the scaling of waiting-time distributions of the negative IMF B_z component. *Annales Geophysicae*, 24, 2735–2741. <https://doi.org/10.5194/angeo-24-2735-2006>
- Bowman, A. W., & Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford, UK: Oxford University Press.
- Braun, W. J., & Murdoch, D. J. (2016). *A First Course in Statistical Programming With R*. New York: Cambridge University Press.
- Bruno, R., Carbone, V., Primavera, L., Malara, F., Sorriso-Valvo, L., Bavassano, B., & Veltri, P. (2004). On the probability distribution function of small-scale interplanetary magnetic field fluctuations. *Annales Geophysicae*, 22, 3751–3769. <https://doi.org/10.5194/angeo-22-3751-2004>
- Bruno, R., & Carbone, V. (2013). The solar wind as a turbulence laboratory. *Living Reviews in Solar Physics*, 10, 2.
- Bruno, R., & Carbone, V. (2013). *Turbulence in the Solar Wind*. Cham, Switzerland: Springer.
- Burlaga, L. F. (2001). Lognormal and multifractal distributions of the heliospheric magnetic field. *Journal of Geophysical Research*, 106(A8), 15,917–15,928. <https://doi.org/10.1029/2000JA000107>
- Burlaga, L. F., & King, J. H. (1979). Intense interplanetary magnetic fields observed by geocentric spacecraft during 1969–1975. *Journal of Geophysical Research*, 84(A11), 6633–6640. <https://doi.org/10.1029/JA084iA11p06633>
- Burlaga, L. F., & Ness, N. F. (1998). Magnetic field strength distributions and spectra in the heliosphere and their significance for cosmic ray modulation: Voyager 1, 1980–1994. *Journal of Geophysical Research*, 103(A12), 29,719–29,732. <https://doi.org/10.1029/98JA02682>
- Burlaga, L. F., & Szabo, A. (1999). Fast and slow flows in the solar wind near the ecliptic at 1 AU. *Space Science Reviews*, 87, 137–140. <https://doi.org/10.1023/A:1005186720589>
- Chapman, S. C., & Hnat, B. (2007). Quantifying scaling in the velocity field of the anisotropic turbulent solar wind. *Geophysical Research Letters*, 34, L17103. <https://doi.org/10.1029/2007GL030518>
- Chapman, S. C., & Nicol, R. M. (2009). Generalized similarity in finite range solar wind magnetohydrodynamic turbulence. *Physical Review Letters*, 103, 241101. <https://doi.org/10.1103/PhysRevLett.103.241101>
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. London: Springer.
- Dmitriev, A. V., Suvorova, A. V., & Veselovsky, I. S. (2009). Statistical characteristics of the heliospheric plasma and magnetic field at the Earth's orbit during four solar cycles 20–23. In H. E. Johannson (Ed.), *Handbook on Solar Wind: Effects, Dynamics and Interactions* (pp. 81–144). New York: NOVA Science Publishers.
- Farrell, W. M., Thompson, R. F., Lepping, R. P., & Byrnes, J. B. (1995). A method of calibrating magnetometers on a spinning spacecraft. *IEEE Transactions on Magnetics*, 31(2), 966–972. <https://doi.org/10.1109/20.364770>
- Feynman, J., & Ruzmaikin, A. (1994). Distributions of the interplanetary magnetic field revisited. *Journal of Geophysical Research*, 99(A9), 17,645–17,651. <https://doi.org/10.1029/94JA01098>
- Gilchrist, W. G. (2000). *Statistical Modelling with Quantile Functions*. Boca Raton, FL: Chapman and Hall.
- Gogoberidze, G., Chapman, S. C., Hnat, B., & Dunlop, M. W. (2012). Impact of measurement uncertainties on universal scaling of MHD turbulence. *Monthly Notices of the Royal Astronomical Society*, 426, 951–955. <https://doi.org/10.1111/j.1365-2966.2012.21775.x>
- Goldstein, M. L., Roberts, D. A., & Matthaeus, W. H. (1995). Magnetohydrodynamic turbulence in the solar wind. *Annals Review Astrophysics*, 33, 283–325. <https://doi.org/10.1146/annurev.aa.33.090195.001435>
- Hassler, D. M., Dammach, I. E., Lemaire, P., Brekke, P., Curdt, W., Mason, H. E., ... Wilhelm, K. (1999). Solar wind outflow and the chromospheric magnetic network. *Science*, 283, 5403. <https://doi.org/10.1126/science.283.5403.810>
- Hnat, B., Chapman, S. C., Gogoberidze, G., & Wicks, R. T. (2011). Scale-free texture of the fast solar wind. *Physical Review E*, 84, 065401R. <https://doi.org/10.1103/PhysRevE.84.065401>

- Hnat, B., Chapman, S. C., & Rowlands, G. (2005). Compressibility in solar wind plasma turbulence. *Physical Review E*, *94*, 204502. <https://doi.org/10.1103/PhysRevLett.94.204502>
- Horbury, T. S., Balogh, A., Forsyth, R. J., & Smith, E. J. (1995). Observations of evolving turbulence in the polar solar wind. *Geophysical Research Letters*, *22*(23), 3401–3404. <https://doi.org/10.1029/95GL03550>
- Horbury, T. S., Forman, M. A., & Oughton, S. (2005). Spacecraft observations of solar wind turbulence: An overview. *Plasma Phys. Controlled Fusion*, *47*(12B), B703–B717. <https://doi.org/10.1088/0741-3335/47/12B/552>
- Hush, P., Chapman, S. C., Dunlop, M. W., & Watkins, N. W. (2015). Robust statistical properties of the size of large burst events in AE. *Geophysical Research Letters*, *42*, 9197–9202. <https://doi.org/10.1002/2015GL066277>
- Kiyani, K. H., Osman, K. T., & Chapman, S. C. (2015). Dissipation and heating in solar wind turbulence: From the macro to the micro and back again. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, *373*, 20140155. <https://doi.org/10.1098/rsta.2014.0155>
- Klein, L. W., & Burlaga, L. F. (1982). Interplanetary magnetic clouds at 1 AU. *Journal of Geophysical Research*, *87*(A2), 613–624. <https://doi.org/10.1029/JA087iA02p00613>
- Lockwood, M., Stamper, R., & Wild, M. N. (1999). A doubling of the Sun's coronal magnetic field during the past 100 years. *Nature*, *399*, 437–439. <https://doi.org/10.1038/20867>
- Luhmann, J. G., Li, Y., Arge, C. N., Gazis, P. R., & Ulrich, R. (2002). Solar cycle changes in coronal holes and space weather cycles. *Journal of Geophysical Research*, *107*(A8), 1154. <https://doi.org/10.1029/2001JA007550>
- Marsch, E., & Tu, C.-Y. (1997). Intermittency, non-Gaussian statistics and fractal scaling of MHD fluctuations in the solar wind. *Nonlinear Processes in Geophysics*, *4*(2), 101–124. <https://doi.org/10.5194/npg-4-101-1997>
- McComas, D. J., Angold, N., Elliot, H. A., Livadiotis, G., Schwadron, N. A., Skoug, R. M., & Smith, C. W. (2013). Weakest solar wind of the space age and the current “mini” solar maximum. *Astrophysical Journal*, *779*(1), 2. <https://doi.org/10.1088/0004-637X/779/1/2>
- Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, *2*(1), 226–251. <https://doi.org/10.1080/15427951.2004.10129088>
- Moloney, N. R., & Davidsen, J. (2014). Stationarity of extreme bursts in the solar wind. *Physical Review E*, *89*, 052812. <https://doi.org/10.1103/PhysRevE.89.052812>
- Osman, K. T., Kiyani, K. H., Chapman, S. C., & Hnat, B. (2014). Anisotropic intermittency of magnetohydrodynamic turbulence. *Astrophysical Journal Letter*, *783*(2), L27. <https://doi.org/10.1088/2041-8205/783/2/L27>
- Richardson, I. G., Berdichevsky, D., Desch, M. D., & Farrugia, C. J. (2000). Solar-cycle variation of low density solar wind during more than three cycles. *Geophysical Research Letters*, *27*(23), 3761–3764. <https://doi.org/10.1029/2000GL000077>
- Riley, P., & Love, J. L. (2016). Extreme geomagnetic storms: Probabilistic forecasts and their uncertainties. *Journal of Geophysical Research: Space Weather*, *15*, 53–64. <https://doi.org/10.1002/2016SW001470>
- Rouillard, A. P., Davies, J. A., Lavraud, B., Forsyth, R. J., Savani, N. P., Bewsher, D., ... Eyles, C. J. (2010). Intermittent release of transients in the slow solar wind: 1. Remote sensing observations. *Journal of Geophysical Research*, *115*, A04103. <https://doi.org/10.1029/2009JA014471>
- Schwenn, R. (2006). Solar wind sources and their variations over the solar cycle. *Space Science Reviews*, *124*, 51–56. <https://doi.org/10.1007/s11214-006-9099-5>
- Schwenn, R. (2010). The solar cycle. *Living Reviews in Solar Physics*, *7*, 1.
- Sheeley, N. R. Jr., Wang, Y.-M., Hawley, S. H., Brueckner, G. E., Dere, K. P., Howard, R. A., ... Biesscker, D. A. (1997). Measurements of flow speeds in the corona between 2 and 30 R_{\odot} . *Astrophysical Journal*, *484*, 472–478. <https://doi.org/10.1086/304338>
- Slavin, J. A., & Smith, E. J. (1983). Solar cycle variations in the interplanetary magnetic field. In M. Neugebauer (Ed.), *Solar Wind Five* (pp. 323–331). Washington, DC: Rep. NASA Conf. Publ. 2280.
- Sorriso-Valvo, L., Carbone, V., Giuliani, P., Veltri, P., Bruno, R., Antoni, V., & Martinez, E. (2001). Intermittency in plasma turbulence. *Planetary and Space Science*, *49*(12), 1193–1200. [https://doi.org/10.1016/S0032-0633\(01\)00060-5](https://doi.org/10.1016/S0032-0633(01)00060-5)
- Tindale, E., & Chapman, S. C. (2016). Solar cycle variation of the statistical distribution of the solar wind e parameter and its constituent variables. *Geophysical Research Letters*, *43*, 5563–5570. <https://doi.org/10.1002/2016GL068920>
- Tsubouchi, K., & Omura, Y. (2007). Long-term occurrence probabilities of intense geomagnetic storm events. *Space Weather*, *5*, S12003. <https://doi.org/10.1029/2007SW000329>
- Tsurutani, B. T., Echer, E., & Gonzalez, W. D. (2011). The solar and interplanetary causes of the recent minimum in geomagnetic activity (MGA23): A combination of midlatitude small coronal holes, low IMF B_z variances, low solar wind speeds and low solar magnetic fields. *Annales Geophysicae*, *29*, 839–849. <https://doi.org/10.5194/angeo-29-839-2011>
- Tu, C.-Y., & Marsch, E. (1995). *MHD Structures, Waves and Turbulence in the Solar Wind*. Dordrecht, Netherlands: Springer.
- Veselovsky, I. S., Dmitriev, A. V., & Suvorova, A. V. (2010). Algebra and statistics of the solar wind. *Cosmic Research*, *48*(2), 115–130. <https://doi.org/10.1134/S0010952510020012>
- Veselovsky, I. S., Dmitriev, A. V., & Suvorova, A. V. (2010). Lognormal, normal and other distributions produced by algebraic operations in the solar wind. In M. Maksimovic et al. (Eds.), *AIP Conference Proceedings 12th International Solar Wind Conference* (Vol. 1216, pp. 152–155). Saint-Malo, France: American Institute of Physics. <https://doi.org/10.1063/1.3395824>
- Viall, N. M., Spence, H. E., Vourlidas, A., & Howard, R. (2010). Examining periodic solar-wind density structures observed in the SECCHI heliospheric imagers. *Solar Physics*, *267*, 175–202. <https://doi.org/10.1007/s11207-010-9633-1>
- Vörös, Z., Leitner, M., Narita, Y., Consolini, G., Kovács, P., Tóth, A., & Lichtenberger, J. (2015). Probability density functions for the variable solar wind near the solar cycle minimum. *Journal of Geophysical Research: Space Physics*, *120*, 6152–6166. <https://doi.org/10.1002/2015JA021257>
- Wheatland, M. S. (2000). The origin of the solar flare waiting-time distribution. *Astronomy Journal Letter*, *536*(2), L109–L112. <https://doi.org/10.1086/312739>
- Wilk, M. B., & Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. *Biometrika*, *55*(1), 1–17. <https://doi.org/10.2307/2334448>
- Xu, F., & Borovsky, J. E. (2004). A new four-plasma categorization scheme for the solar wind. *Journal of Geophysical Research*, *120*, 70–100. <https://doi.org/10.1002/2014JA020412>