Draft version December 13, 2024
Typeset using LATEX twocolumn style in AASTeX631

De-Biasing Structure Function Estimates From Sparse Time Series of the Solar Wind: A Data-Driven Approach

Daniel Wrench [1] and Tulasi N. Parashar [1]

[1] Victoria University of Wellington
Kelburn, Wellington 6012, New Zealand

ABSTRACT

Structure functions, which represent the moments of the increments of a stochastic process, are essential complementary statistics to power spectra for analysing the self-similar behaviour of a time series. However, many real-world environmental datasets, such as those collected by spacecraft monitoring the solar wind, contain gaps, which inevitably corrupt the statistics. The nature of this corruption for structure functions remains poorly understood—indeed, often overlooked. Here we simulate gaps in a large set of magnetic field intervals from Parker Solar Probe in order to characterize the behaviour of the structure function of a sparse time series of solar wind turbulence. We quantify the resultant error with regards to the overall shape of the structure function, and its slope in the inertial range. Noting the consistent underestimation of the true curve when using linear interpolation, we demonstrate the ability of an empirical correction factor to de-bias these estimates. This correction, "learnt" from the data from a single spacecraft, is shown to generalize well to data from a solar wind regime elsewhere in the heliosphere, producing smaller errors, on average, for missing fractions > 25%. Given this success, we apply the correction to gap-affected Voyager intervals from the inner heliosheath and local interstellar medium, obtaining spectral indices similar to those from previous studies. This work provides a tool for future studies of fragmented solar wind time series, such as those from Voyager, MAVEN, and OMNI, as well as sparsely-sampled astrophysical and geophysical processes more generally.

1. INTRODUCTION

1.1. *Gaps in solar wind time series*

The solar wind is a supersonic plasma that continuously flows outward from the Sun (Parker 1958). As well as playing a key part in our understanding of space weather, the solar wind has a special role in astrophysics more broadly. It is the only system in which we can study *in situ*, using spacecraft, the cosmically ubiquitous phenomenon of weakly collisional plasma turbulence (Bruno & Carbone 2013; Matthaeus 2021). Unfortunately, as with countless other datasets in the earth and space sciences, time series of the solar wind are often plagued by data gaps. Telemetry constraints, instrument failures and data filtering/conditioning all result in time series of plasma parameters that are variably incomplete. For example, data from the twin Voyager spacecraft, which provide our only measurements of the outer heliosphere and interstellar medium, have daily gaps of 12-16 hours due to the limited communication with ground stations from such distances (Gallana et al. 2016; Cuesta 2020). Our record of the solar wind at Mars is also regularly contaminated by multi-hour gaps, during the periods in which the MAVEN spacecraft flies through the planet's magnetosphere (Azari et al. 2024). These and other significantly fragmented datasets are highlighted in Table 1.

The aforementioned causes of gaps in solar wind datasets mean that, generally speaking, we can treat these gaps as "missing at random".[1] While this reduces the potential bias, any significant discontinuity still presents a challenge for various analyses in heliophysics. For example, it hinders our ability to forecast space weather events (Kataoka & Nakano 2021; Smith et al. 2022), to understand the coupling of the solar wind with planetary magnetospheres (Magrini et al. 2017; Lockwood et al. 2019; Azari et al. 2024), or to

Corresponding author: Daniel Wrench
daniel.wrench@vuw.ac.nz

[1] I.e., the presence of gaps is unrelated to the variables of interest—but is related to external factors, such as the time of day in the case of periodic gaps (Little & Rubin 2019).

Dataset	Location	Typical % of magnetic field data missing	Reference
Helios 1 & 2	0.3-1 au	50-60%	Venzmer & Bothmer (2018)
OMNI	1 au	67% until 1995, 8% thereafter	Lockwood et al. (2019)
MAVEN	Mars orbit	60-80%	Azari et al. (2024)
Voyager 1 & 2	1-140 au	70%	Gallana et al. (2016)

Table 1. Description of solar wind datasets that are particularly affected by missing data. au = astronomical units (from the Sun). Note that OMNI is a compilation of data from a range of spacecraft at 1 au; the increase in availability after 1995 was due to the Wind spacecraft coming online.

study plasma turbulence using scale-dependent statistics (Gallana et al. 2016; Fraternale 2017; Dorseth et al. 2024), which in turn are necessary for accurate models of cosmic-ray propagation (e.g., Engelbrecht & Burger 2013; Engelbrecht et al. 2022). Therefore, it is crucial to investigate whether we can increase the amount of reliable information that can be extracted from such scientifically invaluable datasets. In this study we focus on the extraction of robust turbulence statistics. Here, being able to analyse long intervals is particularly important, as this allows us to sample as full a range of scales as possible (Dorseth et al. 2024). To this end, we investigate, and demonstrate a method of addressing, the effect of data gaps on the *structure function.*

1.2. *Structure functions and turbulence*

The pth order structure function at lag τ gives the pth moment of the distribution of increments at that lag, $\mathcal{P}(\Delta x_\tau)$, where $\Delta x_\tau(t) = x(t+\tau) - x(t)$ is the increment and $x(t)$ is the value of some (scalar or vector) variable x at time t (or equivalently, at a point in space $\mathbf{r}$)

$$\begin{aligned} S_p(\tau) &= \int_{-\infty}^{\infty} |\Delta x_\tau|^p \mathcal{P}(\Delta x_\tau) d(\Delta x_\tau), \\ &= \langle |\Delta x_\tau|^p \rangle, \end{aligned} \tag{1}$$

where angle brackets denote an ensemble average. For an ergodic process, this can be replaced with a time average

$$S_p(\tau) = \frac{1}{N(\tau)} \sum_{i=1}^{N(\tau)} |x(t_i+\tau) - x(t_i)|^p, \tag{2}$$

where $N(\tau)$ represents the sample size at a given τ. If either $x(t_i)$ or $x(t_i+\tau)$ is missing due to a data gap, the corresponding increment is excluded, reducing the effective sample size. In addition to assuming ergodicity, the calculation of the structure function assumes stationarity of the increments (Yaglom 2004). When converting from temporal to spatial lags, it is also assumed that Taylor's hypothesis is valid (Taylor 1938). These assumptions—ergodicity, stationarity, and Taylor's hypothesis—are commonly made in solar wind studies, though their validity varies depending on the context (Matthaeus & Goldstein 1982; Jagarlamudi et al. 2019; Klein et al. 2014; Isaacs et al. 2015).

The second-order structure function, $S_2(\tau)$, is a popular statistic in many fields for characterizing variation across time or space. This is in part due to the practical advantages it offers over other second-order scale-domain statistics that provide equivalent information. With a simpler construction than the power spectrum and less restrictive stationarity requirements than the autocorrelation function, $S_2(\tau)$ allows for easy identification of the range of the scales that contribute to the variation of a source (Yaglom 1957; Schulz-DuBois & Rehberg 1981; Dudok de Wit et al. 2013). Theoretically, $S_2(\tau) \to 2\sigma^2$ as $\tau \to \infty$, where σ^2 is the variance of the signal. The scale at which $S_2(\tau)$ flattens is approximately equal to the correlation length, a measure of the outer scale of the system.[2] The identification of these so-called "characteristic scales" is one of the key applications of $S_2(\tau)$ in astrophysics, where it is commonly employed to study the light curves of active galactic nuclei (e.g., Kozłowski 2016; De Cicco et al. 2022). In geostatistics, $S_2(\tau)$ is referred to as the variogram and is used as a model of spatial variation in the interpolation scheme known as kriging (Matheron 1963; Webster & Oliver 2007).

The structure function has special significance in the analysis of self-similar processes. For a process that exhibits fractal behaviour, we observe $S_p \propto \tau^{\zeta(p)}$, where $\zeta(p)$ is a straight line for monofractal behaviour and a nonlinear function for multifractal behaviour (Frisch 1995).[3] In the case of turbulence, the field in which structure functions were first introduced in 1941, classical theory predicts $\zeta(p) = p/3$ in the inertial range (Kolmogorov 1941; Frisch 1995). Under certain assump-

[2] Typically, the correlation length is calculated using the autocorrelation, $R(\tau)$. Under an assumption of weak stationarity, required for $R(\tau)$ but not $S_2(\tau)$, the two functions are related by $S_2(\tau)/\sigma^2 = 2[1 - R(\tau)]$.

[3] The scaling exponent ζ is related to the Hurst exponent H by $\zeta(p) = pH$.

tions[4], we can relate this power-law scaling to that of the power spectrum $E(k) \propto k^{-\beta}$ via $\beta = \zeta(2) + 1$ (Pope 2000), giving the famous $\beta = 5/3$ power law of turbulence (Kolmogorov 1941). This relationship has been exploited in multiple solar wind studies by converting the structure function into an "equivalent spectrum" (Chasapis et al. 2017; Chhiber et al. 2018; Roberts et al. 2022; Thepthong et al. 2024).

Therefore, statistical analysis of turbulence often involves fitting power laws to structure functions and comparing the corresponding scaling exponents with theoretical predictions. This includes their wide application to a range of astrophysical flows, including the interstellar medium (e.g., Boldyrev et al. 2002; Padoan et al. 2003), intracluster medium (e.g., Li et al. 2020; Gatuzz et al. 2023), and, as highlighted in the present work, the solar wind (e.g., Horbury & Balogh 1997; Bigazzi et al. 2006; Chen et al. 2012; Pei et al. 2016). As well as the aforementioned practical advantages, higher-order structure functions are particularly well-suited to probing increment distributions in finer detail.[5] In particular, the *intermittency* of turbulent fluctuations is of interest. Intermittency refers to a greater propensity for particularly large fluctuations, i.e., a heavy-tailed probability distribution. This can be directly quantified via the kurtosis of the distribution, which is given by the normalized fourth-order structure function $S_4(\tau)/S_2(\tau)$ (Frisch 1995). Although intermittency is a well-known phenomenon in turbulent flows, it is of particular interest in the solar wind, due to its role in understanding the sites and mechanisms of energy dissipation in weakly collisional plasmas (TenBarge & Howes 2013; Matthaeus et al. 2015; Chhiber et al. 2018; Bruno 2019).

1.3. *Current approaches to handling data gaps*

The power spectrum has received much more attention than the structure function with regards to the effect of gaps, due to the more immediate obstacle gaps pose to its calculation and the ubiquity of the power spectrum across science and engineering. We briefly review this work now. Spectral analyses of the solar wind typically handle small gaps in time series (around a few percent in length) using linear interpolation (e.g., Vršnak et al. 2007; Chen et al. 2020; Carbone et al. 2021). For larger gaps, the performance of various interpolation methods and alternative spectral estimators have been compared (Munteanu et al. 2016; Magrini et al. 2017). (For applications outside of heliophysics, see Carré & Porter (2010); Mao et al. (2024); Arévalo et al. (2012); Babu & Stoica (2010).) Using such techniques, accurate spectral estimation has been claimed for solar wind datasets with missing fractions of up to 50% (Dorseth et al. 2024), 70% (Gallana et al. 2016; Fraternale 2017; Fraternale et al. 2019) or even 80% (McKee 2020).

One prominent gap-handling technique is the Blackman-Tukey method, which calculates the power spectrum as the Fourier transform of the autocorrelation function (Blackman & Tukey 1958). The autocorrelation is a function of τ, which means it can be readily calculated from discontinuous time series. The closely-related structure function shares this advantage, which has facilitated confident use of the statistic when faced with small gaps (see, e.g., Horbury & Balogh (1997); Bigazzi et al. (2006), for solar wind studies, and Takahashi et al. (2000); Zhang et al. (2002); Seta et al. (2023); Mckinven et al. (2023) for astrophysical studies). More recently, structure functions have been claimed to be robust for much larger missing fractions. For example, in atmospheric physics, a variant of the structure function based on the Haar wavelet was recommended as a spectral estimator for sparse turbulent-like signals (Mossad et al. 2024). For solar wind data, a pair of studies have suggested that up to 70% data loss can have only a negligible effect on statistics derived from the structure function, which we discuss presently. Burger & McKee (2023) created a synthetic dataset with known spectral properties, and then decimated it according to Voyager gap distributions of 64% and 68% sparsity. The resultant spectral index and power in the inertial range, as well as the correlation time, deviated by no more than 5% from their true values, as shown in their Table 1. A similar result was also found for data gaps from the IMP and ACE spacecraft (Burger et al. 2022, Table 1). However, these results were only for a single synthetic dataset, and, as acknowledged by the authors, one would expect "somewhat" different results for different realizations of their turbulence simulation.

Fraternale et al. (2019), on the other hand, reported that both the amount and distribution of missing data in Voyager datasets makes computation of the structure function "nontrivial". They showed that the periodic gaps present in these time series lead to regular oscillations in the sample size, which in turn produce artefacts in time-domain statistics such as the structure

[4] Weak stationarity, zero mean dataset, frequency range from 0 to ∞, and $1 < \beta < 3$ (Emmanoulopoulos et al. 2010, Appendix B.)

[5] This said, higher-order spectra have been developed and are reported to be more adept than structure functions at handling the effects of non-stationarity and large-scale structures (Carbone et al. 2018).

function.[6] Proceeding to calculate the structure function without any interpolation, the authors took this behaviour into account by using a statistical significance threshold based on relative sample size when calculating the spectral index. The statistical convergence of structure functions affected by gaps has also been explored (Fraternale & Pogorelov 2021, Appendix B).

In an astrophysical context, Emmanoulopoulos et al. (2010) also challenged the assumption that structure functions are immune to missing data, as part of a wide-ranging critique of over-interpreting structure functions when studying the variability of blazars. In a brief qualitative analysis, it was shown that gaps severely affect structure function estimates in an unpredictable manner that is dependent on the specific time series (see their Figure 12). Therefore, they concluded that extensive simulation is necessary to account for this behaviour.

In order to inform strategies for more robust structure function estimation, we perform such simulation and thereby thoroughly test the resilience of structure functions to gaps. We also analyse the effect of linear interpolation, and investigate whether a simple correction can be made to "de-bias" its consistent underestimation. This provides an alternative to the suite of techniques developed by Fraternale (2017) and others for improving spectral estimates, as well as more "black-box" approaches, such as the neural network model by Wrench et al. (2022) that showed limited ability to predict solar wind structure functions.

As our ground truth, we use magnetic field measurements from the Parker Solar Probe and Wind spacecraft, as described in Section 2. In Section 3, we describe the extensive gap simulation of these intervals, followed in Section 4 by the results on a set of case studies and the overall statistical picture, as well as an application to Voyager intervals. Section 5 describes our conclusions, including limitations and future directions of this work.

2. DATA

The aforementioned investigations into the effects of gaps on spectral estimation used simulated time series, real solar wind intervals, or a combination of the two. Here, we restrict ourselves to real-world intervals, so as to avoid the simplifying assumptions when working with simulated data, such as Gaussian behaviour. We also use considerably more intervals than the other works cited.

We conduct the bulk of our analysis using data from Parker Solar Probe (PSP), a mission launched in 2018 to study the origins of the solar wind by flying very close to the Sun (less than 10 solar radii at closest approach). This data provides us with long, continuous time series required to perform comprehensive gap simulation for a range of turbulence realizations. Specifically, we obtain the vector time series of the magnetic field $[B_x, B_y, B_z]$, as measured by its fluxgate magnetometer instrument at a native cadence of 256 samples/second (Fox et al. 2016; Bale et al. 2016). We use data from the years 2019-2023.

In order to test the ability of our de-biasing algorithm to generalize to turbulence in different regions of the heliosphere, we also construct a test set using data from Wind. Wind is a spacecraft situated at the L1 Lagrange point that has been continuously measuring the near-Earth solar wind since May 2004, and has contributed significantly to our understanding of turbulence in the solar wind (e.g., Woodham et al. 2018; Verdini et al. 2018; Wilson III et al. 2021). We use data collected by the Magnetic Field Experiment (Lepping et al. 1995) at a native cadence of 11 samples/second during the period May-December 2016.

Finally, we apply our de-biasing algorithm to Voyager data as an example of how it could be used in practice. We take two highly fragmented, approximately week-long intervals from Voyager 1: one from 118 au in the inner heliosheath with 65% missing data, and one from 154 au in the local interstellar medium with 86% missing data. Both are measured at native cadence of 1 sample every 48 seconds (0.021 samples/second).

Turbulence parameters such as the correlation time (Cuesta et al. 2022) and the magnetic fluctuation amplitude (Chen et al. 2020) can show a large degree of variability throughout the heliosphere (even within a localized region of space or a single orbit). Therefore, we perform two types of standardization on each interval, in order to improve the likelihood that our results generalize to different spacecraft in different regions of the heliosphere. Firstly, to account for the changing correlation time, we standardize each interval to a consistent number of correlation times over a consistent number of points. In this work, each interval is made to contain 10 correlation times across 10,000 points. Secondly, we standardize power levels (i.e., fluctuation amplitudes) by normalizing each interval to have a mean of 0 and variance of 1. In this sense, we have normalized both "axes" of each interval, with the aim of making our analysis system-agnostic. The steps in this process and an illustrated example are given in the Appendix. Following this procedure, we extract 10,731 standardized intervals of solar wind from PSP, and 165 from Wind.

[6] This behaviour was illustrated for the autocorrelation function by Gallana et al. (2016, Supplementary Information) and Dorseth et al. (2024).

3. METHOD

After standardizing the intervals, we compute the second-order structure function (hereafter SF) for the complete data, calling this our "true" SF, $S_2(\tau)$. This is computed using Eq. (2), with $p = 2$, using each $3{\times}10{,}000$ magnetic field vector. The SFs are calculated up to $\tau = 2000$ (i.e., two correlation lengths), so as to ensure that there is a large (pre-gapping) sample size at each lag, and that the SF covers the full inertial range. As an additional measure of estimation accuracy, we also compute the slope of the SF over the expected inertial range $50 \leq \tau \leq 500$, corresponding to 5-50% of a correlation length.

We then create 25 copies of each time series and introduce random gaps. As well as increasing the sample size for our statistical analysis, duplicating the intervals allows us to study the effect of different gap distributions on the same interval. We also simulate gap distributions that combine two types of missing data: uniformly distributed (individual) missing points, and contiguous "chunks" of missing points. The differing effects of the two types have been demonstrated by Emmanoulopoulos et al. (2010) and Dorseth et al. (2024). In both of these studies, uniformly distributed gaps were shown to merely add noise to the SF—an effect that is relatively easily ameliorated by linear interpolation. Contiguous gaps, on the other hand, such as the multi-hour gaps in the Voyager and MAVEN datasets, significantly distort the shape of the SF. For greater fidelity to spacecraft data, we apply both gap types to the same time series, but emphasize the contiguous case. Specifically, for a given total gap percentage (TGP), which we vary up to as much as 95%, we require that at least 70% of that amount must be removed via contiguous chunks. The exact proportions are chosen randomly.

From these gapped intervals, we compute our SF estimates, $\hat{S}_2(\tau)$. Firstly, we use Eq. (2) without any interpolation: this is our "naive" estimate, $\hat{S}_2^{\text{naive}}(\tau)$, following the common approach in the literature (e.g., Horbury & Balogh 1997; Bigazzi et al. 2006; Mckinven et al. 2023). Secondly, we apply linear interpolation, calling this our "LINT" estimate, $\hat{S}_2^{\text{LINT}}(\tau)$. This allows us to understand the behaviour of this very simple and common technique for handling gaps in time series.

The existing literature on quantifying the effect of gaps on SFs has been limited to the effect on derived statistics, i.e., the inertial range slope and correlation length (Burger & McKee 2023). For comparison with these results, we also compute the error of slope estimates in the inertial range. However, we also note the importance of accurately estimating the amplitude and shape of the entire SF curve. For example, the entire SF is required to compute the kurtosis. Therefore, we also evaluate the error of the overall SF. We quantify this error for a given τ using the percentage error, PE, defined as follows:

$$\text{PE}(\tau) = \frac{\hat{S}_2(\tau) - S_2(\tau)}{S_2(\tau)} \times 100. \tag{3}$$

The overall error of an SF estimate is given by the mean absolute percentage error, MAPE,

$$\text{MAPE} = \frac{1}{n_\tau} \sum_{\tau=1}^{n_\tau} |\text{PE}(\tau)|, \tag{4}$$

where n_τ is the number of lags over which the SF has been computed. Later in this paper, when creating our correction factor, we will define additional error metrics.

4. RESULTS

To begin, it is useful to recall the basic issue at hand: gaps in a time series reduce the sample size $N(\tau)$ of each lag distribution $\mathcal{P}(\Delta x_\tau)$. Unlike simply having a shorter interval, gaps result in different lags being depleted to different degrees, depending on the size and location of the gaps. Uniformly distributed gaps will tend to reduce $N(\tau)$ uniformly across all lags; whereas contiguous gaps, which we emphasize here, result in *uneven* depletion of the distributions. In either case, reduction in sample size affects the variance of said distribution, and therefore the value of the SF. Here we give examples of this effect on individual time series, before proceeding to the overall trends.

4.1. *Effect of gaps: case studies*

Fig. 1 shows a case study of the effect of gaps on three standardized intervals from the Wind spacecraft.[7] For interval (1), with about one quarter of the data removed, the SF remains relatively unaffected for both $\hat{S}_2^{\text{LINT}}(\tau)$ and $\hat{S}_2^{\text{naive}}(\tau)$, which both remain smooth and close to the true curve. Looking more closely at $PE(\tau)$, we can see a contrast between the two estimators. $\hat{S}_2^{\text{LINT}}(\tau)$ has the worst error at small lags (maximum PE of -25% at $\tau = 1$), whereas $\hat{S}_2^{\text{naive}}(\tau)$ performs worst at large lags (maximum of 30% at $\tau \approx 800$). Ultimately, it is this worse performance at larger lags for interval (1) that makes the relative overall error of $\hat{S}_2^{\text{naive}}(\tau)$ about 10 times larger than that of $\hat{S}_2^{\text{LINT}}(\tau)$ (MAPE = 19.8% vs. 2.3%).

[7] We discuss case studies from our Wind test dataset, rather than our PSP training set, in order to later compare with the corrected versions.

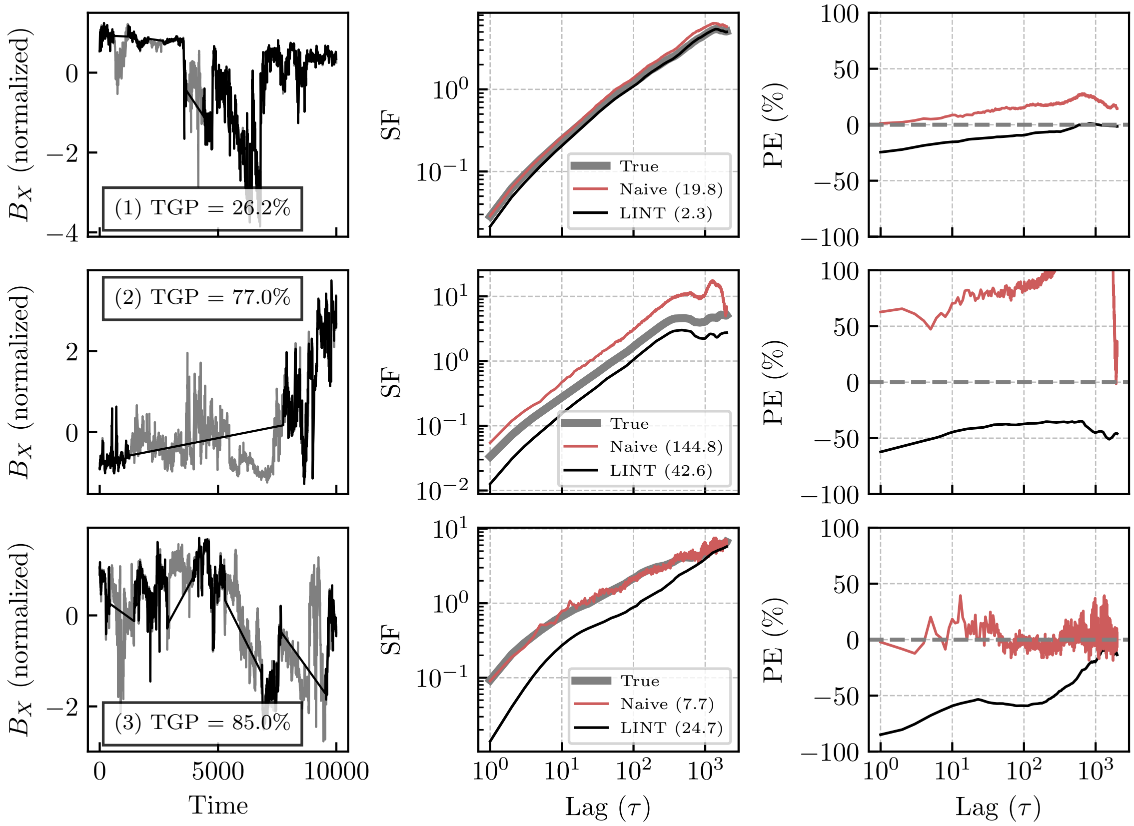

Figure 1. Case studies of the effect of increasing amounts of missing data on the structure function of three Wind intervals. The left-hand column shows the original complete interval (grey), and the interpolated gapped interval (black). Only one of the three vector components used in the calculation is shown, for visualization purposes. The middle column shows the SF from the complete interval ("true", thick grey), as well as the two estimates: $\hat{S}_2^{\mathrm{naive}}(\tau)$ ("naive", red), and $\hat{S}_2^{\mathrm{LINT}}(\tau)$ ("LINT", black). The mean absolute percentage error of each estimate is given in brackets. The percentage errors as a function of lag are given in the right-hand column.

For interval (2), with 77% of data missing, mostly via one large gap, we see the SF estimates start to diverge considerably from the true values. The naive approach significantly overestimates the SF, particularly at $\tau > 200$ where overestimation exceeds 100%. This overestimation is the result of removing a relatively quiescent segment of the original interval, leaving behind data with comparatively large fluctuations. $\hat{S}_2^{\mathrm{LINT}}(\tau)$, of course, has the opposite effect, once again dragging down the SF by smoothing out all fluctuations in the removed segment. In this example, we see greatest underestimation from $\hat{S}_2^{\mathrm{LINT}}(\tau)$ at $\tau = 1$ of about -60%, before reducing slightly to about -40% where it remains mostly steady. However, both estimators do match the true shape reasonably well, including the bend-over point at a few hundred lags. This example is in line with the findings of Burger & McKee (2023), who showed that the inertial range slope *can* remain relatively unaffected by gaps of more than 60%. However, the statistical analysis to follow will show that this is far from guaranteed.

Finally, for interval (3) with 85% missing data across four main chunks, $\hat{S}_2^{\mathrm{LINT}}(\tau)$ shows even greater underestimation at the smallest scales. The naive approach, meanwhile, shows noisy, high-frequency fluctuations, but mostly remains centered on the true curve. This gives it a much lower overall error than the previous case studies (MAPE = 7.7, compared with 19.8 for interval (1) and 144.8 for interval (2)). Interval (3) highlights that ignoring the gaps is a more unpredictable strategy than interpolation: sometimes the variances at each lag are relatively unaffected (albeit noisier); other times there is more noticeable distortion, as in interval (2).

(We note that the errors observed in Fig. 1 are not as extreme as those illustrated in Fig. 12 of Emmanoulopoulos et al. (2010), despite similar missing fractions. This is likely because their intervals were 1/5 the length of ours and therefore more affected by gaps.)

We now move to a statistical analysis to look at overall trends in behaviour across many different turbulence time series and gap distributions.

4.2. *Effect of gaps: statistical analysis*

The statistical analysis visualized in Fig. 2 shows the SF estimation errors from our large PSP training set. In addition to the PE for each individual SF estimate, we also show in the top row the mean percent error, or ensemble average, across all N SFs at each lag τ:

$$\mathrm{MPE}(\tau) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{PE}(\tau)_i. \tag{5}$$

Referring to this ensemble average, we see in Fig. 2a that the average error remains close to zero for all lags. This clearly demonstrates the unbiased nature of the naive SF estimator, such that its expected value $E[\hat{S}_2^{\mathrm{naive}}(\tau)] = S_2(\tau)$. Despite the often very large errors caused by gaps at large lags, often approaching 100% underestimation and exceeding 100% overestimation, there is no consistent bias away from the true value in either direction. Moreover, while greater errors tend to be correlated to TGP (note the error vs. color of the trendlines), there is unpredictable variability in this dependence, as well as in the amplitude of the error. The fact that high TGP can sometimes result in low errors could help explain the surprising results of Burger & McKee (2023), where TGPs of 64% and 68% resulted in negligible changes in derived statistics. However, we show here that this appears to be an atypical case.

Fig. 2b shows the errors from linearly interpolating the gaps. We see a very similar overall absolute error for the set of intervals (19.1 vs. 18.98), but a very different picture with regards to the direction of this error. Up to $\tau = 100$ (10% of a correlation time), this estimator shows consistent underestimation—$E[\hat{S}_2^{\mathrm{lint}}(\tau)] < S_2(\tau)$—that increases reasonably smoothly with TGP.

The decrease in the variance of the lag distributions $\mathcal{P}(\Delta x_\tau)$ is expected given that drawing straight lines across gaps is equivalent to smoothing the function and removing variation; the same phenomenon has been observed in the power spectrum (Fraternale et al. 2019). We see greater underestimation at small lags because, in spite of having a larger sample size, these distributions are much more distorted by long periods of gaps. When τ is small, it is more likely that both of the values $x(t), x(t+\tau)$ in the difference $x(t) - x(t+\tau)$ occur on the same interpolated line, and the longer this line, the smaller this new difference will be. This results in a dramatic shift of the increment distribution $\mathcal{P}(\Delta x_\tau)$ towards the center and therefore an excessive decrease in the variance of this distribution. This further results in an underestimation of the SF at small lags.[8]

Above approximately lag 100, the ensemble average error gets closer to zero as we begin to see overestimation for some intervals at these larger lags, but the average bias remains negative. At this point, the correlation between TGP and error also becomes weaker.

Eq. (5) is a 1D average error. We also calculate a 2D error, as a function of lag and the gap percentage at a given lag, $\mathrm{GP}(\tau)$—as previously noted, this is dependent on both the size and location of gaps. We bin these two variables and calculate the mean PE of the $\hat{S}_2^{\mathrm{LINT}}(\tau)$ estimate in each bin $\mathcal{B}$:

$$\mathrm{MPE}(\mathcal{B}) = \frac{1}{N(\mathcal{B})}\sum_{\tau,\mathrm{GP}\,\in\,\mathcal{B}}\mathrm{PE}_{\mathrm{LINT}}(\tau,\mathrm{GP}), \tag{6}$$

where $\mathcal{B}$ is a 2D bin defined according to a range of τ and GP. Each bin is then colored according to its corresponding $\mathrm{MPE}(\mathcal{B})$, resulting in the heatmaps shown in Fig. 2c and d.

For the naive estimates, the almost blank heatmap in Fig. 2c reinforces the unbiasedness of simply ignoring the gaps. We only see a very small average error for $\tau > 800$ and $\mathrm{GP} > 80\%$. For the $\hat{S}_2^{\mathrm{LINT}}(\tau)$ errors in Fig. 2d, we see the minimum (near-0) average error at large lags and low missing % in the bottom right, which then becomes an increasingly negative error (underestimating the SF) at smaller lag and higher missing %, moving towards the top left of the figure. As already described, smaller lags see stronger underestimation as it is more likely that most of the increments are computed from interpolated segments.

4.3. *Computation of correction factor*

Noting the consistent bias observed in $\hat{S}_2^{\mathrm{LINT}}(\tau)$, we investigate whether this bias is in fact consistent enough that it can be utilised to correct any given estimate of the SF from a sparse time series. In other words, we test

[8] We also examined the error of two alternative estimators of the SF, Cressie-Hawkins and Dowd (Cressie & Hawkins 1980; Dowd 1984; Webster & Oliver 2007). These estimators, from geostatistics, are designed to be more robust to outliers and skew than the traditional estimator we use here (also known as the Matheron method-of-moments estimator). It was found that both of these estimators were in fact more sensitive to gaps, producing larger errors, and did not have the same unbiased property of the traditional estimator.

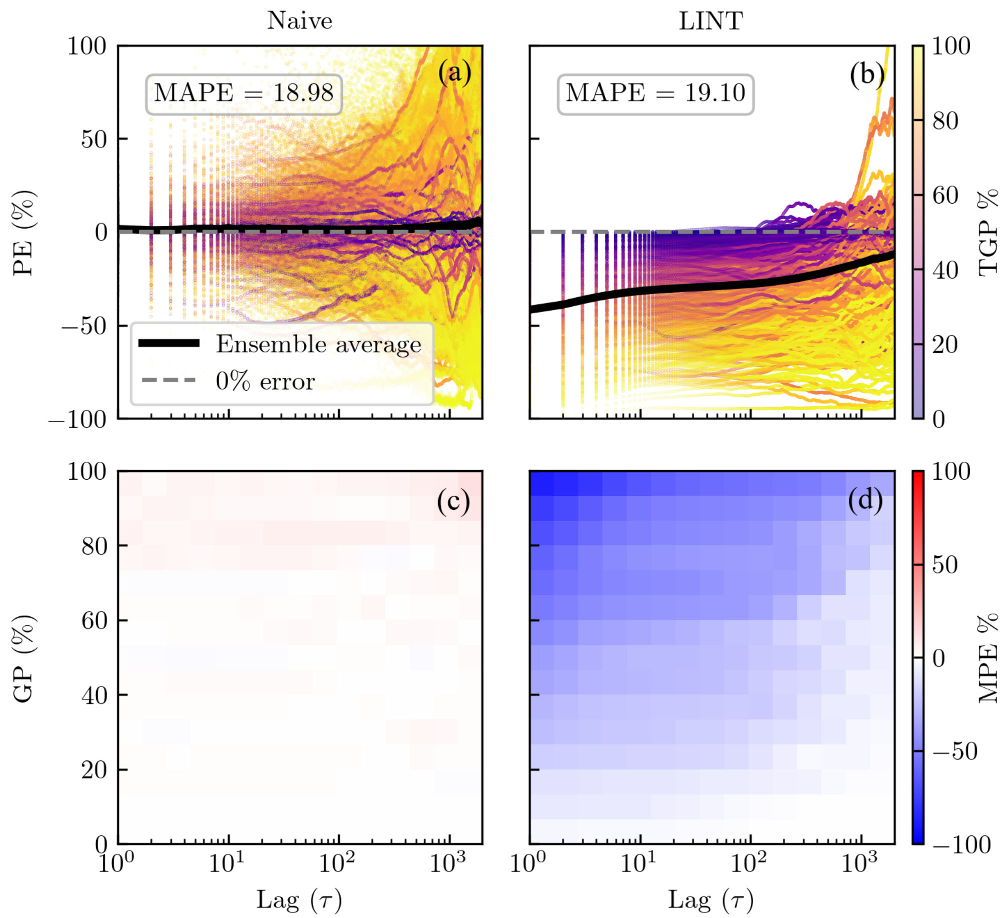

Figure 2. Two representations of relative error as a function of lag and missing fraction, as calculated from the PSP training set. Results for the $\hat{S}_2^{\text{naive}}(\tau)$ are given in (a) and (c) and $\hat{S}_2^{\text{LINT}}(\tau)$ in (b) and (d). Percentage error (PE) trendlines are given in (a) and (b) for a subset of 775 intervals, as in the third column of Fig. 1, colored according to the total gap percentage (TGP) of that interval. The black lines show an ensemble average of each trendline (see Eq. (5)). A 0% error line (dashed grey) is shown for reference. Subplots (c) and (d) show mean percentage error (MPE) for each combination of lag-specific gap percentage (GP, 15 linear bins) and lag (15 logarithmic bins)—see Eq. (6) for the calculation. These heatmaps are calculated from the full training set of ≈250,000 intervals.

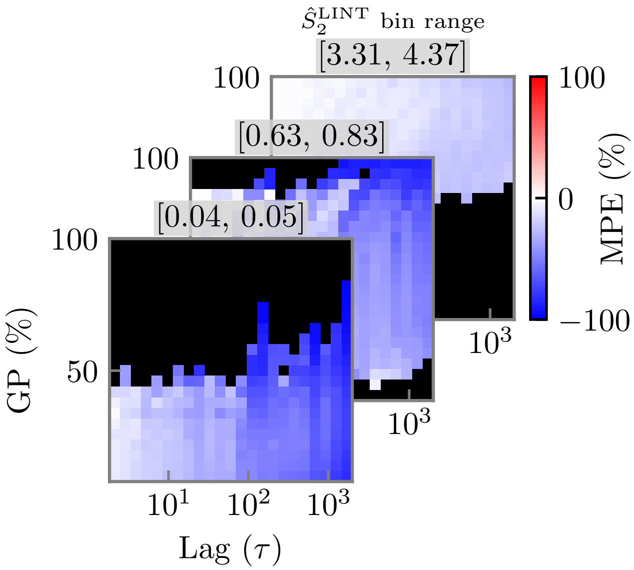

Figure 3. Three slices of the 3D "error cube" used for calculation of the final correction factor α using Eq. (7). The 2D error heatmaps of the $\hat{S}_2^{\mathrm{LINT}}(\tau)$ estimator, as given in Fig. 2d, are now additionally computed across 25 bins of the estimated SF value by said estimator. Bin ranges are given in square brackets; the full correction factor uses 25 such bins across all 3 dimensions of GP, lag, and $\hat{S}_2^{\mathrm{LINT}}(\tau)$. Black regions indicate unsampled bins. The shifting error as one moves up the SF (across bins) is clear.

the ability of a correction factor to "de-bias" $\hat{S}_2^{\mathrm{LINT}}(\tau)$. Currently, as shown in Fig. 2d, we calculate the average error introduced at a given lag (τ) by a given reduction in the sample size at that lag (GP). In order to improve the specificity of this correction, we also calculate error as a function of the value, or "power", of the estimated SF itself, $\hat{S}_2^{\mathrm{LINT}}(\tau)$: very small estimates of power will have significant negative errors, whereas larger estimates will have small or even positive errors. Therefore, we introduce a third variable, $\hat{S}_2^{\mathrm{LINT}}(\tau)$, into Eq. (6), resulting in a $25 \times 25 \times 25$ 'error cube'. The results for a subset of bins along this additional dimension can be seen in Fig. 3. It is clear looking across bins that the average error changes: as the values of $\hat{S}_2^{\mathrm{LINT}}(\tau)$ bins increase, the MPE for a given (τ, GP) gets closer to zero.

We then convert the MPE in each bin into a multiplicative correction factor, using the following equation:

$$\alpha(\mathcal{B}) = \frac{100}{100 + \mathrm{MPE}(\mathcal{B})}. \tag{7}$$

Any values of $\hat{S}_2^{\mathrm{LINT}}(\tau)$ in the test set which fall into this bin are then multiplied by α to attempt to return it to its "true" value. This gives us our corrected SF:

Method	SF MAPE (SD)	Slope APE (SD)
Naive	16.8 (16.8)	12.8 (14.1)
LINT	17.3 (17.8)	13.0 (13.9)
Corrected	8.7 (8.2)	9.0 (10.4)

Table 2. Performance of each method on the Wind test set. SF MAPE is the overall SF estimation error, quantified by the mean MAPE over all structure functions. Slope APE is the overall error in the estimated slope of the SF, quantified using the absolute percentage error (APE), averaged over all structure functions. SD = standard deviation.

$$\hat{S}_2^{\mathrm{corr}}(\tau, \mathrm{GP}) = \alpha(\mathcal{B})\hat{S}_2^{\mathrm{lint}}(\tau, \mathrm{GP}). \tag{8}$$

For example, a bin with MPE=-30% (i.e., underestimating the true SF by 30% on average) will have a corresponding $\alpha \approx 1.43$.

4.4. *Validation of correction factor on Wind data*

The procedure was first tested on a set of unseen PSP test data. The results when applied to the Wind test set are essentially the same. Hence, we discuss only the Wind results here to avoid repetition and given the importance of Wind data as a litmus test for the generalization power of the procedure. The Wind test set, once standardized and duplicated using the procedure in Section 2, consisted of 4125 intervals. The key "hyper-parameter" to tune in developing this correction factor was the number of bins, which represents a trade-off between sample size and specificity when it comes to the precise correction factor for each combination of variables. We trialled 10, 15, 20, and 25 bins. In a similar trade-off, we compared only binning on GP and τ, versus binning along a 3rd dimension of power.

The model was evaluated on the test set using two metrics, the mean MAPE of $\hat{S}_2(\tau)$, and the mean APE of the inertial range slope, across all the intervals in the test set.

Using these metrics, it was found that using all 3 variables and 25 bins to compute α gave the best performance. The overall results for this estimator are given in Table 2. This shows that $\hat{S}_2^{\mathrm{corr}}(\tau)$ indeed improves the overall SF estimation, and, to a lesser extent, its shape in the inertial range, as quantified by the error on the slope.

However, it is more useful to understand the performance of each estimator as a function of data sparsity. As before, we start by showing case studies. In Fig. 4, we perform the correction on the same three Wind intervals studied in Fig. 1. We also provide a measure of uncertainty in the corrections, using the variation in errors used to calculate α. Specifically, we obtain the

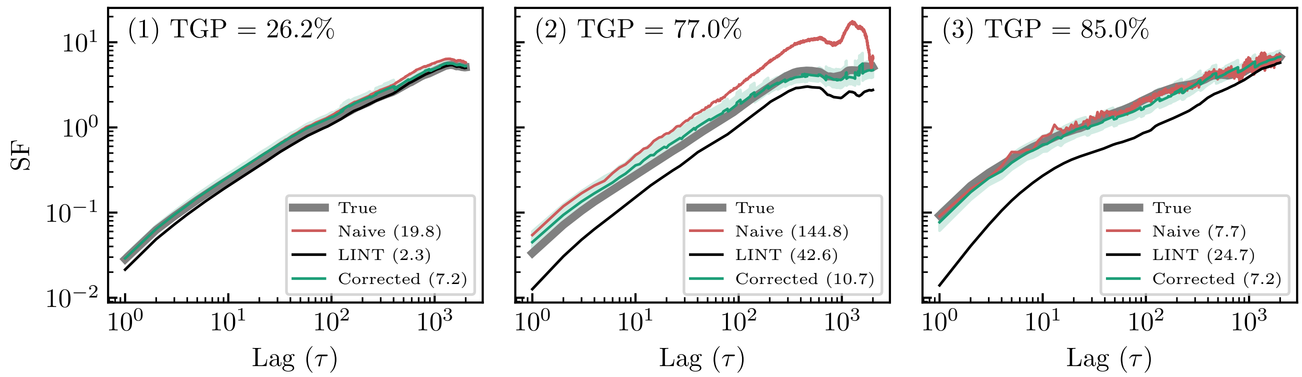

Figure 4. Case studies of applying the correction factor to $\hat{S}_2^{\mathrm{LINT}}(\tau)$ for three Wind intervals (the same intervals studied in Fig. 1). The shaded green region indicates $\hat{S}_2^{\mathrm{corr}}(\tau) \pm$ two standard deviations.

upper and lower limits of $\hat{S}_2^{\mathrm{corr}}(\tau)$by adding and subtracting two times the standard deviation of the PEs in each bin $\mathcal{B}$ from $MPE(\mathcal{B})$ in Eq. (7).

For interval (1), $\hat{S}_2^{\mathrm{corr}}(\tau)$ is an improvement on $\hat{S}_2^{\mathrm{naive}}(\tau)$ but is worse than $\hat{S}_2^{\mathrm{LINT}}(\tau)$, having a higher MAPE of 7.2 vs. 2.3. However, as we significantly increase the TGP, $\hat{S}_2^{\mathrm{corr}}(\tau)$ is superior to both of the original estimators. In interval (2), we can see the correction has achieved precisely its goal: taking advantage of the consistent underestimation but relatively accurate shape maintained by $\hat{S}_2^{\mathrm{LINT}}(\tau)$, the correction has essentially just translated the SF upwards to very closely match its true power level. While it has slightly over-corrected at small lags and under-corrected at large lags for this example, the confidence region overlaps with the true SF nicely. Finally, for interval (3) we again see good performance, accurately translating smaller lags upwards by a larger amount than large lags. It still maintains an edge over $\hat{S}_2^{\mathrm{naive}}(\tau)$ here, even though that estimate is unusually accurate for such a high TGP.

The test set errors for each method as a function of missing fraction is given in Fig. 5. As expected, all methods show increasing MAPE, and increasing variance in MAPE, with greater TGP. While noting this increasing variance, we fit polynomial regression lines to give an indication of the range of TGP each estimator is best suited for. The results suggest that $\hat{S}_2^{\mathrm{LINT}}(\tau)$ is superior to no gap handling at TGP<60%, beyond which $\hat{S}_2^{\mathrm{naive}}(\tau)$ typically has lower MAPE. This was seen in the case study of interval (3) in Fig. 1: because $\hat{S}_2^{\mathrm{LINT}}(\tau)$ shows more predictable underestimation, it can produce worse estimates than $\hat{S}_2^{\mathrm{naive}}(\tau)$ for large TGPs. But as shown by the boxplots, $\hat{S}_2^{\mathrm{naive}}(\tau)$ also has the largest errors. For TGP>25%, $\hat{S}_2^{\mathrm{corr}}(\tau)$ shows the lowest average error. We see greater advantage of this method with higher TGP, with the MAPE remaining below about 50% all the way up to TGP=95%. This is approximately half the maximum error observed for $\hat{S}_2^{\mathrm{LINT}}(\tau)$ TGP=95%, and drastically lower than the maximum error of $\hat{S}_2^{\mathrm{naive}}(\tau)$. Finally, using this MAPE metric, the corrected SF is also clearly superior to the neural network model studied previously by Wrench et al. (2022). While both that and the current method are data-driven, "learning" the errors associated with the SFs from gapped data, the current approach is much more transparent, and performs more intuitively with increasing missing data.

4.5. *Application to Voyager data*

As mentioned previously, the effect of gaps in Voyager data on turbulence statistics has had some attention in the literature, but has been limited to either the power spectrum (Fraternale 2017; Gallana et al. 2016; Fraternale et al. 2019), or, in the case of quantitative examination of the SF, a single gap simulation (Burger & McKee 2023). Given the success of our empirical, PSP-derived correction factor, we now apply it two highly sparse intervals from Voyager 1.

Prior to calculating the SFs using Eq. (8), the error heatmaps shown in Fig. 3 are first smoothed using a Gaussian filter. This reduces any discontinuities in $\hat{S}_2^{\mathrm{corr}}(\tau)$ that result from corrections for adjacent lags being somewhat different. Then, using the method developed by Thepthong et al. (2024), we convert $\hat{S}_2^{\mathrm{corr}}(\tau)$ into an equivalent spectrum, and calculate the spectral slope in the range 100-700 lags. This overlaps with the critical range for spectral recovery of $f \in [10^{-5}, 10^{-4}]$ Hz as identified by Fraternale & Pogorelov (2021).

The results for the inner heliosheath interval are given in Fig. 6. We observe a corrected SF that appears reasonable: we no longer observe the sharp dip in $\hat{S}_2^{\mathrm{naive}}(\tau)$ that appears to be an artifact; but we retain the approximate power of $\hat{S}_2^{\mathrm{naive}}(\tau)$—power which is sharply reduced in $\hat{S}_2^{\mathrm{LINT}}(\tau)$. From this, we obtain a very smooth

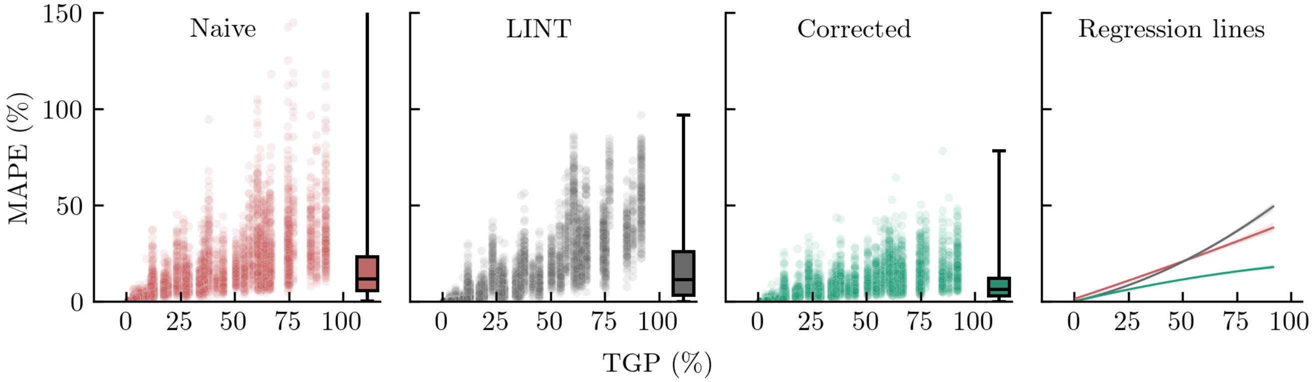

Figure 5. Error as a function of TGP for each of the three SF estimators for the Wind test set. On the right axis of each scatterplot is a boxplot showing the univariate distribution of errors: note some of the points for the $\hat{S}_2^{\text{naive}}(\tau)$ errors are outside the plotted area, hence the upper tail of its boxplot is obscured. The final panel shows order-2 polynomial regression lines fitted to each scatterplot. A 99% confidence region is overlain on each line, but is very narrow and therefore barely visible.)

equivalent spectrum, from which an inertial range slope of -1.5 is determined via a least-squares fit.

It is a similar story for the local interstellar medium interval (Fig. 7). The naive estimate has a large number of oscillations, particularly at lags between 1,000 and 10,000; these are absent from $\hat{S}_2^{\text{LINT}}(\tau)$ and consequently the corrected SF. Once again, the corresponding equivalent spectrum is very smooth. This allows us to confidently fit a power law, which in this case has a slope of -1.45.

In line with results of Gallana et al. (2016, Supplementary Information, Fig. 1), these examples show that interpolation significantly reduces the spurious oscillations inherited from the gap distribution, and our technique then further corrects for the decreased power. The spectral indices obtained here are approximately in line with Fraternale & Pogorelov (2021). Two primary differences between our approach and theirs are: i) they used a significantly longer interval. Our method can be trained and run on those kinds of intervals. However, we decided to test our learned weights directly, and hence have a shorter interval; ii) they decomposed the fluctuations into parallel and perpendicular directions based on the local magnetic field. We only work with the total magnetic spectrum. However, we also find the slopes in the interstellar medium to be slightly shallower than Kolmogorov (see their Figures 5 & 6).

5. DISCUSSION AND SUMMARY

Data gaps are best avoided when performing time series analysis, but this is not always possible when long intervals are required. As a result, our understanding of a variety of astrophysical and geophysical processes relies on extracting robust statistics from sparsely sampled time series. An area of research where this is particularly relevant is that of solar wind turbulence in the outer heliosphere and interstellar medium, where precious *in situ* data points are few and far between. In this study we have been able to provide new estimates of structure functions and spectral indices for these regions, based on a comprehensive examination of the statistical biases introduced by gaps.

We began by conducting an extensive gap simulation of a large number of solar wind intervals from Parker Solar Probe. In order to produce results that were as general as possible, i.e., not specific to one spacecraft or gap distribution, datasets were standardized and data points removed both randomly and in contiguous chunks. From these simulations, we demonstrated the starkly different effects of ignoring gaps vs. linearly interpolating them when calculating SFs, with regards to both the magnitude and direction of errors. As shown in Fig. 2, the "naive" approach of ignoring the errors, which is commonly thought to be satisfactory, is indeed an unbiased estimator. However, in this context, this simply means that there is no *statistical tendency* to over- or underestimate; the errors are still extremely unpredictable and can frequently be far in excess of 100%.

Linear interpolation, on the other hand, while very effective for small missing fractions, has a clear tendency to underestimate due to its smoothing effect. This effect is particularly damaging at small lags. This lag-dependence results in artificial scaling laws, some of the key parameters in turbulence analysis. However, we have shown that this bias is predictable enough that it can, to an extent, *be learnt and corrected for.* In a data-driven approach to the problem, we calculated the average estimation error from interpolated intervals as a function of lag, % missing (at each lag, therefore accounting for the specific distribution of gaps), and power, or estimated SF value. This was then used to de-

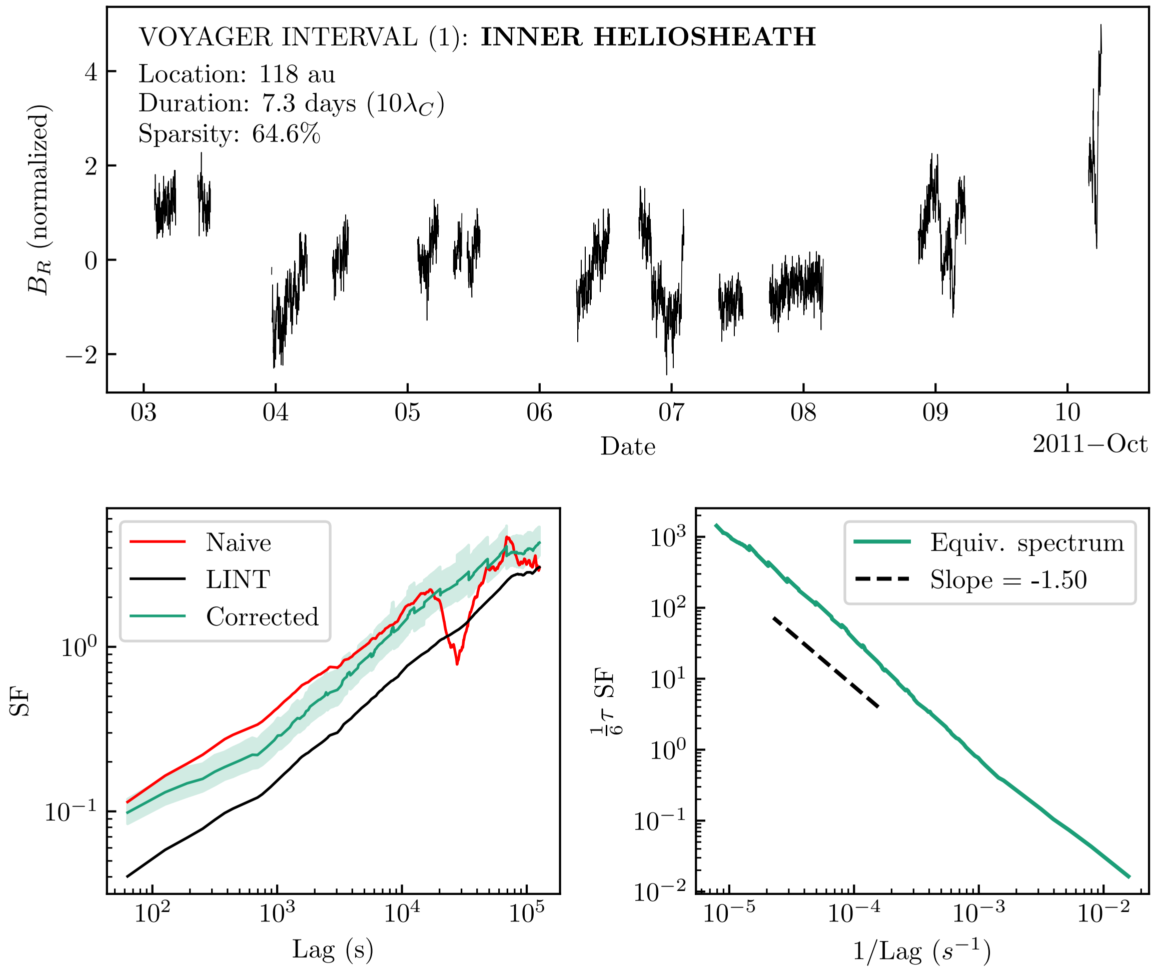

Figure 6. Top: normalized interval of Voyager magnetic field data from the inner heliosheath. Only one of the three vector components used in the calculation is shown, for visualization purposes. Bottom left: $\hat{S}_2^{\mathrm{naive}}(\tau)$, $\hat{S}_2^{\mathrm{LINT}}(\tau)$, and $\hat{S}_2^{\mathrm{corr}}(\tau)$ for the given interval. Bottom right: equivalent spectrum, calculated from $\hat{S}_2^{\mathrm{corr}}(\tau)$ according to the procedure given in Thepthong et al. (2024); formula given by y-axis label.

rive an empirical multiplicative correction factor for the interpolated SF. The improvement in estimation accuracy was proven on a test set from the Wind spacecraft, showing a typical reduction in error of about 50% compared with ignoring or interpolating gaps. Ultimately, we recommend our correction procedure over these other methods for missing fractions of greater than 25%.

The success of the learnt correction factor on the Wind test set gave us confidence to expect good results for other unseen datasets. Therefore, finally, we applied it to two intervals from Voyager, and converted these to equivalent spectra. This provides an alternative approach to spectral estimation from sparse solar wind intervals (for the other main approach, see Fraternale 2017). This brief application hints at the future potential of this work to unlock hitherto inaccessible scientific value from sparse datasets, of solar wind turbulence and beyond.

5.1. *Limitations and future work*

We do not simulate periodic gaps in this work; the effects of these have been shown elsewhere, at least for the power spectrum and autocorrelation function (Gallana et al. 2016), for which the distortion of the SF is very similar.

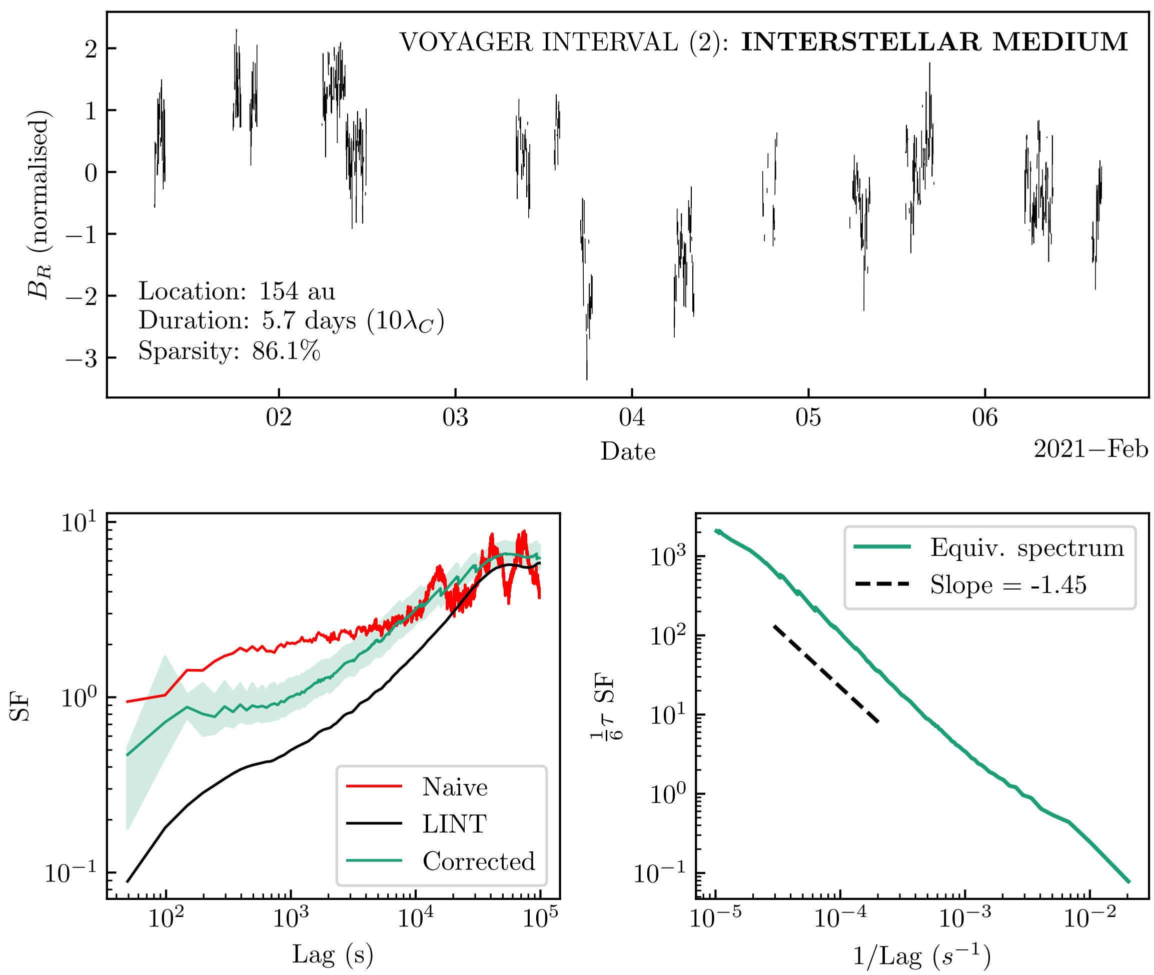

Figure 7. Top: normalized interval of Voyager magnetic field data from the local interstellar medium. Only one of the three vector components used in the calculation is shown, for visualization purposes. Bottom left: $\hat{S}_2^{\rm naive}(\tau)$, $\hat{S}_2^{\rm LINT}(\tau)$, and $\hat{S}_2^{\rm corr}(\tau)$ for the given interval. Bottom right: equivalent spectrum, calculated from $\hat{S}_2^{\rm corr}(\tau)$ according to the procedure given in Thepthong et al. (2024); formula given by y-axis label

While useful for making the results general to different spacecraft, the standardization we performed on the intervals based on the correlation time does somewhat limit these results. It means that the correction factor released with this paper can only be applied to time series comprising 10 correlation times spanning 10,000 points. However, as noted in the Data Product section below, this is a parameter in the code that can be easily changed and the corrections re-evaluated accordingly. Furthermore, we note that the standardization procedure uses correlation time estimations, which could be imprecise for originally sparse datasets. However, its use only for standardization to ≈10 correlation times gives us confidence that the error produced by this would be small.

We also only tested one interpolation method. Of course, more sophisticated interpolation techniques than linear interpolation are available, including stochastic types that are able to provide uncertainties (Azari et al. 2024) and capture higher-order statistics (Lübke et al. 2023). We did not address those here due to wanting to focus on thoroughly evaluating the commonly-used linear interpolation. A comparison with these more sophisticated models would be useful to investigate under a similar train-test paradigm.

DATA PRODUCT

In order to enable robust analysis of sparse solar wind time series for turbulence research, we share the derived correction factor with the community. We provide the specific values obtained from the PSP data and applied to Wind and Voyager, as well as a notebook demonstrating how to apply it to the Voyager dataset. Noting that these apply specifically to standardized intervals comprising 10 correlation lengths across 10,000 points, we also provide the codes used to produce the correction factor, thereby allowing for customized corrections for different length intervals. These are all provided on GitHub[9] under a 2-Clause BSD License and are archived in Zenodo (Wrench 2024).

ACKNOWLEDGMENTS

The authors thank the PSP, Wind, and Voyager instrument teams for the data and NASA GSFC's Space Physics Data Facility for providing access to it. The authors also wish to acknowledge the use of the Rāpoi high performance computing (HPC) cluster, provided by Victoria University of Wellington; and the New Zealand eScience Infrastructure (NeSI) HPC facilities, consulting support and training services. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme. Finally, we thank Mark Bishop and Jago Edyvean for their helpful comments on the manuscript.

AUTHOR CONTRIBUTIONS

DW: Conceptualization, Analysis, Drafting, Editing
TNP: Conceptualization, Supervision, Editing

APPENDIX

STANDARDIZATION PROCEDURE

Here we describe the steps of the interval standardization procedure that was outlined in Section 2. An example of a raw and standardized interval is given in Fig. 8.

1. Take an interval of magnetic field measurements corresponding to a large number of correlation times, according to typical values for the correlation time from the literature. (For this we use the entire interval covered by each raw file, which conveniently contains about 40-50 correlation times for both PSP and Wind.)

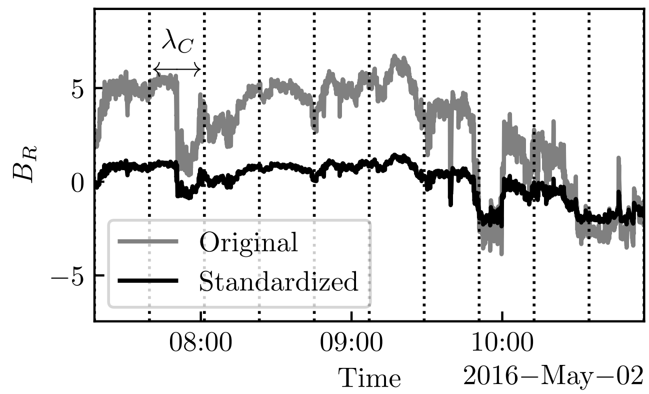

Figure 8. Example of the standardization process for an interval of Wind magnetic field data from 2016. The correlation scale for the entire 24-hour interval was calculated as 22min, using the integral method. The interval was then resampled to 1.3s, to correspond to 10 correlation times across 10,000 points, as indicated by the vertical dotted lines. This allowed for division into two sub-intervals of 10,000 points, the first of which is shown here. Each sub-interval was then standardized to have mean 0 and variance 1, giving the time series in black. (The correlation scale and final intervals used vector data, but only the radial component is shown here for demonstration purposes.)

2. Calculate the *local* correlation time of this entire interval using the integral method (Wrench et al. 2024).

3. Resample the interval such that 10,000 points corresponds to 10 of these correlation times.

4. Split the re-sampled interval into sub-intervals of length 10,000 (typically 2-4 of these per original interval).

5. If any sub-interval has more than 1% missing data, discard it. Otherwise, fill in any gaps with linear interpolation, such that each sub-interval is 100% complete.

6. Normalize the sub-intervals to have a mean of 0 and variance of 1.

7. Calculate the SF from each sub-interval as described in Section 3. We note that the above procedure results in normalized structure functions, as have been used by Chen et al. (2012).

[9] `sf_gap_analysis` codebase: https://github.com/daniel-wrench/sf_gap_analysis.

REFERENCES

Arévalo, P., Churazov, E., Zhuravleva, I., Hernández-Monteagudo, C., & Revnivtsev, M. 2012, Monthly Notices of the Royal Astronomical Society, 426, 1793

Azari, A., Abrahams, E., Sapienza, F., et al. 2024, Journal of Geophysical Research: Machine Learning and Computation, 1, e2024JH000155

Babu, P., & Stoica, P. 2010, Digital Signal Processing, 20, 359, doi: 10.1016/j.dsp.2009.06.019

Bale, S., Goetz, K., Harvey, P., et al. 2016, Space science reviews, 204, 49

Bigazzi, A., Biferale, L., Gama, S. M. A., & Velli, M. 2006, The Astrophysical Journal, 638, 499, doi: 10.1086/498665

Blackman, R. B., & Tukey, J. W. 1958, Bell System Technical Journal, 37, 185

Boldyrev, S., Åke Nordlund, & Padoan, P. 2002, The Astrophysical Journal, 573, 678, doi: 10.1086/340758

Bruno, R. 2019, Earth and Space Science, 6, 656

Bruno, R., & Carbone, V. 2013, Living Rev. Solar Phys, 10, doi: 10.12942/lrsp-2013-2

Burger, R., & McKee, S. 2023, Advances in Space Research, 71, 4916, doi: 10.1016/j.asr.2023.01.033

Burger, R., Nel, A., & Engelbrecht, N. 2022, The Astrophysical Journal, 926, 128

Carbone, F., Sorriso-Valvo, L., Alberti, T., et al. 2018, The Astrophysical Journal, 859, 27

Carbone, F., Sorriso-Valvo, L., Khotyaintsev, Y. V., et al. 2021, Astronomy & Astrophysics, 656, A16

Carré, J., & Porter, E. K. 2010, arXiv preprint arXiv:1010.1641

Chasapis, A., Matthaeus, W. H., Parashar, T. N., et al. 2017, The Astrophysical Journal Letters, 844, L9, doi: 10.3847/2041-8213/aa7ddd

Chen, C., Bale, S., Bonnell, J., et al. 2020, The Astrophysical Journal Supplement Series, 246, 53

Chen, C. H. K., Mallet, A., Schekochihin, A. A., et al. 2012, The Astrophysical Journal, 758, 120, doi: 10.1088/0004-637X/758/2/120

Chhiber, R., Chasapis, A., Bandyopadhyay, R., et al. 2018, Journal of Geophysical Research: Space Physics, 123, 2018JA025768, doi: 10.1029/2018JA025768

Cressie, N., & Hawkins, D. M. 1980, Journal of the international Association for Mathematical Geology, 12, 115

Cuesta, M. E. 2020, Master's thesis, University of Delaware

Cuesta, M. E., Parashar, T. N., Chhiber, R., & Matthaeus, W. H. 2022, The Astrophysical Journal Supplement Series, 259, 23, doi: 10.3847/1538-4365/ac45fa

De Cicco, D., Bauer, F. E., Paolillo, M., et al. 2022, Astronomy & Astrophysics, 664, A117

Dorseth, M., Perez, J. C., Bourouaine, S., Palacios, J. C., & Raouafi, N. E. 2024, A&A, 689, A117, doi: 10.1051/0004-6361/202449869

Dowd, P. 1984, in Geostatistics for Natural Resources Characterization: Part 1 (Springer), 91–106

Dudok de Wit, T., Alexandrova, O., Furno, I., Sorriso-Valvo, L., & Zimbardo, G. 2013, Space Science Reviews, 178, 665

Emmanoulopoulos, D., McHardy, I. M., & Uttley, P. 2010, Monthly Notices of the Royal Astronomical Society, 404, 931, doi: 10.1111/J.1365-2966.2010.16328.X

Engelbrecht, N., & Burger, R. 2013, The Astrophysical Journal, 772, 46

Engelbrecht, N. E., Effenberger, F., Florinski, V., et al. 2022, Space Science Reviews, 218, 33

Fox, N. J., Velli, M. C., Bale, S. D., et al. 2016, Space Science Reviews, 204, 7, doi: 10.1007/s11214-015-0211-6

Fraternale, F. 2017, PhD thesis, Politecnico di Torino, doi: 10.6092/polito/porto/2687873

Fraternale, F., & Pogorelov, N. V. 2021, The Astrophysical Journal, 906, 75

Fraternale, F., Pogorelov, N. V., Richardson, J. D., & Tordella, D. 2019, The Astrophysical Journal, 872, 40, doi: 10.3847/1538-4357/aafd30

Frisch, U. 1995, Turbulence: The Legacy of A. N. Kolmogorov (Cambridge University Press). https://books.google.co.nz/books?id=-JcjT4wYgfgC

Gallana, L., Fraternale, F., Iovieno, M., et al. 2016, Journal of Geophysical Research A: Space Physics, 121, 3905, doi: 10.1002/2015JA021830

Gatuzz, E., Mohapatra, R., Federrath, C., et al. 2023, Monthly Notices of the Royal Astronomical Society, 524, 2945, doi: 10.1093/mnras/stad2039

Horbury, T. S., & Balogh, A. 1997, Nonlinear Processes in Geophysics, 4, 185, doi: 10.5194/NPG-4-185-1997

Isaacs, J. J., Tessein, J. A., & Matthaeus, W. H. 2015, Journal of Geophysical Research: Space Physics, 120, 868, doi: 10.1002/2014JA020661

Jagarlamudi, V. K., de Wit, T. D., Krasnoselskikh, V., & Maksimovic, M. 2019, The Astrophysical Journal, 871, 68

Kataoka, R., & Nakano, S. 2021, Geophysical Research Letters, 48, e2021GL096275

Klein, K., Howes, G., & TenBarge, J. 2014, The Astrophysical Journal Letters, 790, L20

Kolmogorov, A. N. 1941, Proceedings of the USSR Academy of Sciences, 301

Kozłowski, S. 2016, The Astrophysical Journal, 826, 118

Lepping, R. P., Slavin, J. A., Ness, N. E., et al. 1995, Space Science Reviews, 71, 207

Li, Y., Gendron-Marsolais, M.-L., Zhuravleva, I., et al. 2020, The Astrophysical Journal Letters, 889, L1, doi: 10.3847/2041-8213/ab65c7

Little, R. J., & Rubin, D. B. 2019, Statistical analysis with missing data, Vol. 793 (John Wiley & Sons)

Lockwood, M., Bentley, S. N., Owens, M. J., et al. 2019, Space Weather, 17, 133, doi: 10.1029/2018SW001856

Lübke, J., Friedrich, J., & Grauer, R. 2023, Journal of Physics: Complexity, 4, 015005

Magrini, L. A., Domingues, M. O., & Mendes, O. 2017, Brazilian Journal of Physics, 47, 167, doi: 10.1007/s13538-017-0486-z

Mao, R., Lee, J. E., & Edwards, M. C. 2024, arXiv preprint arXiv:2410.05571

Matheron, G. 1963, Economic geology, 58, 1246

Matthaeus, W. 2021, Physics of Plasmas, 28

Matthaeus, W. H., & Goldstein, M. L. 1982, Journal of Geophysical Research, 87, 10347, doi: 10.1029/JA087iA12p10347

Matthaeus, W. H., Wan, M., Servidio, S., et al. 2015, "Philosophical Transactions of the Royal Society A", 373, doi: 10.1098/rsta.2014.0154

McKee, S. 2020, Master's thesis, North-West University

Mckinven, R., Gaensler, B., Michilli, D., et al. 2023, The Astrophysical Journal, 950, 12

Mossad, M., Strelnikova, I., Wing, R., & Baumgarten, G. 2024, Atmospheric Measurement Techniques, 17, 783

Munteanu, C., Negrea, C., Echim, M., & Mursula, K. 2016, Annales Geophysicae, 34, 437

Padoan, P., Boldyrev, S., Langer, W., & Åke Nordlund. 2003, The Astrophysical Journal, 583, 308, doi: 10.1086/345351

Parker, E. N. 1958, Astrophysical Journal, vol. 128, p. 664, 128, 664

Pei, Z., He, J., Wang, X., et al. 2016, Journal of Geophysical Research: Space Physics, 121, 911

Pope, S. B. 2000, Turbulent Flows (Cambridge University Press), doi: 10.1017/CBO9780511840531

Roberts, O. W., Alexandrova, O., Sorriso-Valvo, L., et al. 2022, Journal of Geophysical Research: Space Physics, 127, doi: 10.1029/2021JA029483

Schulz-DuBois, E., & Rehberg, I. 1981, Applied Physics A, 24, 323

Seta, A., Federrath, C., Livingston, J. D., & McClure-Griffiths, N. 2023, Monthly Notices of the Royal Astronomical Society, 518, 919

Smith, A., Forsyth, C., Rae, I., et al. 2022, Space Weather, 20, e2022SW003098, doi: https://doi.org/10.1029/2022SW003098

Takahashi, T., Kataoka, J., Madejski, G., et al. 2000, The Astrophysical Journal, 542, L105

Taylor, G. I. 1938, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 164, 476

TenBarge, J. M., & Howes, G. 2013, The Astrophysical Journal Letters, 771, L27

Thepthong, P., Pongkitiwanichakul, P., Ruffolo, D., et al. 2024, The Astrophysical Journal, 962, 37

Venzmer, M., & Bothmer, V. 2018, Astronomy & Astrophysics, 611, A36

Verdini, A., Grappin, R., Alexandrova, O., & Lion, S. 2018, The Astrophysical Journal, 853, 85, doi: 10.3847/1538-4357/aaa433

Vršnak, B., Temmer, M., & Veronig, A. M. 2007, Solar Physics, 240, 315

Webster, R., & Oliver, M. A. 2007, Geostatistics for environmental scientists (John Wiley & Sons)

Wilson III, L. B., Brosius, A. L., Gopalswamy, N., et al. 2021, Reviews of Geophysics, 59, e2020RG000714, doi: https://doi.org/10.1029/2020RG000714

Woodham, L. D., Wicks, R. T., Verscharen, D., & Owen, C. J. 2018, The Astrophysical Journal, 856, 49, doi: 10.3847/1538-4357/aab03d

Wrench, D. 2024, sf_gap_analysis, Zenodo, doi: 10.5281/zenodo.14431555

Wrench, D., Parashar, T. N., Oughton, S., Lange, K. D., & Frean, M. 2024, The Astrophysical Journal, 961, 182, doi: 10.3847/1538-4357/ad118e

Wrench, D., Parashar, T. N., Singh, R. K., Frean, M., & Rayudu, R. 2022, Space Weather, 20, doi: 10.1029/2022SW003200

Yaglom, A. M. 1957, Theory of Probability and its Applications, 2, 273

—. 2004, An introduction to the theory of stationary random functions (Courier Corporation)

Zhang, Y., Treves, A., Celotti, A., et al. 2002, The Astrophysical Journal, 572, 762