Inherent Length Scales of Periodic Mesoscale Density Structures in the Solar Wind Over Two Solar Cycles

L. Kepko¹, N. M. Viall¹, and K. Wolfinger²

¹NASA Goddard Space Flight Center, Greenbelt, MD, USA, ²Department of Physics, University of Colorado Boulder, Boulder, CO, USA

Abstract It is now well established through multiple event and statistical studies that the solar wind at 1 AU contains periodic, mesoscale \((L \sim 100-1,000 \, \text{Mm})\) structures in the proton density. Composition variations observed within these structures and remote sensing observations of similar structures in the young solar wind indicate that at least some of these periodic structures originate in the solar atmosphere as a part of solar wind formation. Viall et al. (2008, https://doi.org/10.1029/2007JA012881) analyzed 11 years of data from the Wind spacecraft near L1 and demonstrated a recurrence to the observed length scales of periodic structures in the solar wind proton density. In the time since that study, Wind has collected 14 additional years of solar wind data, new moment analysis of the Wind SWE data is available, and new methods for spectral background approximation have been developed. In this study, we analyze 25 years of Wind data collected near L1 and produce occurrence distributions of statistically significant periodic length scales in proton density. The results significantly expand upon the Viall et al. (2008, https://doi.org/10.1029/2007JA012881) study and further show a possible relation of the length scales to solar “termination” events.

Plain Language Summary The plasma and magnetic field in the solar atmosphere flows away from the Sun, filling interplanetary space. This plasma is called the solar wind, and it constantly bombards all of the planets in the solar system. The solar wind is composed of mesoscale structures—larger than scales where particle dynamics are important, but smaller than global scales—of increased density, and therefore pressure. A subgroup of mesoscale density structures is of order the size of Earth’s magnetosphere, and often quasi-periodic. These periodic density structures are an important driver of dynamics in Earth’s space environment. In this study, we examine the statistics of the size scales of these structures using 25 years, or approximately two solar cycles, of solar wind data measured by the Wind spacecraft. We confirm earlier work showing a persistence of particular length scales of the periodicities and find a possible relation of the length scales to the end of a Hale magnetic cycle. In addition to their driving of magnetospheric dynamics, periodic density structures are a tracer of solar wind formation. Their lengths scales and evolution are an important constraint of solar wind formation.

1. Introduction

The solar wind contains structures at vastly different scales, from small scale 1–2 min magnetic holes (Winterhalter et al., 1994) to large-scale coronal mass ejections and stream interaction regions (Kilpua et al., 2017; Richardson, 2018). There is a rich spectrum of structure between these two extremes, at “mesoscales,” which here we define as scale sizes \(L \sim 100\) to several 1,000 Mm, or equivalently time scales of \(t \sim\) a few minutes to several hours. Solar wind structures throughout this range of “mesoscale” sizes have been identified in Solar Terrestrial Relations Observatory (STEREO) white light remote sensing data down to the resolution of the imager (DeForest et al., 2018; Viall & Vourlidas, 2015). They have also been observed in situ, in the form of magnetic field flux rope structures as small as 50 Mm (Murphy et al., 2020), in plasma density at scales between 50 and 2,000 Mm (Stansby & Horbury, 2018), and in combinations of magnetic and plasma signatures (Borovsky, 2008; Matteo et al., 2019; Rouillard, Lavraud, et al., 2010; Rouillard, Kouloumvakos, et al., 2020; Sanchez-Diaz et al., 2019).

A subset of mesoscale solar wind structures are quasi-periodic proton density enhancements, termed periodic density structures (PDSs). They were initially discovered through event studies that showed a direct correspondence between magnetospheric pulsations in the mHz range (periodicities of a few minutes to a few hours) and a one-to-one correlation with discrete frequencies in the solar wind density observed in

©2020. 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.
the upstream solar wind (Kepko et al., 2002; Kepko & Spence, 2003). The apparent frequency of a PDS as it flows past Earth or an in situ spacecraft is related to the radial length scale of the structure as \( f_{\text{pds}} = \frac{V_{\text{sw}}}{L_{\text{pds}}} \), where \( f_{\text{pds}} \) is simply the inverse of the \( \Delta T \) between each density enhancement. Numerous event studies have observed direct links between the periodicities in solar wind density and periodicities in radar (Fenrich & Waters, 2008; Stephenson & Walker, 2002), ionospheric (Dyrdal et al., 2008), and ground magnetometer (Villante et al., 2007; Villante & Tiberi, 2016) observations, at frequencies from \( \sim 4 \) mHz. Viall, Kepko, et al. (2009) identified statistically significant frequencies observed in 11 years of Wind proton density data near L1 and 10 years of dayside GOES magnetospheric \( B_z \) data. They showed that both the solar wind and dayside magnetosphere contained recurrent, similar sets of observed frequencies between \( \sim 0.5 \) and 4.0 mHz, which lie in the Pc5–Pc6 frequency range. These mHz frequencies correspond to the smaller mesoscale structures, \( \sim 100–1,000 \) Mm, at nominal solar wind speeds.

These length scales are on the order of the dayside magnetosphere (\( \sim 80 \) Mm) and larger, and therefore quasi-statically drive magnetospheric pulsations through periodic dynamic pressure changes. Hence, even the smallest mesoscale solar wind structures are effective at creating a global magnetospheric response. It is this coherent, global magnetospheric response to solar wind structures at this size scale that motivates our lower limit definition for “mesoscale.” We note that the equivalent frequency of an 80 Mm structure at nominal solar wind speeds is \( \sim 4 \) mHz. At shorter length scales, the solar wind structures are smaller than the dayside magnetosphere, and the interaction is no longer quasi-static. Therefore, there is a general split between solar wind directly driven oscillations at \( f < 4 \) mHz, and internally supported oscillations, such as cavity mode or fieldline resonances, at around \( f > 4 \) mHz (Hartinger et al., 2013).

Since the initial papers describing the existence of PDSs in the solar wind, there have been several attempts to identify their source. A key measurement is the occurrence distributions (ODs) of statistically significant frequencies and length scales observed in solar wind proton density measurements. Viall et al. (2008) found statistically significant bands of periodic length scales, and Viall, Kepko, et al. (2009) found similar bands in frequency. These distributions of spectral peaks in solar wind density consists of three sources: in situ generated structures (e.g., via turbulence); “false positives” at a rate determined by the chosen confidence thresholds and appropriateness of the background spectral fit; and PDSs injected through the process of solar wind formation. The first two of these sources would generate a smoothly varying distribution of observed periodicities, rather than the recurrent sets found by Viall et al. (2008) and Viall, Kepko, et al. (2009), while the third could produce localized OD peaks. Although it is theoretically possible that there exists an MHD instability that could generate periodic structures in transit to 1 AU, for example, a slow mode wave (Hollweg et al., 2014), to date, there has been no published observations of such instabilities creating periodicities on mesoscales. Furthermore, Viall, Spence, et al. (2009) found fewer recurrent solar wind periodicities analyzing the data in time-frequency space than Viall et al. (2008) did analyzing the same data in length scale-wave number space. This suggests structures advecting with solar wind streams, rather than locally generated oscillations or waves at particular frequencies.

Multiple lines of evidence suggest that periodic solar wind density structures are tracers of solar wind formation. In situ observations show composition, magnetic field, and electron strahl changes that indicate magnetic reconnection effects that could only have occurred during solar wind release and acceleration (Kepko et al., 2016; Matteo et al., 2019; Viall, Spence, et al., 2009; ). Matteo et al. (2019), using Helios data, found anisotropic temperature changes within PDSs that are not observed near L1, consistent with solar formation followed by temperature isotropization while in transit. Remote imaging studies using the Solar Terrestrial Relations Observatory (STEREO)/Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI) white light instruments have identified PDSs in the solar corona as close as 2.5 solar radii, observed as they accelerate with the surrounding solar wind (DeForest et al., 2016, 2018; Viall et al., 2010; Viall & Vourlidas, 2015). Rouillard, Kouloumvakos, et al. (2020) confirmed the relationship between mesoscale density structures observed in images and in situ by tracking larger streamer blobs with embedded approximately hour-long structures from STEREO SECCHI to their impact with Parker Solar Probe. In short, it is now clear that the solar wind is often formed of quasi-periodic mesoscale plasma density structures released as a part of solar wind formation.

Three factors motivate this investigation. First, while previous studies used only 11 years of data, 25 years of Wind solar wind data are now available, which allows an examination of evolution of the recurrent length
scales as a function of two complete solar cycles. Second, the Wind SWE data that the Viall et al. (2008) and Viall, Kepko, et al. (2009) statistical studies analyzed have been reprocessed (Kasper et al., 2006), providing an opportunity to test the accuracy and precision of those previous results. Third, recent progress on techniques used to identify statistically significant spectral peaks has shown that there are limitations to using the AR(1) background assumption and suggests a different background assumption may be more appropriate (Vaughan et al., 2011).

2. Methods

2.1. Data Processing and Quality Checks

We follow the general process of data preparation and spectral analysis as the Viall et al. (2008) study. We used the proton number density and proton velocity measured by the Solar Wind Experiment (SWE) Faraday Cup onboard the Wind spacecraft (Ogilvie et al., 1995) to examine the characteristics of mesoscale PDSs between ~80 and 1,000 Mm observed over the full lifetime of Wind to this point, from 1995–2019. In the time since the Viall et al. (2008) study, Kasper et al. (2006) developed a new fitting technique to calculate separately the moments of the proton and Helium distributions from the Wind Faraday Cup data. This new data set, since it takes into account the bi-Maxwellian nature of the solar wind, provides a more accurate measure of the proton number density and velocity. The primary impact of that reprocessing on this study is that the velocity increased on average by a few percent, which increases the length scales by a few percent, and the proton density decreased slightly. These changes are slightly more pronounced when the velocity is high.

For continuity with and comparison to the Viall et al. (2008) study, we follow the same processing steps prior to the spectral analysis to produce length series segments 9,072 Mm in length, overlapping by 252 Mm. We first converted the time series of solar wind proton density, \( n(t) \), to a length series, \( L(t) \), by multiplying each time step by the radial velocity, \( v_r(t) \). Since each step has a different velocity, this produces an irregularly sampled series that is not compatible with Fourier analysis and must be interpolated to a fixed \( \Delta L \). Yet, due to the wide spread in solar wind speeds, resampling to a single common length step would lead to oversampling at low speeds, and undersampling at high speeds. We therefore produced two sets of interpolated, evenly sampled segments. Segments with \( v_r \geq 550 \text{ km/s} \) were termed “fast” segments, with \( \Delta L_f = 56.7 \text{ Mm} \), while those with \( v_r \leq 550 \text{ km/s} \) were termed “slow” segments, with \( \Delta L_s = 35.4 \text{ Mm} \). For slow wind, 9,072 Mm is approximately 6 hr of data at the median slow solar wind speed, and the 35.4 Mm \( \Delta L \) is approximately equivalent to the SWE instrument sampling rate (typically 90–100 s) converted to length. Similarly, for the fast wind 9,072 Mm is approximately 4 hr, and 56.7 Mm is the equivalent sampling rate multiplied by the median fast speed. While the Nyquist between between “fast” and “slow” segments is different, the combination of \( \Delta L \) and number of points in each segment keeps the spectral resolution the same. Note that the categorization of fast and slow data segments is not an attempt at a physics-based classification of solar wind type; it is well known that speed is not the best physics-based classification (Borovsky, 2012; Roberts et al., 2020; Zurbuchen et al., 2002). Rather, these two categories are only the result of the effective sampling rate of the data segments.

Figure 1 shows both a slow (panels a–c) and fast (panels d–f) segment of solar wind data comparing the original (blue) and reprocessed (red) SWE data as a time series, and both data sets converted into a length series (panels c and f). These segments are typical of other intervals in that they exhibit the very slight increase of a few percent in velocity in the reprocessed data. The reprocessed data also show differences in higher-frequency variations, particularly for the fast wind (see Figure 1d).

For each data segment, we imposed data quality requirements to minimize spurious spectral signals, and do not analyze segments that failed the data quality check. We required that the Wind spacecraft be located at least 50 Earth radii (\( R_E \)) upstream of Earth, to exclude any solar wind collected within or near Earth’s magnetosphere, or that could be contaminated with foreshock activity. This reduced the number of segments during the early part of the Wind mission, when it occasionally enters Earth’s magnetosphere. We remove single-point data spikes and interpolated over them. We excluded any segment that contained more than 10% flagged or missing data over the entire segment, or 3% consecutive flagged or missing data. Finally, we excluded segments that contained discontinuous jumps (e.g., shocks) in the number density, since this would introduce “ringing” in the spectra. To determine a discontinuous jump, we subtracted a third-order
polynomial fit to the data segment and discarded segments that contained changes in 5-point running averages that exceeded 3.7 standard deviations of the detrended median. The fraction of segments that passed these quality control checks is shown in Figure 2. There is a slight decrease in the number of segments that passed these checks using the reprocessed SWE data for the slow wind compared to the original data used by Viall et al. (2008).

**Figure 1.** Comparison of the original (blue) and reprocessed (red) solar wind data from Wind SWE for a representative segment, for both slow (left) and fast (right). Reprocessed data show slightly lower density (a and d), slightly higher velocity (b and e), and the high-frequency variations are of lower amplitude than the original data.

![Figure 1](image1.png)

**Figure 2.** The percentage of the slow and fast solar wind length series segments that passed the quality control checks and that were analyzed for periodic density structures. We also include the percentage of segments that passed these same quality checks in the original Viall et al. (2008) study. The differences are due to the reprocessed Wind SWE data.

![Figure 2](image2.png)
2.2. Spectral Analysis and Peak Detection

We perform spectral analysis on each segment that passed the quality checks. We identify statistically significant spectral speaks using an amplitude test and a harmonic \( F \) test. For the amplitude test, we calculate the spectra, estimate the background fit, then identify statistically significant peaks above this background. We use the segments in Figures 1c and 1f to demonstrate the process and present the results in Figure 3.

Estimation of the spectra relies on the multitaper method (MTM), in which multiple, orthogonal Slepian tapers are convolved with the data segment to provide multiple, independent estimates of the spectra (Thomson, 1982). While producing a robust spectral estimate, this technique reduces the effective frequency resolution of the data as a function of the number of tapers chosen, \( K \), to \( 2p f_R \), where \( f_R = 1/(N \Delta L) \) is the Rayleigh frequency and \( p = (K + 1)/2 \). In this study we used five Slepian tapers, leading to an effective resolution of \( 6f_R = 6.6 \times 10^{-4} \text{Mm}^{-1} \). We zeropad the data segments by a factor of 10 prior to calculating the spectral estimates. In Figures 3a and 3c we plot MTM spectra for the fast and slow length series segments shown in Figure 1, for both the original and reprocessed data. Note that the \( X \) axis is in units of wave number \( \text{Mm}^{-1} \), and we also list the equivalent length scale. Both the original and reprocessed data sets show similar spectral characteristics at the longer length scales (lower wave numbers) but differ slightly at the smaller length scales (higher wavenumbers); the differences are more pronounced in the fast wind spectra. These trends are generally persistent across all segments and are consistent with the reprocessed data having lower noise.

Viall et al. (2008), following Mann and Lees (1996), modeled the spectral background under the assumption that the observations \( x_t \), at point \( t \), followed an autoregressive AR(1) process, such that

\[ x_t = \rho x_{t-1} + \epsilon_t, \]

where \( \rho = 0.95 \). This model provides a reasonable fit to the data, and the residuals from the model can be used to estimate the background spectral density. The \( F \) test is then applied to the residuals to identify statistically significant peaks above the background. We have plotted an AR(1) background fit for both data sets, with the 95% confidence level. Peaks that simultaneously pass the amplitude and \( F \) test are marked with half circles for both original (blue) and reprocessed (red) data.

Figure 3. MTM spectra and \( F \) test for both the slow (left) and fast (right) segment shown in Figure 1, and for both the original (blue) and reprocessed (red) Wind data. We have plotted an AR(1) background fit for both data sets, with the 95% confidence level. Peaks that simultaneously pass the amplitude and \( F \) test are marked with half circles for both original (blue) and reprocessed (red) data.
In Figures 3b and 3d we show the test, an issue we discuss in the next section. The choice of the background noise model, while not affecting the false positives from the two tests are uncorrelated, requiring that a signal pass both tests is analogous to testing at the 99.75% confidence level, our confidence levels are determined relative to that background. AR(1) background fits and 95% confidence levels for the original and reprocessed data sets, for the fast and slow segments, are shown in Figures 3a and 3c, overlaid on the MTM amplitude spectra. The background AR(1) fit for both the original and reprocessed data are quite similar for the slow wind, with calculated values of $a = 0.836$ and $a = 0.846$, respectively. For the fast wind, however, the spectra and AR(1) fits are quite different, due to reduced high-frequency power in the reprocessed data, with $a = 0.792$ and $a = 0.883$ for the original and reprocessed data, respectively. For both fast and slow wind, the AR(1) background fits lie well above the background at shorter scales (higher wave number), suggesting AR(1) may not be a good background assumption. We return to this in the next section.

The determination of a significant spectral peak, in this example wave numbers that have spectral power that exceed the 95% confidence threshold, is complicated by two issues. First, by definition power spectrum and confidence levels produce false positives at the rate determined by the confidence thresholds (Mann & Lees, 1996; Thomson, 1982). That is, for each frequency tested for significance, for a 95% test, for example, there is a 5% probability of exceeding the threshold. These false positives would be randomly distributed in frequency and therefore could not produce the types of preferential ODs identified by Viall et al. (2008). To minimize these “false positives,” in addition to the amplitude test, we apply a second type of spectral test, the harmonic $F_t$ test, which is independent of the background fit (Mann & Lees, 1996). The amplitude test requires a signal to have strong power but does not explicitly test the discrete nature of the power enhancement. On the other hand, the harmonic $F_t$ tests for phase coherent signals but does not test the power contained in those signals. As in Viall et al. (2008) we require that a spectral peak pass both the narrow band (amplitude) and $F_t$ test simultaneously to be considered significant and counted in our statistics.

The precise value of the peak we identify is fixed to the maximum $F_t$ test frequency within the spectral amplitude band that exceeds the threshold. Because a peak has to pass both, independent, tests simultaneously at the 95% level, our confidence threshold in application is significantly higher than 95%. Assuming that the false positives from the two tests are uncorrelated, requiring that a signal pass both tests is analogous to testing at a 99.75% confidence threshold. The second issue in identifying significant spectral peaks is that the choice of the background noise model, while not affecting the $F_t$ test, affects the narrow band (amplitude) test, an issue we discuss in the next section.

In Figures 3b and 3d we show the $F_t$ test for the representative segments, and we indicate peaks that pass both the narrow band and $F_t$ test at the 95% level with half circles. Note that many peaks pass the harmonic $F_t$ test with little power and are therefore not identified as significant in this combined test. Similarly, there are several amplitude peaks that exceed the amplitude threshold, but not the $F_t$ test. For example, the amplitude peak at $L = 200$ Mm in the slow wind, while significant in terms of spectral amplitude, was not considered phase coherent by the $F_t$ test and therefore was not considered significant. Since the $F_t$ test is a test for phase coherence, our study likely undercounts solar wind signals that have significant power but are not precisely phase coherent. As such, results that use this technique should be considered a lower bound.

### 2.3. Background Estimation

The narrow band (amplitude) spectral test is a measure of the power of a discrete signal relative to a background spectra. The AR(1) process assumption (Equation 1) is widely used, since it is reasonable to expect a physical system to have memory. However, whether that memory takes the precise form of the AR(1) in any particular segment of solar wind data is currently impossible to know a priori. Indeed, Figures 3a and 3c show that the AR(1) does not fit the highest and lowest wave numbers well. We find this to be a persistent

\[ x(t_i) = a x(t_{i-1}) + \epsilon_i, \]

where $a$ is the degree of correlation between sequential data points and $\epsilon$ is random noise with zero mean (white noise). The limit of $a = 0$ produces a purely white noise spectrum, while larger values of $a$ produce more strongly red-noise data series. The analytical spectrum of (1) is

\[ S_{AR1}(f) = S_0 \frac{1-a^2}{1-2a \cos(\pi f f_N) + a^2}, \]

where $S_0 = \sigma^2/(1-a^2)$ is the average value of the power spectrum and $\sigma^2$ is the variance of the white noise. We fit (2) via least squares to the spectra computed using the MTM to produce an estimation of the background under the assumption of red + white noise, and confidence levels are determined relative to that background. AR(1) background fits and 95% confidence levels for the original and reprocessed data sets, for the fast and slow segments, are shown in Figures 3a and 3c, overlaid on the MTM amplitude spectra. The background AR(1) fit for both the original and reprocessed data are quite similar for the slow wind, with calculated values of $a = 0.836$ and $a = 0.846$, respectively. For the fast wind, however, the spectra and AR(1) fits are quite different, due to reduced high-frequency power in the reprocessed data, with $a = 0.792$ and $a = 0.883$ for the original and reprocessed data, respectively. For both fast and slow wind, the AR(1) background fits lie well above the background at shorter scales (higher wave number), suggesting AR(1) may not be a good background assumption. We return to this in the next section.

The determination of a significant spectral peak, in this example wave numbers that have spectral power that exceed the 95% confidence threshold, is complicated by two issues. First, by definition power spectrum and confidence levels produce false positives at the rate determined by the confidence thresholds (Mann & Lees, 1996; Thomson, 1982). That is, for each frequency tested for significance, for a 95% test, for example, there is a 5% probability of exceeding the threshold. These false positives would be randomly distributed in frequency and therefore could not produce the types of preferential ODs identified by Viall et al. (2008). To minimize these “false positives,” in addition to the amplitude test, we apply a second type of spectral test, the harmonic $F_t$ test, which is independent of the background fit (Mann & Lees, 1996). The amplitude test requires a signal to have strong power but does not explicitly test the discrete nature of the power enhancement. On the other hand, the harmonic $F_t$ tests for phase coherent signals but does not test the power contained in those signals. As in Viall et al. (2008) we require that a spectral peak pass both the narrow band (amplitude) and $F_t$ test simultaneously to be considered significant and counted in our statistics.

The precise value of the peak we identify is fixed to the maximum $F_t$ test frequency within the spectral amplitude band that exceeds the threshold. Because a peak has to pass both, independent, tests simultaneously at the 95% level, our confidence threshold in application is significantly higher than 95%. Assuming that the false positives from the two tests are uncorrelated, requiring that a signal pass both tests is analogous to testing at a 99.75% confidence threshold. The second issue in identifying significant spectral peaks is that the choice of the background noise model, while not affecting the $F_t$ test, affects the narrow band (amplitude) test, an issue we discuss in the next section.

In Figures 3b and 3d we show the $F_t$ test for the representative segments, and we indicate peaks that pass both the narrow band and $F_t$ test at the 95% level with half circles. Note that many peaks pass the harmonic $F_t$ test with little power and are therefore not identified as significant in this combined test. Similarly, there are several amplitude peaks that exceed the amplitude threshold, but not the $F_t$ test. For example, the amplitude peak at $L = 200$ Mm in the slow wind, while significant in terms of spectral amplitude, was not considered phase coherent by the $F_t$ test and therefore was not considered significant. Since the $F_t$ test is a test for phase coherence, our study likely undercounts solar wind signals that have significant power but are not precisely phase coherent. As such, results that use this technique should be considered a lower bound.

### 2.3. Background Estimation

The narrow band (amplitude) spectral test is a measure of the power of a discrete signal relative to a background spectra. The AR(1) process assumption (Equation 1) is widely used, since it is reasonable to expect a physical system to have memory. However, whether that memory takes the precise form of the AR(1) in any particular segment of solar wind data is currently impossible to know a priori. Indeed, Figures 3a and 3c show that the AR(1) does not fit the highest and lowest wave numbers well. We find this to be a persistent
characteristic of the AR(1) fit when applied to the solar wind number density data. In effect, this bias imposes a slightly higher or lower confidence threshold than 95% across the spectra and indicates that the solar wind may not be modeled well as an AR(1) process for the ∼6 hr windows we consider here.

The paleoclimatology community has studied the AR(1) background assumption extensively, where the choice of noise model impacts the ability to detect cycles in the stratigraphic record. In response to these concerns, Vaughan et al. (2011) suggest a bending power law (BPL) background spectrum fit

\[ S_{\text{BPL}}(f) = \frac{Nf^{-\beta}}{1 + (f/f_b)^{\gamma-\beta}} \]  

(3)

which has the AR(1) as a special case and performs well in mixed noise spectra. Here \( N \) is the normalization, \( \beta \) is the spectral slope index at low frequencies, \( \gamma \) is the spectral slope index at high frequencies, and \( f_b \) is the frequency at which the bend occurs. For low values of \( f_b \), the BPL reduces to a straight power law with spectral slope \(-\gamma\).

The BPL fit, and the 95% confidence level, is shown in Figures 4a and 4b in green for the same segments shown in Figure 3. Note how the BPL is a better representation of the background at both the higher and lower wave numbers compared to AR(1) (red). We also plot a straight power law (blue) with spectral slope, \( p \), for both slow and fast segments for reference. There is consistency in the identified peaks using the

Figure 4. A comparison of three different background assumptions for the solar wind intervals shown in Figure 1. Shown are an AR(1) (red), a BPL (green), and power law (blue). For clarity, we have not plotted the 95% confidence levels. Peaks that simultaneously pass the amplitude and F test at 95% are marked for the different fits. The spectral background model parameters are \( N = 24.33, \beta = -0.51, \gamma = 1.87, f_b = 1.8 \times 10^{-4} \) Mm\(^{-1}\) for BPL slow wind; \( p = -1.74 \) for PL slow wind; \( N = 0.2, \beta = 0.2, \gamma = 2.26, f_b = 3.2 \times 10^{-4} \) for BPL fast wind; and \( p = -2.1 \) for PL fast wind.
different background assumptions, with the BPL assumption producing fewer peaks in the slow wind segment. This tendency for BPL to identify fewer significant peaks than AR(1), particularly at lower frequencies, is a consistent feature across the entire 25-year study.

The BPL is flexible in that it allows for an AR(1) solution, a single power law, and a host of solutions in between. Since the BPL approximates the solar wind background spectra better than AR(1), and because it is more versatile than a straight power law, we utilize BPL as one of the two background assumptions we use for our statistical study. For consistency with Viall et al. (2008), we also run the analysis with an AR(1) background estimate.

2.4. Occurrence Distributions

We applied the data processing and spectral analysis methods described above to the reprocessed solar wind measured by the Wind spacecraft from 1995–2019. For each segment we determine statistically significant peaks that pass the amplitude and \( F \) tests simultaneously, for both BPL and AR(1) background assumptions. We create separate ODs of the statistically significant lengths (inverse wave numbers) identified using the AR(1) + \( F \) test and BPL + \( F \) test criteria. For each set, we compute ODs over overlapping, 3-year intervals, with bins of width \( 6f_R \), the effective resolution of the MTM with our choice of \( K = 5 \), stepping by \( 3f_R \) for each subsequent bin. The inverse of wave number is length, and the wave number resolution bins of \( 6f_R \) corresponds to 3.5 Mm near the Nyquist and 1,500 Mm near zero wave number. For each 3-year window, we applied the bootstrap technique (\( N = 500 \)) to estimate the uncertainty of local peaks on the histogram and calculated a median histogram, median fit (5 point moving mean), and standard deviation from these 500 instantiations.

To demonstrate this process we show the median histograms, representing an OD, for 1995–1998 for both the fast and slow solar wind in Figures 5a and 5b, with \( 2\sigma \) standard deviation bars determined via the bootstrap method (Efron & Tibshirani, 1993). Visually, these histograms exhibit locally enhanced counts for

Figure 5. Three-year occurrence distributions for 1995–1997 for the slow and fast solar wind calculated for both the AR(1) (red) and BPL (blue) spectral background assumptions. Vertical bars represent +2\( \sigma \) standard deviation. Length scales that are greater than 2\( \sigma \) above the median fit (dashed lines) are shown in thick lines in (a) and (c), where we have extended the significant length scale range by \( f_R/2 \) in either direction. The residual distributions, obtained by subtracting the median fits from the occurrence distributions, are shown in (b) for the slow and (d) for the fast wind. Circles denote points that exceed 2\( \sigma \).
particular length-scale bands, with strong correlation between the occurrence enhancements using the AR(1) and BPL spectral background fits. The residuals (Figures 5b and 5d) highlight the similarity in local occurrence enhancements between the AR(1) and BPL histograms, despite the differences in the overall shape of the ODs. We use the bootstrapped ODs to determine statistically significant occurrence enhancements as those points that are >2σ above the background fit. These are highlighted with circle in Figures 5b and 5d, and with thick lines in Figures 5a and 5c.

Importantly, although the AR(1) and BPL background models produce different overall shapes of the ODs, they produce similar residuals, and similar occurrence enhancements are identified as statistically significant with the bootstrap method for each. For the slow wind, the OD determined with the AR(1) assumption exhibits a steep slope on the short length scale (higher wave number) end, consistent with the examples shown in Figures 3 and 4. The BPL assumption does not exhibit this bias, which provides confidence for local occurrence enhancements within this region (between ~90 and 150 Mm). For example, there is a local occurrence enhancement identified in the BPL OD near 110 Mm as >2σ significant, on top of a relatively flat part of the distribution. In the AR(1) OD, this shows up as a relatively small local enhancement, and appears in the residual histogram as well, but is not significant at the 2σ level. In addition, the ODs produced with the BPL assumption identify ~50% fewer significant peaks than those with the AR(1) assumption. This trend is consistent throughout the 25 year interval and indicates that the BPL is likely a better approximation for the solar wind background spectra, with fewer false positive detections. Despite the difference between the AR(1) and BPL results in absolute counts, the relative amplitude of the enhancements in the OD is similar between the two background model assumptions.

3. Results

We ran the entire 25 year Wind SWE data set through the analysis process described in section 2. Figure 6 shows the percentage of analyzed segments that contained at least one statistically significant length scale for the two different fits, for both fast and slow wind, compared to the results of Viall et al. (2008).
enhancements (i.e., the persistent length scales) that are $>2\sigma$ above the OD with thick lines in Figure 7. For example, the histograms for 2017–2019 slow wind in Figure 7 show in the BPL histogram three clear peaks below 100 Mm, and two broad peaks near 130 and 160 Mm. The histogram derived from the AR(1) assumption shows the first two peaks below 100 Mm and the two broad peaks near 130 and 160 Mm, but at a reduced relative amplitude compared to the BPL histogram.

To compare between the two background assumptions, we plot the significant length scales identified in both the AR(1) and BPL-derived ODs as significant at the $2\sigma$ level as horizontal bars in Figure 8. Lengths that were identified concurrently in the ODs of both model fits are shown in Figure 8 as solid black bars; these same lengths are highlighted in the individual panels as darker shades of red and blue. Many of the ODs...
exhibit local enhancements at the smallest length scales, very near the Nyquist, and we shade those particular length scales lighter to emphasize they may not be statistically significant. Excluding these length scales below 80 Mm, we find for slow wind that 67% of the BPL lengths are contained in the AR(1) distributions, and 54% of the AR(1) are in BPL. For fast wind, excluding length scales <130 Mm, these are 79% and 88% for both BPL and AR(1), respectively. The primary differences are at the ends of the spectral range analyzed and follow the general pattern identified in the example shown in Figure 4. At the long length scale end (low wave number), fewer significant peaks were identified with the BPL (blue) background assumption, while at the short length scale end (high wave number), fewer peaks were identified with the AR(1) (red) background assumption. Another difference occurs in the early part of the mission, with AR(1) finding a band near 150 Mm in the slow wind that is not apparent in the BPL results.

The new results are consistent with the previous results of Viall et al. (2008) that covered the years 1995–2005 using the original Wind data. Figure 9 shows the concurrently identified significant length scales from Figure 8 with the AR(1) derived results from Viall et al. (2008). For the slow wind (Figure 9a), both studies identified significant lengths near 130 and 170 Mm, and an additional set near 330 Mm. The differences between the original and reprocessed data occur primarily in the first three rows, covering years 1995–2001, during the earliest portion of the Wind mission. The fast wind results compare very well to the previous Viall et al. (2008) results, with three sets of length scales near 100, 300, and 400–500 Mm detected in both the original and reprocessed Wind data. The slight shift to shorter length scales in the 80–500 Mm bands in the reprocessed data results is due to a reduced central peak in the OD in the reprocessed data compared to the original data.

4. Discussion

The histograms shown in Figure 7 represent ODs of significant length scales observed in the solar wind near L1 over two solar cycles. The overall shape of these distributions exhibits a consistent pattern across the full 25 years of Wind data (Figure 7). For the slow wind, the statistically significant length scales identified using the BPL background assumption exhibit comparatively few counts at the longer length scales (>300 Mm), and a broad peak near the center of the distribution (100–200 Mm). The AR(1)-derived histograms exhibit a steep slope at the smaller length scales, followed by a slow decline at the longer length scales. The histograms for the fast wind length scales show a similar, although less pronounced, trend. Recall that at long wave number, the bin width (6fR) becomes comparable to the length scales. Future work examining longer data segments is required to understand the nature of the shape of the OD over these longer (greater than ~500 Mm) length scales.

In addition to these overall trends, the ODs exhibit local enhancements of length scales identified as significant. These are highlighted in the OD histograms in Figure 7 and pulled out separately in Figure 8 as bars.
The contour plots (a) and (b) are the addition of the normalized (to the peak value) residuals for both the BPL and AR(1) derived occurrence distributions for the slow and fast wind. Red indicates areas of enhancement observed in both OD residuals, blue indicates areas where both found length scales significantly below the background fit, and green indicates regions near the background or areas where BPL and AR(1) were in disagreement. The bars superimposed on (a) and (b) are from Figures 8e and 8f and indicate length scales that exceeded the background by 2σ. The schematic (c) is a pictorial representation of (a) and (b) combined and includes the 3 year running average of sunspot number and locations of the terminator.

Figure 10. The contour plots (a) and (b) are the addition of the normalized (to the peak value) residuals for both the BPL and AR(1) derived occurrence distributions for the slow and fast wind. Red indicates areas of enhancement observed in both OD residuals, blue indicates areas where both found length scales significantly below the background fit, and green indicates regions near the background or areas where BPL and AR(1) were in disagreement. The bars superimposed on (a) and (b) are from Figures 8e and 8f and indicate length scales that exceeded the background by 2σ. The schematic (c) is a pictorial representation of (a) and (b) combined and includes the 3 year running average of sunspot number and locations of the terminator.
Figures 7 and 8 together provide evidence for persistent bands of significant periodic length scales. To highlight these trends, we have plotted colored contour plots, along with the normalized residuals from which these length scales were determined, in Figures 10a and 10b. The residuals here are the addition of the normalized OD residuals from the BPL and AR(1) background assumptions. The plotted values are \((OD_{BPL} - \hat{f}(t)_{BPL}) + (OD_{AR(1)} - \hat{f}(t)_{AR(1)})\), where OD is the 3 year OD and \(\hat{f}(t)\) is the OD fit for the two spectral background assumptions. Length-scale occurrence enhancements that were detected in both OD residuals would add together (red), while parts of the distributions that are less correlated would tend to zero (green).

Figures 10a and 10b show clear patterns of periodic length scales that evolve over the full 25 years of Wind SWE data. In the slow wind, \(L \sim 90\) Mm (VI), \(L \sim 130–140\) Mm (III), and \(L \sim 170–190\) Mm (II) are all observed for the majority of the 25 year data set, with some noticeable variations we discuss below. There are two smaller bands near \(L \sim 210\) Mm (IV) in the middle years and between 310 and 350 Mm in the later years, and a sloped band between 250 and 400 Mm (I) for the first half of the interval. An additional band appears near \(L \sim 120\) Mm in the BPL-derived histograms in Figure 7a but is not apparent in the AR(1)-derived histograms, likely because this region has a very strong slope; there is a similar effect with the \(L \sim 90\) Mm band (see Figure 7). For the fast wind there is an intermittent band between \(L \sim 200–220\) Mm (IV) and two bands (I and V) that are highly sloped in time, suggesting a solar cycle dependence. Band I decreases from 500 to 300 Mm over solar cycle 23, while Band V appears at 400 Mm near the start of Solar Cycle 24. These bands also appear in the “slow” wind results.

Figure 10c shows a pictorial summary of the significant length scale bands, derived by examining the combined bar plots and residual contours of Figures 10a and 10b, and using the additional information of the histograms in Figure 7 to provide visual guidance on persistence. Recall that the separation between fast and slow wind was mathematical, for the purposes of an even sampling rate, rather than the physics of the formation of solar wind of different speeds. For this reason, we have combined the significant length scales observed in the slow and fast wind together. We note that Bands I, IV, and V are observed in both fast and slow wind analysis, suggesting that the creation mechanism of PDSs is not strictly a “slow” (<550 km/s) wind phenomenon.

Many characteristics of the Sun, solar corona, and solar wind are correlated with solar cycle, so unraveling the specific nature of the correlation of PDSs with solar cycle is a topic for future work. Here we speculate on a likely connection. In general, the solar corona is hotter, and its magnetic topology increases in complexity, at solar maximum, as manifested in active regions and their underlying magnetic concentrations, sunspots. To the right of Figures 10a–10c, we show the gradual solar cycle change as measured by 3 year averages of the sunspot number, along with the more abrupt “terminator” events that are the end of a Hale magnetic cycle (McIntosh et al., 2015, 2019). The terminator events are observed as abrupt changes in the distribution of solar EUV brightness and occur when there is no more old cycle polarity flux left on the solar disk. Related, Schonfeld et al. (2017) showed that the amount of hot plasma (plasma greater than \(10^6\) K in the solar corona) abruptly increases at the terminator, due to an increased amount of hot plasma in active regions.

The length-scale bands that we find in this paper exhibit breaks that are associated more closely with terminators than with sunspot minimum, for example, Bands I and II at both ends and Band III for the termination event of solar cycle 22. Additionally, there is a gradual evolution of the characteristic length scales between termination events, most pronounced in Bands I and V. With data from only two, very different, solar cycles, we cannot draw definitive conclusions about the exact relationship between solar wind periodic length scales and the solar cycle, but the result suggests that a relationship exists. The precise details of this relationship would likely become more clear with the next solar cycle.

As reviewed in section 1, there is strong evidence that PDSs originate from the Sun and are associated with magnetic reconnection of plasma from closed-field regions. The evolution of periodic length scales with solar cycle could be the result of changes in the nature of the interchange reconnection that releases the plasma into the solar wind, due to the increase in complexity of the global magnetic topology (Antiochos et al., 2011). The association with the termination event could be the result of the reversal of the polarity of the leading edge of the new active regions. When the leading sunspot has the opposite polarity of the surrounding coronal hole, null point topologies can form out from decaying active regions (e.g., Mason et al., 2019). This magnetic topology is expected to have different interchange reconnection properties than when the
active region has a leading polarity that follows Hale's law. Alternatively, coronal temperature is correlated with solar wind speed, so it could also be that the hotter active regions that occur after the terminator event accelerate solar wind, and any embedded PDSs, differently.

While this study focused specifically on mesoscale structures measured at L1 that exhibit periodicity in density, many other studies have observed mesoscale structures in the solar wind that form at the Sun and advect to 1 AU. A general picture is emerging in which mesoscale structures that form through spatial structures that rotate (Borovsky, 2008, 2020) or time dynamics such as reconnection in the corona (M. J. Owens et al., 2018; M. Owens et al., 2020; Sanchez-Diaz et al., 2016, 2017, 2019; Stansby & Horbury, 2018) are an inherent part of solar wind formation (Viall & Borovsky, 2020).

In a series of papers, Rouillard et al. (2010); Rouillard, Lavraud, et al. (2010); and Rouillard et al. (2011) tracked larger mesoscale structures from their formation in the corona through the inner heliosphere using SECCHI HI images, all the way to their impact at the Earth. They identified the corresponding compositional and magnetic field variations inherent to the structures, which were retained out to 1 AU. This set of studies unequivocally demonstrated that large mesoscale structures created at the Sun survive to 1 AU with identifiable in situ signatures. More recently, Rouillard, Kouloumvakos, et al. (2020) tracked density structures through the STEREO COR2 and HI1 FOVs to their impact at Parker Solar Probe, where they observed a one-to-one correlation between the ~3–4 hr density structures observed remotely and the in situ Parker measurements. They showed that Parker measured additional sequences of small density peaks separated in time by approximately 90–120 min, suggestive of the types of periodic density enhancements at 90 min time scales that have been observed in situ at L1 (Kepko & Spence, 2003; Viall et al., 2008), near Mercury’s orbit with Helios (Matteo et al., 2019), and remotely with STEREO (Viall & Vourlidas, 2015). Many of these event studies exhibited still smaller substructures at tens of minutes (Kepko & Spence, 2003; Kepko & Viall, 2019; Matteo et al., 2019), in the range of the structures studied here. Several studies also found composition signatures, which could only have come from formation at the Sun (Kepko & Viall, 2019; Viall, Spence, et al., 2009). Recent work by Réville et al. (2020) demonstrated that PDSs associated with helmet streamers could be the result of the tearing mode instability at the base of the heliospheric current sheet. They argue that the larger, ~10–20 hr periodicities, as well as the ~1–2 hr periodicities that are observed, are all the result of the tearing mode. Finally, Murphy et al. (2020) demonstrated a distribution of mesoscale solar wind flux ropes observed at Mercury, with time scales of 2.5 min to 4 hr. They concluded that a portion of the distribution was likely related to PDS generation. These studies together demonstrate that the solar wind is often composed of mesoscale density structures and provide ample evidence that structures of order tens of minutes time scales and longer form with the solar wind and survive through the inner heliosphere, out to 1 AU.

Finally, this current study, which focused on the smaller end of the mesoscale range, demonstrated that at least some mesoscale structures are quasi-periodic and occur at repeatable sets of frequencies and/or length scales. We emphasize that these length scales represent PDSs that advect with the solar wind. In the rest frame of a spacecraft or planet, they would appear as a periodic density variations at a frequency determined by \( f_{\text{PDS}} = V_{\text{sw}}/L_{\text{PDS}} \). Statistically, for any particular year, the magnetosphere or a spacecraft would see a spectrum of equivalent frequencies determined by convolving the distribution of solar wind \( V_{\text{sw}} \) with the length scales identified in Figures 8 and 10c from that year. To zeroth order, we can estimate these frequencies using the median solar wind speed for “fast” and “slow” solar wind. These equivalent frequencies are listed at the bottom of Figure 10c. The equivalent frequencies of these structures fall in the few mHz range, which for the magnetosphere is considered the Pc5–Pc6 band. Previously, Viall, Kepko, et al. (2009) studied 11 years of Wind SWE data covering 1995–2005 for evidence of discrete frequency periodicities in the solar wind number density. They found that \( f = 0.7, 1.3 – 1.5, 2.0 – 2.3, \) and \( 4.7 – 4.8 \) mHz occurred most often over that 11 year interval. Figure 10c demonstrates that \( f = 1.4 \) mHz corresponds to Band I in the slow wind, \( f = 2.0 – 2.3 \) mHz corresponds to Band IV in the slow and Band I in the fast, and \( f = 4.7 – 4.8 \) mHz corresponds to Band VI in the slow wind.

Since these are periodic structures in solar wind density, they would periodically compress the magnetosphere via periodic dynamic pressure changes, and we would expect the magnetosphere to show these same sets of frequencies. In the same Viall, Kepko, et al. (2009) study, they also examined GOES magnetospheric magnetic field data for intervals when GOES was near the dayside magnetopause and found in the GOES data a similar set of frequencies to those found in the solar wind. In a direct comparison between Wind
and GOES, they found when a spectral peak was observed in the solar wind, that same peak was observed at GOES 54% of the time. Other statistical studies have similarly identified persistent bands of significant mHz frequencies (e.g., Chisham & Orr, 1997; Francia & Villante, 1997; Ziesolleck & McDermid, 1995). While originally attributed to global cavity modes (e.g., Harrold & Samson, 1992), we now know these ~4 mHz oscillations are largely driven by solar wind PDSs. Since these periodic length scales directly drive the magnetosphere, we would expect the spectrum of discrete mHz oscillations in the magnetosphere to vary year to year as the \( L_{PDS} \) vary. Since the \( L_{PDS} \) has a solar cycle dependence, this would mean the spectrum of discrete mHz waves in the magnetosphere would also have a solar cycle dependence, although the variability of the solar wind speed would produce broad, rather than narrow, enhancements. This slow year-to-year variability, and the distribution of solar wind speeds, can explain year-to-year changes in measured frequencies. In addition, Kepko and Viall (2019) showed that ambient PDSs in the slow solar wind were sometimes compressed and amplified by a faster solar wind stream from behind and that these amplified PDSs had an observable impact on radiation belt particles. These particular PDSs were observed with stream interaction regions, which are known to be important drivers of radiation belt flux enhancements.

5. Conclusions

This study provides further evidence that large portions of the solar wind plasma consist of periodic mesoscale structures, many of which are likely released via magnetic reconnection. Using 25 years of Wind solar wind number density data observed near L1, we have identified bands of periodic length scales that occur more often than others. In the rest frame of a spacecraft or Earth, these periodic mesoscale density structures would appear at frequencies determined by the length scales of the PDSs and the solar wind velocity. Each occurrence of a periodic length scale passed two independent spectral tests at the 95% level, and we tested each occurrence with two different background spectral models. We identify bands of occurrence enhancements that are persistent in time and are significant using both background spectral models (Figure 8c). Bands near \( L \sim 130–140 \) Mm and \( L \sim 170–190 \) Mm were evident in the slow wind, equivalent to frequencies of \( f \sim 3.0 \) and 2.3 mHz in the stationary frame, while bands near 230 and 300 Mm were observed in both the fast and slow segments, equivalent to \( f \sim 1.9 \) and 1.2 mHz, and \( f \sim 3.1 \) and 2.0 mHz, for the slow and fast wind, respectively. Longer length bands were observed between 300 and 300 Mm, decreasing in length over the course of Solar Cycles 22 and 23. The apparent frequencies of these lengths fall in the Pc5–Pc6 pulsation bands, which are known to be important for processes leading to radiation belt particle loss, diffusion, and acceleration (Elkington & Sarris, 2016). The evolution of these bands exhibited changes near solar “terminator” events marking the end of a Hale magnetic cycle (Figure 10), although this is a qualitative association and requires further work. Given the statistical bands of recurrent length scales in the solar wind, it may be possible in the future to produce a statistical model for these solar wind-driven discrete oscillations.

Finally, while our study separated “slow” and “fast” wind based on speed, this was driven by the mathematics of creating length series with a fixed sampling length. Without separating the speed in this manner, length segments corresponding to fast speed would have been undersampled, and slow wind segments would have been oversampled. Therefore, this approach is not suited for, nor designed for, determining how the bands relate to formation of different types of solar wind, nor can it determine whether different physical mechanisms create different bands of periodicities. Indeed, Figure 10 shows that some bands in the “fast” and “slow” wind overlap, indicating a common mechanism for those bands, independent of final wind speed. Future work includes combining our event list of periodicities identified over 25 years of Wind data with clustering analysis (e.g., Ko et al., 2018; Roberts et al., 2020) and studies of compositional changes. This work is already underway.

Data Availability Statement

All data used in this study were obtained from CDAWeb (https://cdaweb.sci.gsfc.nasa.gov/).

References

Antiochos, S. K., Miki, Z., Titov, V. S., Lionello, R., & Linker, J. A. (2011). A model for the sources of the slow solar wind. The Astrophysical Journal, 731(2), 112. http://iopscience.iop.org/0004-637X/731/2/112

Borovsky, J. E. (2006). Flux tube texture of the solar wind: Strands of the magnetic carpet at 1 AU. Journal of Geophysical Research, 113, A08110. https://doi.org/10.1029/2007JA012684

Acknowledgments

L. Kepko and N. Viall acknowledge the Heliophysics Internal Scientist Funding Model and the NASA Guest Investigator program for funding. K. Wolfinger was supported through the NASA internship program. We thank M. Stevens for helpful discussion regarding the Wind Faraday Cup (SWE) data.
