

Probability tails of wavelet coefficients of magnetometer records

P. Kokoszka,¹ I. Maslova,¹ J. Sojka,² and L. Zhu²

Received 17 October 2005; revised 12 January 2006; accepted 7 February 2006; published 8 June 2006.

[1] The ground-based magnetometer network has long been a powerful tool for monitoring and observing the variations of the currents flowing in the magnetosphere-ionosphere (M-I) system. The time series of magnetograms are nonstationary and their frequency behavior changes over time. They are therefore not amenable to traditional time domain or spectral (Fourier) analysis. In recent years, various new mathematical techniques have been developed to analyze magnetometer data and the wavelet technique has stood out as being particularly relevant. In order to correctly make statistical inferences based on wavelet analysis, the wavelet coefficient distributions of magnetograms must be examined. In this work, we apply the discrete wavelet transform to the 1-min magnetometer records and then use several statistical techniques to analyze the probability distributions of the wavelet coefficients. It is found that the distributions of these coefficients for both storm and quiet times are highly nonnormal and can be classified as being heavy tailed. This finding suggests that when applying statistical techniques to the wavelet coefficients of the magnetograms, one must make sure that these techniques are robust to large departures from Gaussianity manifested by the presence of heavy probability tails. It is also found that the tail indexes for storm times are on average smaller than those of quiet times, which reflects the stronger impulsive and nonstationary features in magnetometer data during storm times, and the shifts are most significant for the wavelet coefficients corresponding to physical scales of 4–8 min.

Citation: Kokoszka, P., I. Maslova, J. J. Sojka, and L. Zhu (2006), Probability tails of wavelet coefficients of magnetometer records, *J. Geophys. Res.*, *111*, A06202, doi:10.1029/2005JA011486.

1. Introduction

[2] The currents flowing in the magnetosphere-ionosphere (M-I) form a complicated multiscale system in which a number of individual current components connect and influence each other [Tsyganenko, 2000], and their variations are closely connected to various nonlinear dynamic M-I processes, including magnetic storms and substorms. The monitoring and observing of the variations of the M-I current system has long been an important tool helping us to understand the response of the M-I system to the solar wind driver and explore the electrodynamic processes in the magnetosphere and ionosphere. Among the various observational means, the global network of ground-based magnetometers stands out with unique strengths [Friedrich *et al.*, 1999]. The recorded data are in continuous form covering all UTs and seasons, and almost all electrodynamic processes in the M-I system directly or indirectly leave their traces in the magnetograms. On the other hand, the capability of simultaneously observing many nonlinear M-I electrodynamic

processes by the magnetometer leads to magnetogram records that contain multiple-scale information with complex behavior. An example of a magnetogram is displayed in Figure 1 which shows the relative horizontal component of the magnetic field disturbance, measured at Boulder, Colorado, during a very strong storm which took place between 30 March and 2 April 2001. Owing to the multiscale and nonlinear nature of the M-I current system, the magnetograms are nonstationary and quite impulsive, and their frequency spectrum changes with time, sometimes quite abruptly. Therefore the traditional Fourier analysis, which decomposes signals into infinite sine and cosine components, is not well suited to the analysis of the magnetograms. Windowed Fourier analysis may offer some help, but the signals are still assumed to contain a relatively constant frequency spectrum within windows. In recent years, there has been an increasing interest in using wavelets to analyze various nonlinear geophysical data sets with time-dependent spectral characteristic, including the magnetometer data [e.g., Lui and Najmi, 1997; Wei *et al.*, 2004; Haldoupis *et al.*, 2004; Krankowski *et al.*, 2005].

[3] Most of these studies have not relied on statistical methodology which requires some information about the distributional properties of the wavelet coefficients. In particular, very different statistical methods are required depending on whether the wavelet coefficients have a

¹Department of Mathematics and Statistics, Utah State University, Logan, Utah, USA.

²Center for Atmospheric and Space Science, Utah State University, Logan, Utah, USA.

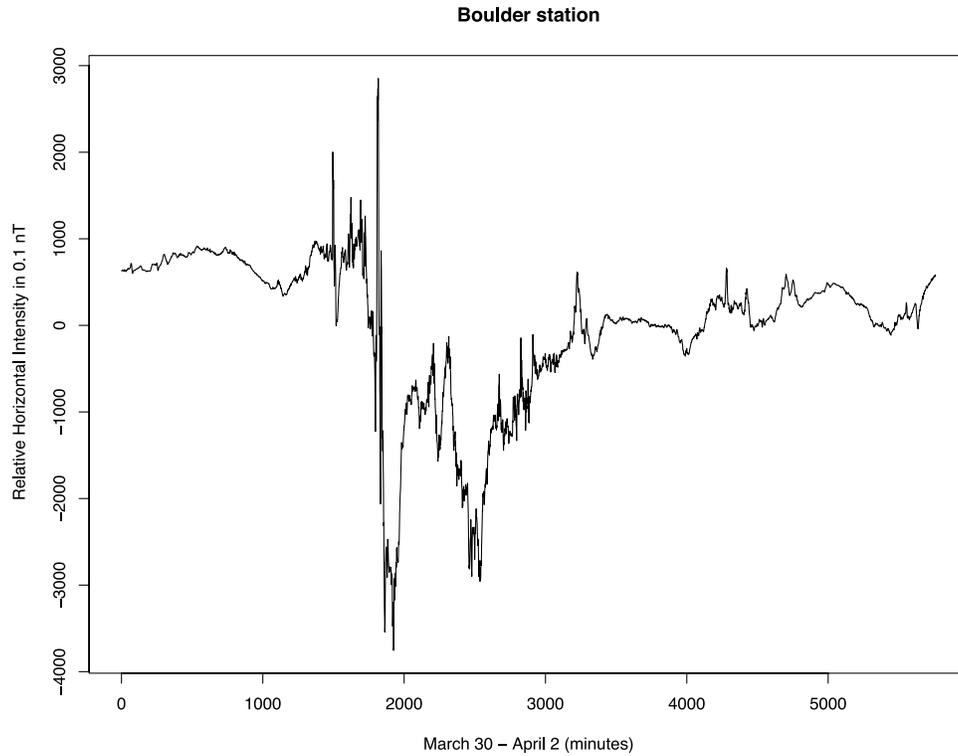


Figure 1. Relative Horizontal intensity in 0.1 nT, Boulder station 30 March to 2 April 2001.

distribution which is close to Gaussian or a distribution with heavy probability tails. We explain the concept of heavy tails by referring to Figure 2. Figure 2a shows a simulated realization of 2048 independent standard normal random variables. As is well known, practically all observations are within three standard deviations of the mean and no unusually large or small observations are seen. By contrast,

in Figure 2b, there are a number of observations which are much larger or much smaller than most observations. This means that there is a large probability that an observation is far away from the mean, or, equivalently, that the density function extends farther away from the mean than a normal density; that is, it has heavier tails. Random variables like those in Figure 2b are therefore called heavy tailed. They

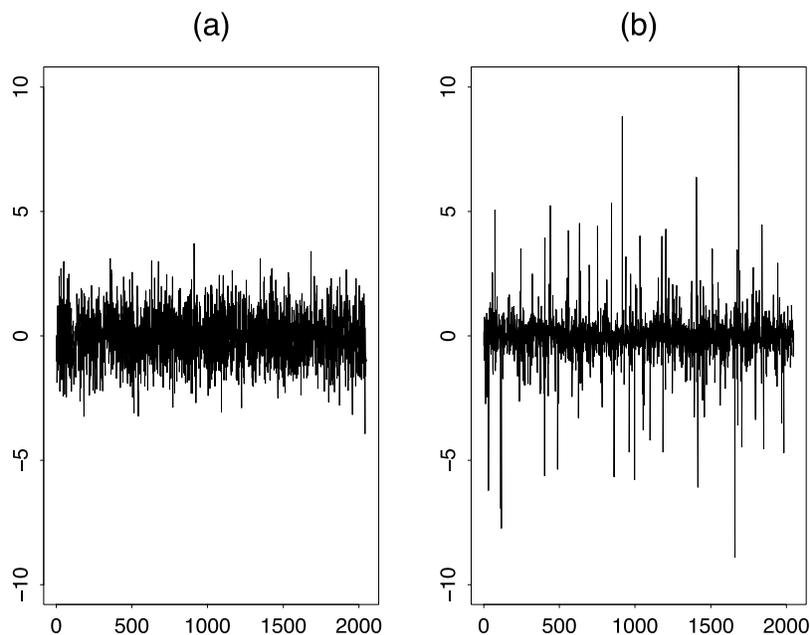


Figure 2. (a) Normal (Gaussian) random variables. (b) Heavy-tailed random variables with tail index $\alpha = 1.85$.

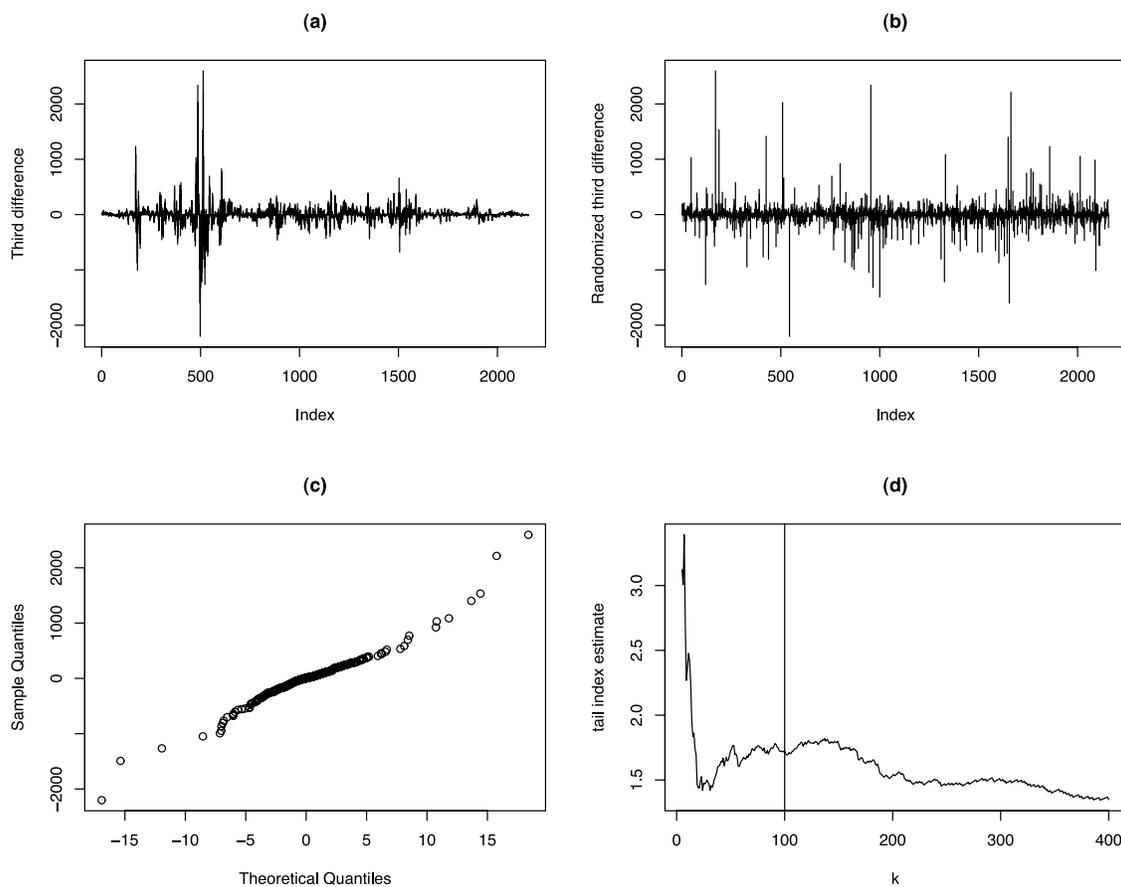


Figure 3. (a) Third difference of the horizontal intensity, Boulder, 30 March to 2 April 2001. (b) Randomized observations from panel Figure 3a. (c) Plot of observations in Figure 3b against the quantiles of Student's t_2 distribution. (d) Hill plot of observations in Figure 3b.

are characterized by the tail index α whose precise definition is given in section 3. Statistical methodology and theory for observations which have heavy tails is often very different than for observations which are approximately normal. It is of particular importance to know if the tail index α is greater or smaller than 2. If $\alpha < 2$, most statistics which use sample standard deviation cannot be used because the sample standard deviation does not converge to the population standard deviation. This issue is taken up in some detail in section 4, but it is important to keep in mind that the differences in statistical methodology go much further; we refer to *Adler et al.* [1998] for a more extensive background.

[4] To establish a connection between magnetograms and heavy-tailed random variables, we refer to Figure 3. Figure 3a shows third differences of a central portion of the magnetogram displayed in Figure 1. Differencing is a standard technique in time series analysis [see, e.g., *Brockwell and Davis*, 2002, pp. 29–35]. If the observations are denoted by X_t , $t = 0, \dots, N - 1$, the series of (first) differences is defined as $\nabla X_t = X_t - X_{t-1}$, $t = 1, \dots, N - 1$. Second differences are differences of the first differences, etc. Differencing typically eliminates a meandering behavior of a time series. For example, the series in Figure 3a has approximately constant mean zero which makes it more regular than the original data displayed in Figure 1. As will

be explained in section 2, wavelet coefficients are in a sense equivalent to local differencing at various scales. However, to focus attention, we consider in this introduction the traditional simple differences, choosing the third difference arbitrarily to illustrate the point (differences of orders one to four have similar properties). The variability of observations in Figure 3a evolves with time reflecting the evolution of the storm. By contrast, the variability of the simulated observations in Figure 2b does not appear to change with time. This is because these observations are independent, unlike the third differences of the magnetogram. To be able to compare these two set of observations, we applied a random permutation to the observations in Figure 3a and displayed the result in Figure 3b. A visual inspection of Figures 2b and 3b reveals a striking similarity and points out that the distribution of the differences (and wavelet coefficients) of magnetograms might be heavy tailed. Most methods for estimating the tail index work best for independent observations, and that is why we always apply a random permutation before studying the tail behavior. Intuitively, randomizing the observations destroys their time dependence and allows isolation of the feature of interest: the distribution of the extremal observations.

[5] Figures 3c and 3d illustrate two methods of estimating the tail index α : QQ and Hill plots. These and other methods are discussed in detail in section 4. If a plot with

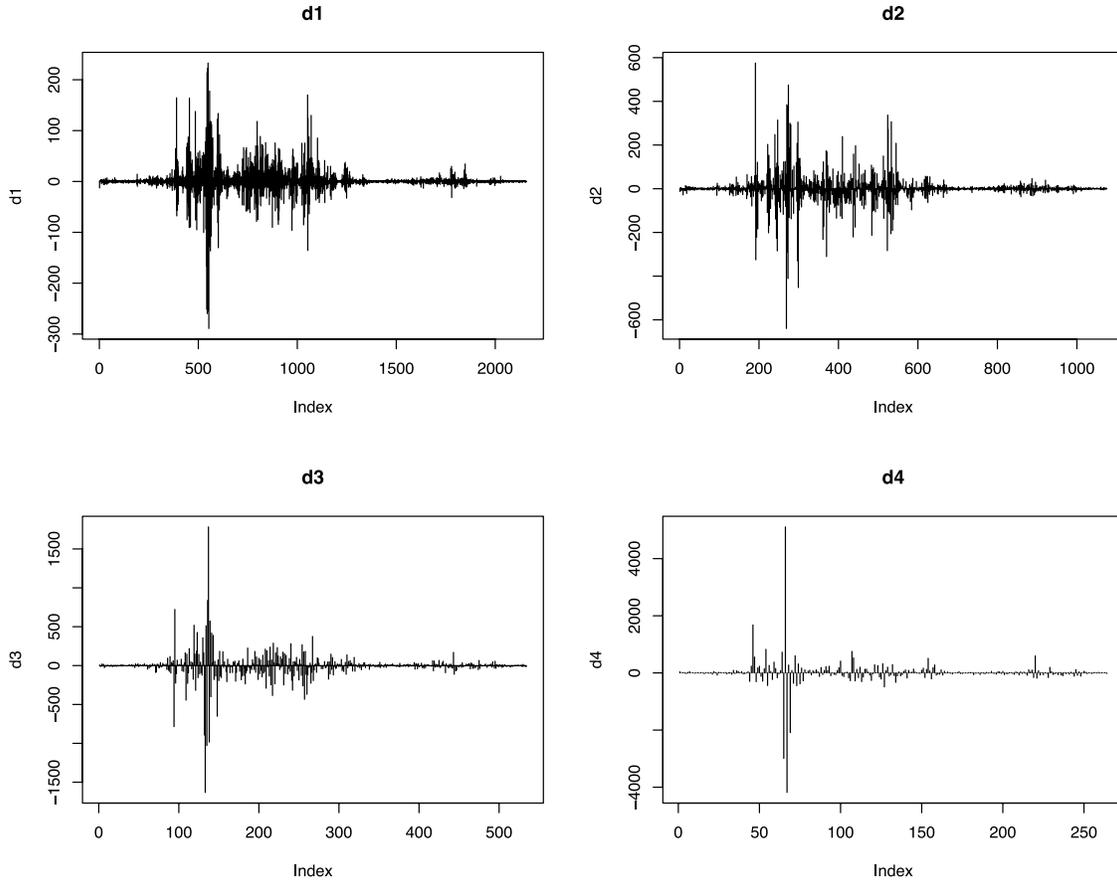


Figure 4. Plots of the DWT coefficients d_j , $j = 1, \dots, 4$, Boulder station, 30 March to 2 April 2001.

respect to a reference distribution follows approximately a straight line, we conclude that the reference distribution is a reasonable approximation to the distribution of the observations. In Figure 3c, the reference distribution (Student's t distribution with two degrees of freedom) has tail index $\alpha = 2$. We thus see that the tail index of the third differences of the magnetogram is approximately equal to 2. A more detailed inspection of Figure 3c, to be presented in section 4, shows that this index is in fact slightly smaller than 2. Figure 3d shows the so-called Hill plot which will also be discussed in detail in section 4. The vertical line at $k = 100$ shows the value of k which we can use to find a tail index estimate on the y axis, which is seen to be about 1.75.

[6] The main goal of this paper is to establish that the wavelet coefficients of magnetogram records are heavy tailed and to estimate their tail index. We also present preliminary evidence showing that the tail behavior is different during magnetically quiet and storm periods.

[7] The wavelet coefficients themselves are defined in section 2. Heavy-tailed distributions are introduced in some detail in section 3. In section 4, we describe statistical techniques used to estimate the tail index and calibrate them to suit our needs. Section 5 contains a detailed analysis of the tail behavior of the wavelet coefficients of the horizontal (H) component of the magnetic field during the storm event shown in Figure 1, while section 6 summarizes the main outcomes of our analysis, provides some extensions, and discusses possible further research. In the following, when

referring to random variables, we use the common abbreviation i.i.d. for “independent identically distributed.”

2. Discrete Wavelet Transform

[8] Suppose X_0, X_1, \dots, X_{N-1} is an observed time series. In our context, X_t is to be thought of as the disturbance of the H component t min after the start of the record under study. In order to define the discrete wavelet transform (DWT), it is convenient to represent the observations as a column vector $\mathbf{X} = [X_0, X_1, \dots, X_{N-1}]^T$. We must also assume that $N = m2^J$, for some integers m and J . The DWT is then a column vector \mathbf{W} of length N (the same as the number of observations), which can be represented as

$$\mathbf{W} = [d_1^T, d_2^T, \dots, d_J^T, s_J^T]^T, \quad (1)$$

where

$$d_j = [d_{j,0}, d_{j,1}, \dots, d_{j,N_j-1}]^T, \quad N_j = 2^{-j}N, \quad j = 1, 2, \dots, J \quad (2)$$

$$s_J = [s_{J,0}, s_{J,1}, \dots, s_{J,N_J-1}]^T, \quad N_J = 2^{-J}N. \quad (3)$$

The wavelet coefficient $d_{j,k}$ reflects the oscillations of the observations around time $2^j k$ and at scale 2^j . The coefficient $s_{J,k}$ is roughly the average of 2^J observations around time

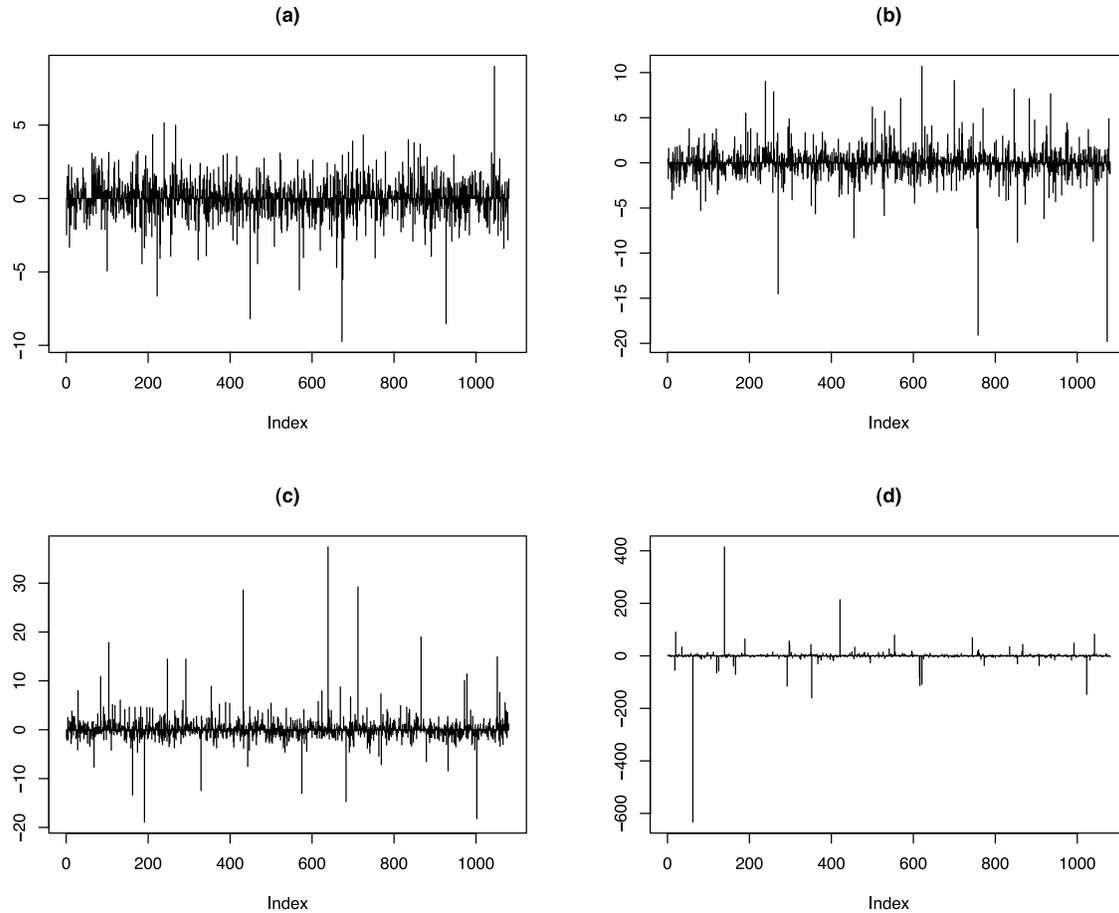


Figure 5. Realizations of sequences of independent identically distributed random variables with tail index (a) $\alpha = 3.5$, (b) $\alpha = 2.5$, (c) $\alpha = 1.5$, and (d) $\alpha = 1$.

$2^j k$. These vague statements can be made precise, we refer to chapter 4 of *Percival and Walden* [2000]. The vector \mathbf{W} is obtained from the vector \mathbf{X} by means of multiplication by an orthonormal matrix (which is implemented as a fast pyramid algorithm). Thus \mathbf{W} can be viewed as \mathbf{X} rotated in the N dimensional space. The objective of the transformation is to see the behavior of \mathbf{X} at different scales.

[9] It is illuminating to compare the DWT to the continuous wavelet transform (CWT), which has been used in recent geophysical research [*Boberg et al.*, 2002]. Denoting by $\psi(t)$ a mother wavelet [see e.g., *Chan*, 1995, chapter 2; *Percival and Walden*, 2000, chapter 1], the value of the CWT at time a and at scale c is defined as

$$D(c, a) = \int_0^N X_t c^{-1/2} \psi((t-a)/c) dt. \quad (4)$$

The values $D(c, a)^2$ are typically plotted on the (a, c) -plane using a color scheme and indicate the energy distribution across time and scale. Note that the above definition assumes that X_t is available in continuous time, so in practice an approximation to the integral must be used. The pyramid algorithm computes the DWT exactly, but it gives only an approximation to the CWT. One can show that in an approximate sense, which can be made precise [see *Percival and Walden*, 2000, chapter 11], $d_{j,k}$ is equal to $D(2^j, 2^j k)$.

The DWT is thus approximately equal to the CWT at dyadic scales and at discrete times, with the number of coefficients $N_j = 2^{-j} N$ decreasing with increasing scale. The DWT is a nonredundant representation of a signal measured in discrete time and, since it is a result of an orthonormal transformation, it exactly preserves the energy of the signal in the sense that

$$\sum_{t=0}^{N-1} X_t^2 = \sum_{j=1}^J \sum_{k=0}^{N_j-1} d_{j,k}^2 + \sum_{k=0}^{N_{J-1}-1} s_{J,k}^2. \quad (5)$$

Moreover the DWT is more amenable to quantitative statistical analysis and possess desirable statistical properties like the approximate decorrelation property. These issues are too extensive to be discussed here; we refer to the monograph of *Percival and Walden* [2000] and to *Veitch and Abry* [1999], *Percival et al.* [2000], *Whitcher et al.* [2002], *Abry et al.* [2002], *Craigmile et al.* [2005], and *Jach and Kokoszka* [2005a] (also A. Jach and P. Kokoszka, Wavelet domain test for long-range dependence in the presence of a trend, submitted to *Statistics*, 2005), to name just a few recent contributions.

[10] In this paper the DWT using the least asymmetric LA(8) filter, also referred to as symmlet [see *Daubechies*, 1992], and $J = 4$ was applied to the magnetogram records.

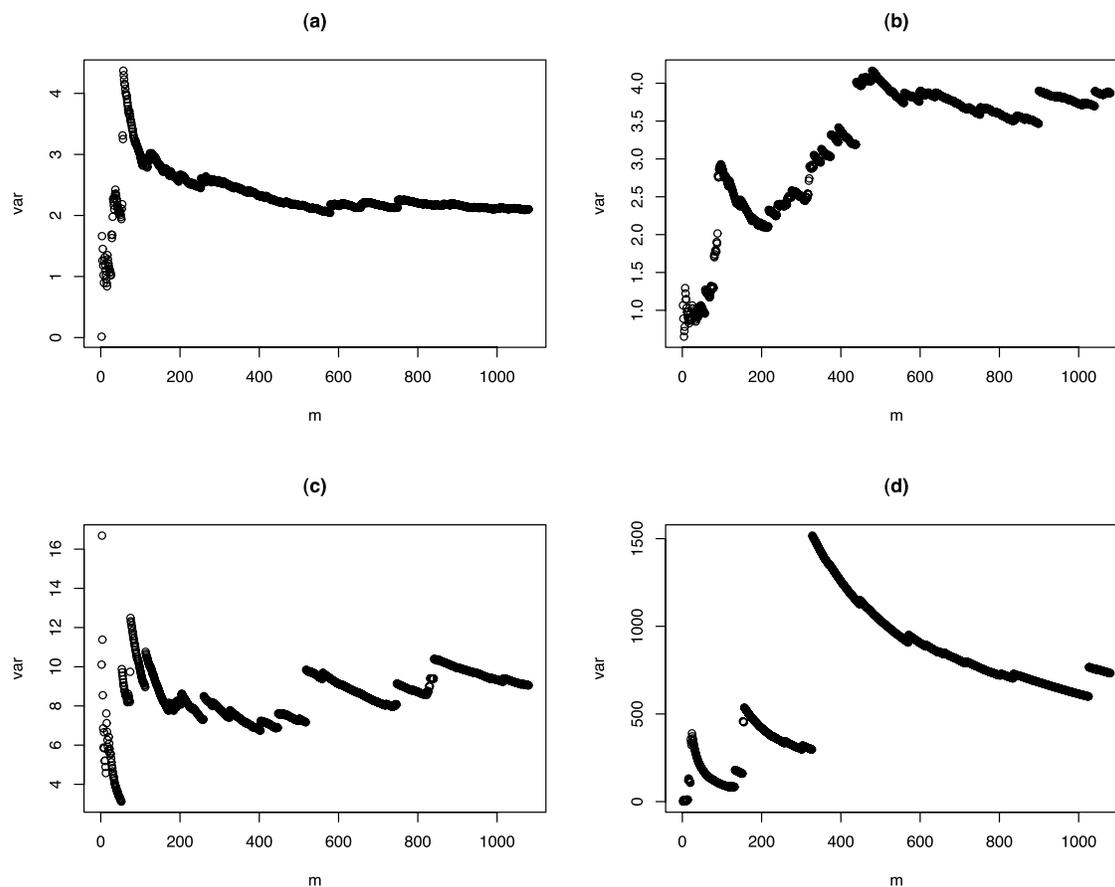


Figure 6. Variance plots of t distributed random variables with (a) $\nu = 3.5$ and (b) $\nu = 2.5$ degrees of freedom, and stable random variables with (c) $\alpha = 1.5$ and (d) $\alpha = 1$.

The parameter L in the definition of the LA(L) filter corresponds to twice the number of local differences. Thus an application of an LA(8) DWT is in a sense equivalent to locally differencing the time series 4 times at the dyadic scales. To make these statements precise, extensive notation and theory is needed which is well beyond the scope of this paper; an interested reader is referred to chapter 4 of *Percival and Walden* [2000]. The choice of a filter is equivalent to a choice of a mother wavelet. As described in section 11.6 of *Percival and Walden* [2000] the main advantage of the symmlets is that the resulting wavelet coefficients do not “drift” relative to the original signal. The choice of the width of the filter is a major issue, since short filters may introduce some undesirable artifacts on one hand. On the other hand, filters with large L can lead to the increase in computational burden, more coefficients influenced by boundary coefficients, etc. Combining all the requirements one usually ends up with LA(8).

[11] We focused on the coefficients $d_{j,k}$ because the coefficients $s_{j,k}$ reflect the smooth component of the magnetograms and are not centered at zero; the $s_{j,k}$ are not used in wavelet-based statistical inferential procedures. Figure 4 presents the coefficients of the DWT of the magnetometer record from Boulder station during 30 March to 2 April 2001; the record itself is shown in Figure 1.

[12] Since the DWT uses the circular filtering that treats the time series as a periodic sequence, the boundary

coefficients should be treated with caution. Using Table 137a of *Percival and Walden* [2000], we removed the boundary coefficients to eliminate their effect on further analysis.

3. Heavy-Tailed Distributions

[13] A random variable X is said to have a heavy-tailed distribution with tail index α if

$$P(|X| > x) = L(x)x^{-\alpha} \quad x > 0, \quad (6)$$

where $\alpha > 0$ and $L(x)$ is a slowly varying function defined by the condition

$$\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1. \quad (7)$$

From the practical point of view, it is convenient to think that $L(x)$ is a function which has a finite positive limit as $x \rightarrow \infty$.

[14] The probability tails of a heavy-tailed random variable thus approach zero much slower than in the case of the normal distribution for which $P(|X| > x) \sim C_1 x^{-1/2} \exp\{-C_2 x^2\}$, for some positive constants C_1 and C_2 . Unlike the normal distribution, condition (6) implies that there is a

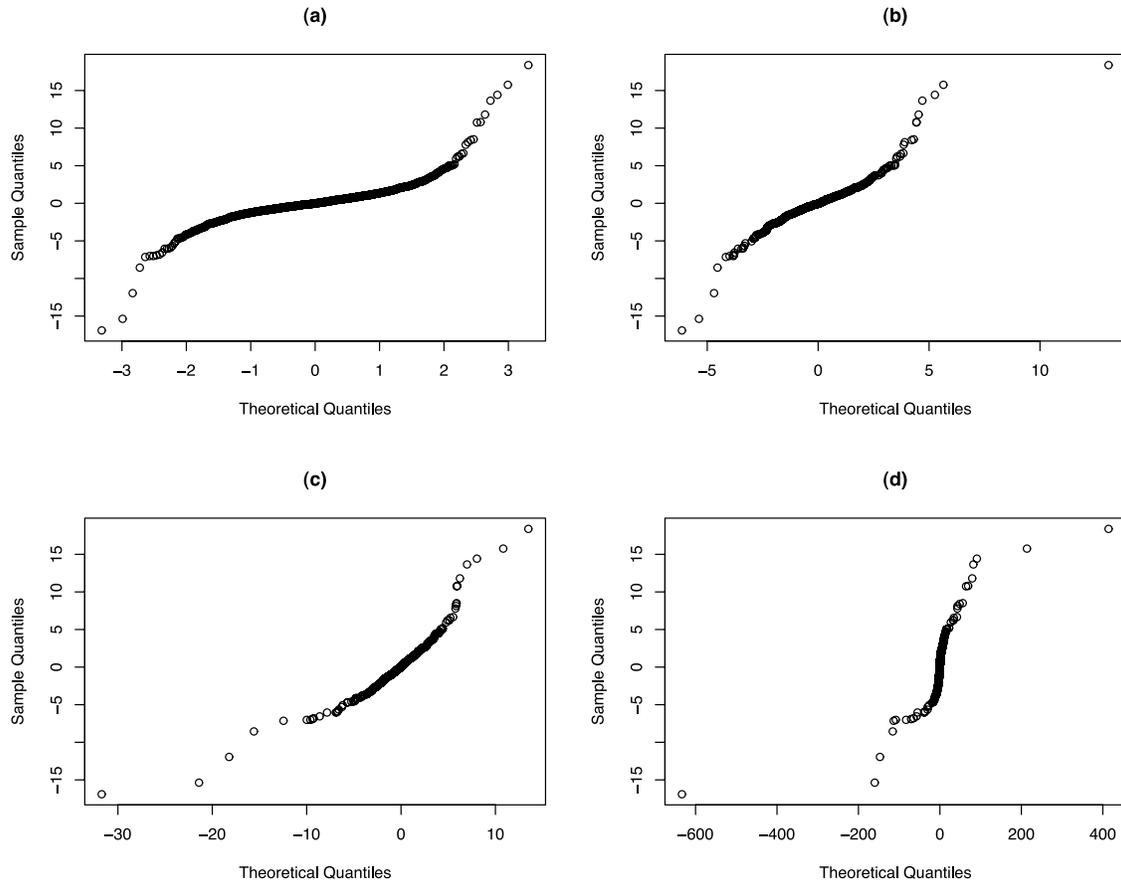


Figure 7. Plots of $s_{\alpha S}$ with $\alpha = 2$ against (a) normal, (b) t_4 , (c) t_2 , and (d) $s_{\alpha S}$, $\alpha = 1$.

large probability that a heavy-tailed random variable takes a value far from the center of the distribution.

[15] As examples of a heavy-tailed distributions we consider the symmetric α -stable ($s_{\alpha S}$) distribution and the t distribution. We use both of them to calibrate the statistical methodology discussed in section 4.

[16] The $s_{\alpha S}$ distribution is defined in chapter 1 of *Samorodnitsky and Taqqu* [1994]. Here we note merely that the tail index α of this distribution satisfies $0 < \alpha < 2$. Consequently, the $s_{\alpha S}$ distribution has infinite variance. Stable distributions are of fundamental importance in probability theory and arise as limits of partial sums in a similar manner as the Gaussian distribution does in the Central Limit Theorem [see *Gnedenko and Kolmogorov*, 1954].

[17] To model heavy-tailed distributions with finite variance, we used t distributed random variables. A random variable T has Student's t distribution with ν degrees of freedom, if its density is

$$f_T(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} (\pi\nu)^{-1/2} (1+t^2/\nu)^{-\frac{\nu+1}{2}} \quad -\infty < t < \infty. \quad (8)$$

The t distribution has finite variance, $\text{Var}(T_\nu) = \frac{\nu}{\nu-2}$, if $\nu > 2$. A simple calculation shows that

$$P(T > x) = \int_x^\infty f_T(t) dt \sim Cx^{-\nu}, \quad (9)$$

so the tail index α for t distribution is equal to the number of the degrees of freedom ν . As ν increases, the density $f_T(t)$ approaches the normal density, but these densities are considered close only for $\nu > 30$.

4. Methods for Detecting and Estimating Heavy Tails

[18] In this section we introduce some techniques used to detect a heavy tailed distribution and to estimate the tail index. First, two visual analysis tools are discussed: the converging variance test used to identify infinite variance distributions and plots used to verify if the data come from a heavy-tailed distribution and to estimate the range of the tail index. Finally, the Hill estimator and its modifications are presented.

[19] In order to describe these techniques, we focus on four series of random variables with known tail indexes α (Figure 5). The first two are t distributed random variables with finite variance and $\alpha = 3.5$ and $\alpha = 2.5$. The other two follow an infinite-variance stable distribution with tail index $\alpha = 1.5$ and $\alpha = 1$.

4.1. Converging Variance Test

[20] The converging variance test is an informal visual data analysis technique used to determine whether the data have infinite variance. The main idea of the test consists of plotting the sample variance σ_m^2 of the first m observations as a function of m . If the data come from the distribution

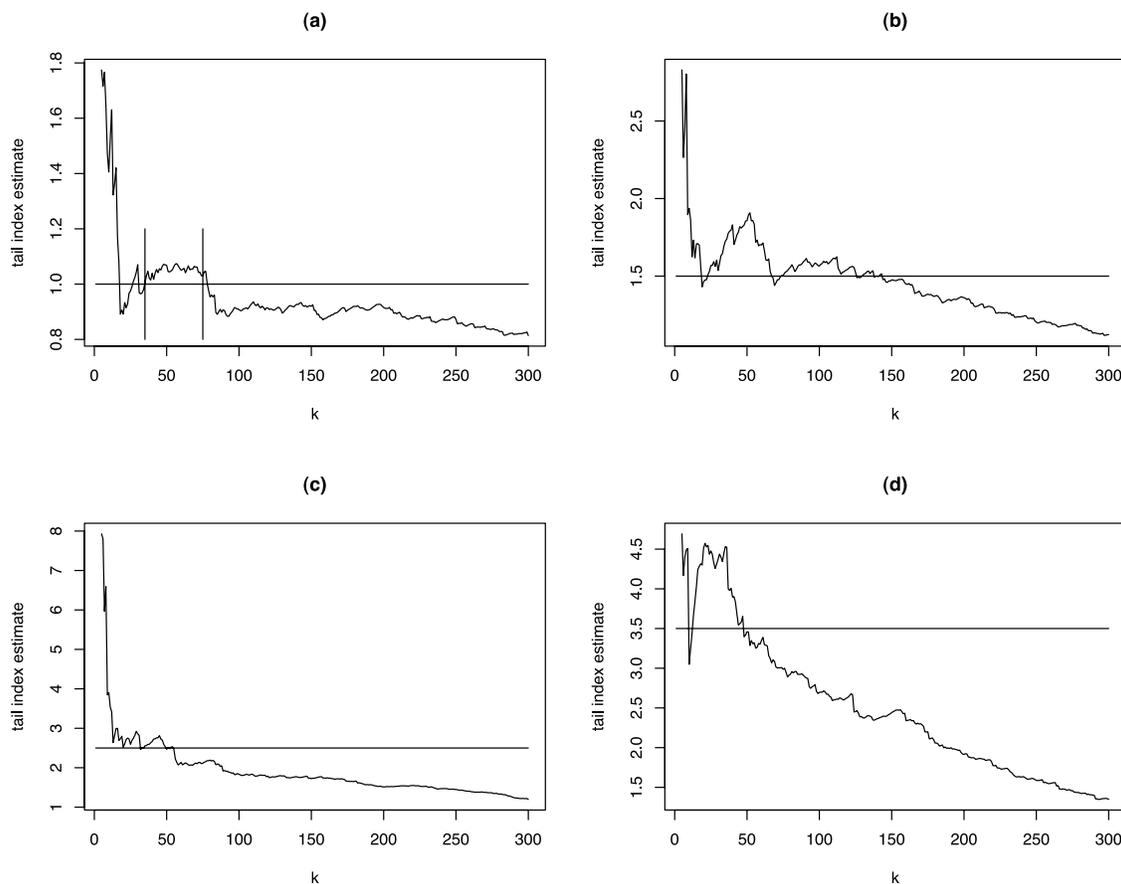


Figure 8. Hill plots for stable random variables with (a) $\alpha = 1$ and (b) $\alpha = 1.5$, and t distributed random variables with (c) $\nu = 2.5$ and (d) $\nu = 3.5$ degrees of freedom.

with infinite variance, σ_m^2 will show large jumps. Otherwise, it will converge to a finite value. Despite the fact that this test was originally designed for i.i.d. random variables, it is also applicable to dependent data, as long as the order of the observations is first randomized [Adler et al., 1998, p. 137].

[21] Variance plots work well for identifying extremely heavy-tailed series. For example, the variance plot of $s \propto s$ random variables with $\alpha = 1$ shows the biggest jumps which indicate the infinite variance of the distribution of the data, Figure 6d. Similarly, for $\alpha = 1.5$, Figure 6c, the variance plot suggests the infinite-variance distribution. When the tail index is around 2 the jumps on the variance plots are not large, but still visible for both infinite and finite variance (see Figure 6b). Finally, when the tail index α exceeds 3, the variance plots clearly converge to a finite value, correctly indicating the finite-variance distribution, see Figure 6a. Thus the converging variance test is useful if the true tail index is not too close to 2: Large jumps may occur owing to chance even if the data have finite variance. Similarly, partial variances may appear to converge to a finite value even when the data have infinite variance. Therefore the converging variance test should be used only as an exploratory tool, or to validate other more quantitative methods.

[22] In section 5.1, the converging variance test is used to find out if the distribution of the DWT coefficients of the

magnetometer records have infinite variance. In the following section, another type of visual analysis is presented.

4.2. Plots

[23] The plots are used to solve a wide variety of problems concerned with exploring the distribution of the data. In this paper, they are used to verify whether the coefficients of the DWT of the magnetometer records follow a distribution with heavy tails. The plot is the graph of the points

$$\left(X_{k,n}, F^{-1}\left(\frac{n-k+1}{n+1}\right) \right), \quad k = 1, \dots, n, \quad (10)$$

where $X_{k,n}$ is the k -th largest observation (order statistic) in a sample of size n , and F is a continuous distribution function (cdf).

[24] By Glivenko-Canteli theorem [Billingsley, 1995, theorem 20.6], which says that the empirical cdf converges

Table 1. Number of the Upper Order Statistics k Used to Compute the Hill Estimator $\hat{\alpha}_{k,n}$ for Selected Series Lengths

Series Length	Number of Upper Order Statistics k
2160	100
1080	125
540	50
270	40

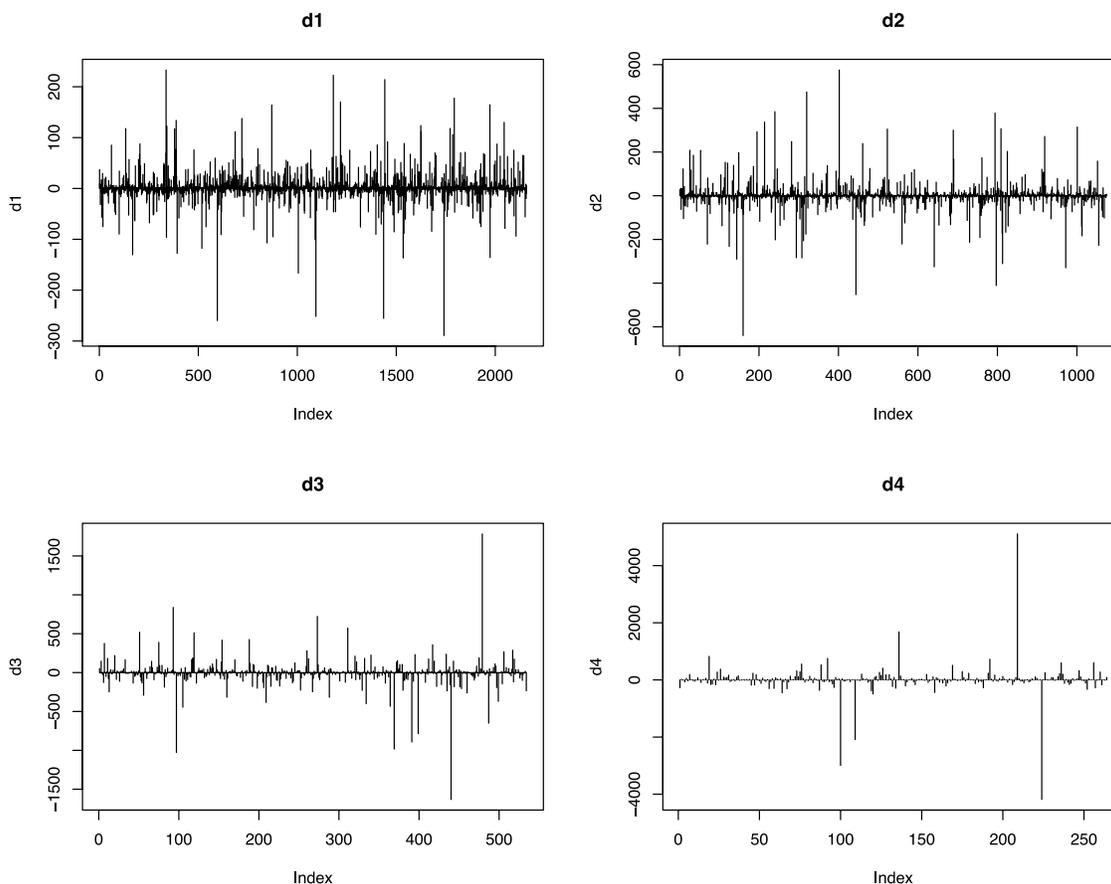


Figure 9. Plots of the randomized DWT coefficients $d_j, j = 1, \dots, 4$.

almost surely to the theoretical one, the plots should be roughly linear if X_i have cdf F . This property is illustrated in Figure 7c, which shows the plot obtained for 1080 simulated t_2 random variables with $\alpha = 2$ and using as F the correct t cdf.

[25] Applying the plot analysis to the data, one compares the plot to a straight line. If the plot curves down on the left side and up on the right, the distribution of the data has a tail that is heavier than the one of the reference distribution F (Figures 7a and 7b). If the plot curves up on the left and down on the right, the tail index of the distribution of the data is larger than the tail index of the reference distribution F ; see Figure 7d. This method applied to the wavelet coefficients (section 5.2) gives results consistent with a more quantitative technique introduced in the next section.

4.3. Hill Estimator

[26] In this section the Hill estimator of the tail index α is introduced and its practical application is discussed. The Hill estimator of α is based on the k upper order statistics $X_{k,n} \leq \dots \leq X_{1,n}$, i.e., k largest observations, and takes the form

$$\hat{\alpha}_{k,n} = \left(\frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1}. \quad (11)$$

In our application, the largest k observations are always positive because the wavelet coefficients have mean zero

and k is very small relative to the total number of coefficients.

[27] The Hill estimator is very natural since different asymptotically equivalent versions of $\hat{\alpha}_{k,n}$ can be derived by various methods, and it has the same form as the maximum likelihood estimator; see chapter 6.4.2. of *Embrechts et al.* [1997]. Moreover, for the Hill estimator, standard statistical properties hold, such as weak consistency [Mason, 1982], strong consistency [Deheuvels et al., 1988] and the asymptotic normality, using which the asymptotic confidence intervals can be constructed [Embrechts et al., 1997, theorem 6.4.6].

[28] To compute the estimate $\hat{\alpha}_{k,n}$, one has to choose an appropriate value of k , which is the main focus of this section. In the asymptotic theory, it is assumed that $k = k(n)$ is a function of the sample size n . The mathematical conditions for k state merely that

$$k(n) \rightarrow \infty, \quad \frac{n}{k(n)} \rightarrow \infty; \quad (12)$$

that is, a sufficiently large number of upper order statistics should be used; however, this number should be asymptotically negligible relative to the sample size.

[29] If conditions (6) and (12) are satisfied, then

$$\sqrt{k}(\hat{\alpha}_{k,n} - \alpha) \xrightarrow{d} \mathcal{N}(0, \alpha^2). \quad (13)$$

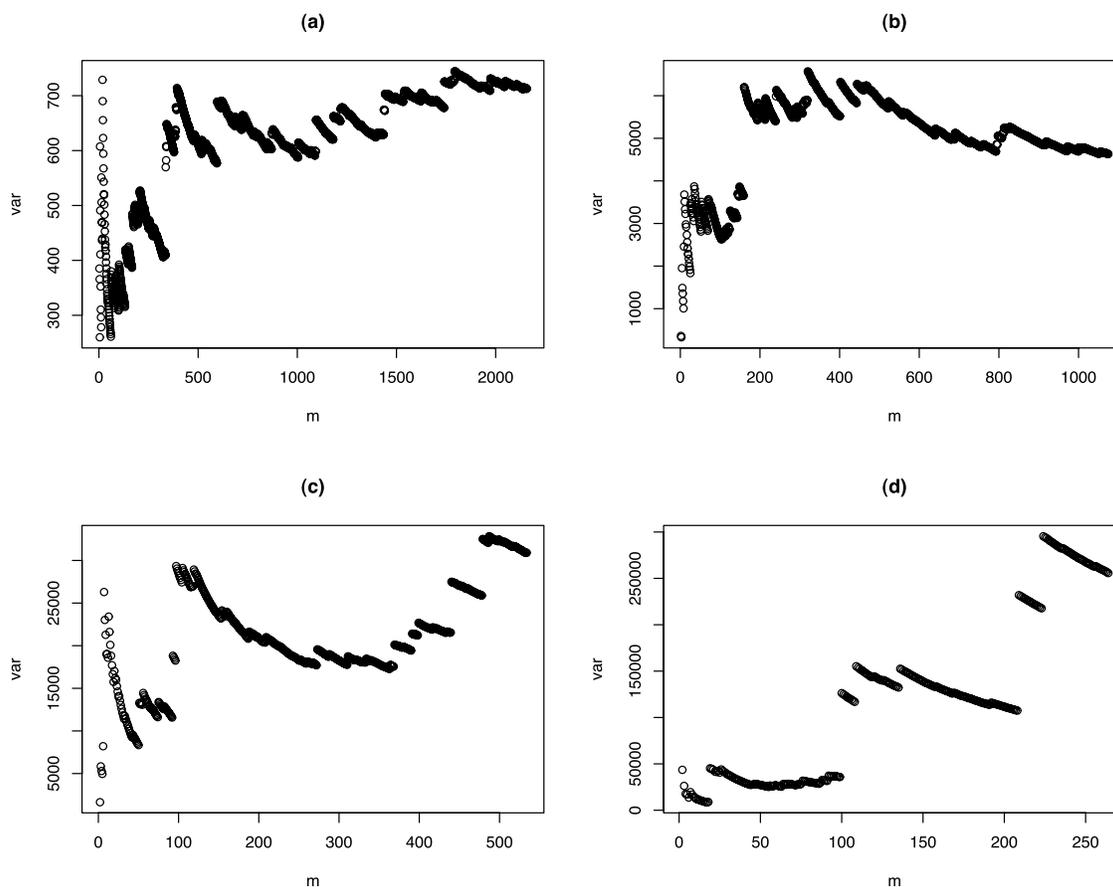


Figure 10. Variance plots of the randomized DWT coefficients (a) d_1 , (b) d_2 , (c) d_3 , and (d) d_4 .

Using this property, one can construct a confidence interval. For example, a 95% (asymptotic) confidence interval for the tail index α is

$$\left(\hat{\alpha}_{k,n} - 1.96 \frac{\hat{\alpha}_{k,n}}{\sqrt{k}}, \quad \hat{\alpha}_{k,n} + 1.96 \frac{\hat{\alpha}_{k,n}}{\sqrt{k}} \right). \quad (14)$$

[30] The Hill estimator is difficult to use in practice owing to its sensitivity to the choice of the number of upper order statistics k . Unfortunately, there is no algorithm for selecting an appropriate k . Therefore, before estimating the tail index α of the wavelet coefficients of the magnetograms, we had to find an appropriate value of k for a given series length n . The optimal k may depend not only on n , but also on the unknown tail index α . We found values of k which give acceptable results for a range of α of interest as follows: First, the Hill plots, which provide the tail index estimate for different k values, were used to find an approximate k value that would give a reasonable tail index estimate for any heavy tailed distribution with $\alpha \in (0, 4]$. “Good” Hill plots tend to have a noticeable horizontal stretch where they are roughly constant over some values of k . An example of such a stretch is marked with two vertical lines in Figure 8a. It is preferable that k is chosen from such a region since it usually gives an accurate tail index estimate. We followed this recommendation and selected the “optimal” k by visually examining a large number of Hill plots for several known values of α (indicated by horizontal lines in Figure 8) and the values of n relevant to the analysis of

the magnetograms. We then validated our choices by performing a simulation study which produced biases and standard errors based on 1000 simulated values of $\hat{\alpha}_{k,n}$. This simulation study, presented in detail by Maslova [2005], allowed us to draw conclusions about the behavior of the Hill estimators. Before discussing these conclusions, we present in Table 1 the values of k we settled on as a function of the sample size n . (The value of n is not to be confused with the number of 1-min observations N ; in section 5, n corresponds to the number of the wavelet coefficients at a given scale.)

[31] The estimates $\hat{\alpha}_{k,n}$ are close to the true value when α is smaller than 1.5. When the true tail index α increases, $\hat{\alpha}_{k,n}$ overestimates it: If α is between 1.5 and 2, $\hat{\alpha}$ will likely be above 2. Thus, for this range of α , the Hill estimator might be misleading, since it suggests the finite variance distribution, when in fact the true distribution has infinite variance. By contrast, $\hat{\alpha}_{k,n}$ underestimates α when the data come from a heavy-tailed distribution with $\alpha > 2$.

[32] A possible solution is transforming the data so that they would have heavier tails, for which $\hat{\alpha}_{k,n}$ gives a more accurate estimate of the tail index α . If X has tail index α , then (6) holds, and so

$$P(X^2 > x) = P(|X| > x^{\frac{1}{2}}) = 2L(x^{\frac{1}{2}})x^{-\frac{\alpha}{2}}. \quad (15)$$

It is easy to see from (7) that $L(x^{1/2})$ is also slowly varying. Hence, if the distribution of X has tail index α , then the tail

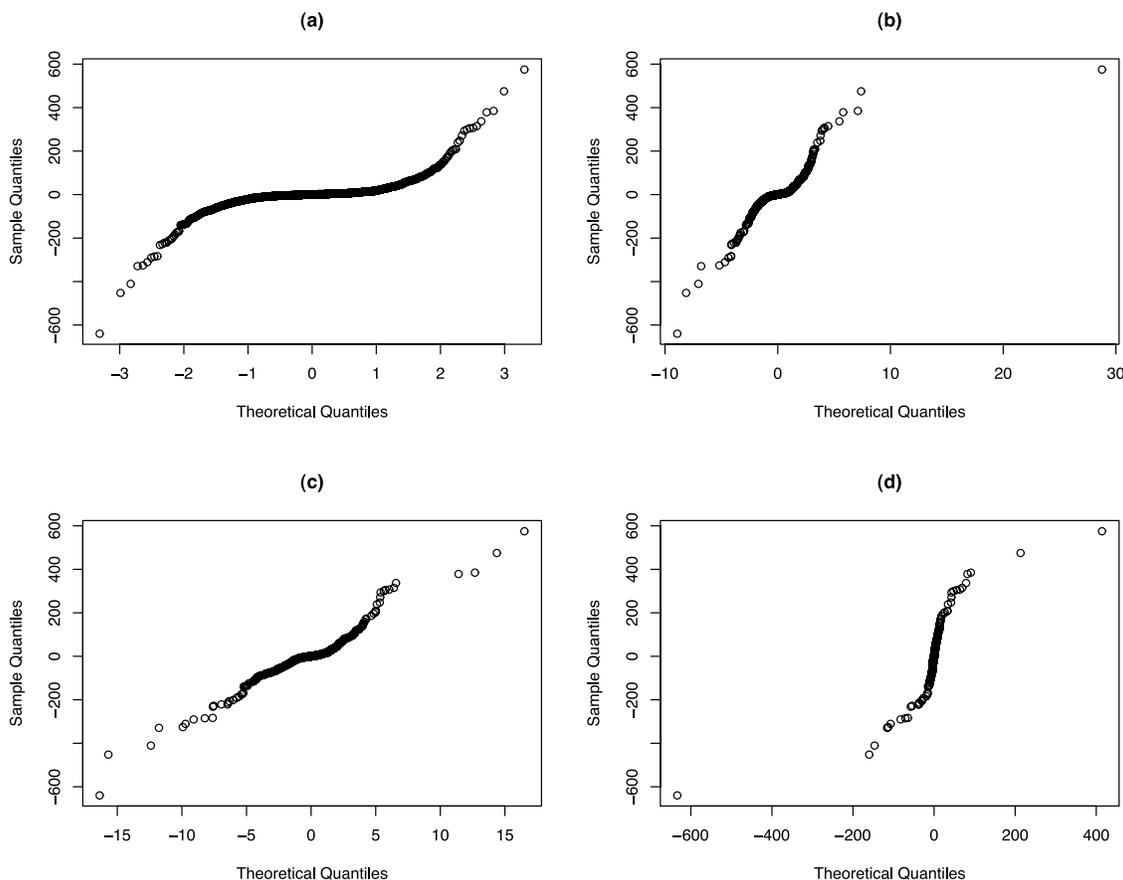


Figure 11. Plots of the randomized DWT coefficients d_2 against (a) normal, (b) t_4 , (c) t_2 , and (d) soS , $\alpha = 1$.

index of X^2 is $\alpha/2$. Applying the transformation $Y = X^2$, the tail index of a new series Y will thus not exceed 2, provided that the original data have the tail index less than 4, as will be seen to be the case in our applications. In this case ($0 < \alpha < 2$) the Hill estimator only slightly overestimates the tail index.

[33] Another way to improve the Hill estimator is to use the smoothing method proposed by *Resnick and Stărică* [1997] which is a simple averaging technique designed to help to minimize the sensitivity of the estimator to the number of the upper order statistics and reduce the high variability of the Hill plot. It consists of averaging the Hill estimator over several values of k , i.e., computing

$$\bar{\hat{\alpha}}_{k,n} = \frac{1}{(u-1)k} \sum_{p=k+1}^{uk} \hat{\alpha}_{p,n}, \tag{16}$$

where $u > 1$. *Resnick and Stărică* [1997] recommend to use $u \in (n^{0.1}, n^{0.2})$, where n is the sample size. We used $u = 4$ for series of length $n = 2160$, $n = 1080$, and $u = 3$ for $n = 540$, $n = 270$.

[34] Owing to the fact that the range of the smoothed plot is reduced in comparison to the classical Hill plot, the method is less sensitive to the choice of k . However, the simulation results of *Maslova* [2005] show that the average Hill estimator severely underestimates the tail index of finite variance distributions. One should keep this in mind when applying this method to the real data.

[35] In conclusion, we can state that the Hill estimator and its two versions discussed above must be used with care and in conjunction with simulations and graphical methods like converging variance test and plots. Unlike the graphical methods, the Hill estimator however offers numerical values of estimates together with a measure of uncertainty (confidence interval).

5. Probability Tails of the Wavelet Coefficients of the H Component

[36] In section 4 we described several methods of detecting and estimating heavy tails based on a sample of i.i.d random variables. Comparison of Figures 4 and 5 shows that the DWT coefficients of the magnetograms are not i.i.d. Therefore we always apply a random permutation to the DWT coefficients of magnetograms. Randomized DWT coefficients are shown in Figure 9 and are seen to resemble the patterns in Figure 5.

[37] The detailed analysis in this section refers to the dynamic range of the H component measured at the Boulder station during the stormy period of 31 March to 2 April 2001. Generalizations to other periods are discussed in section 6.

5.1. Converging Variance Test

[38] Recall that the converging variance test is used to determine whether wavelet coefficients have finite or infinite variance. We compare the variance plots of the wavelet

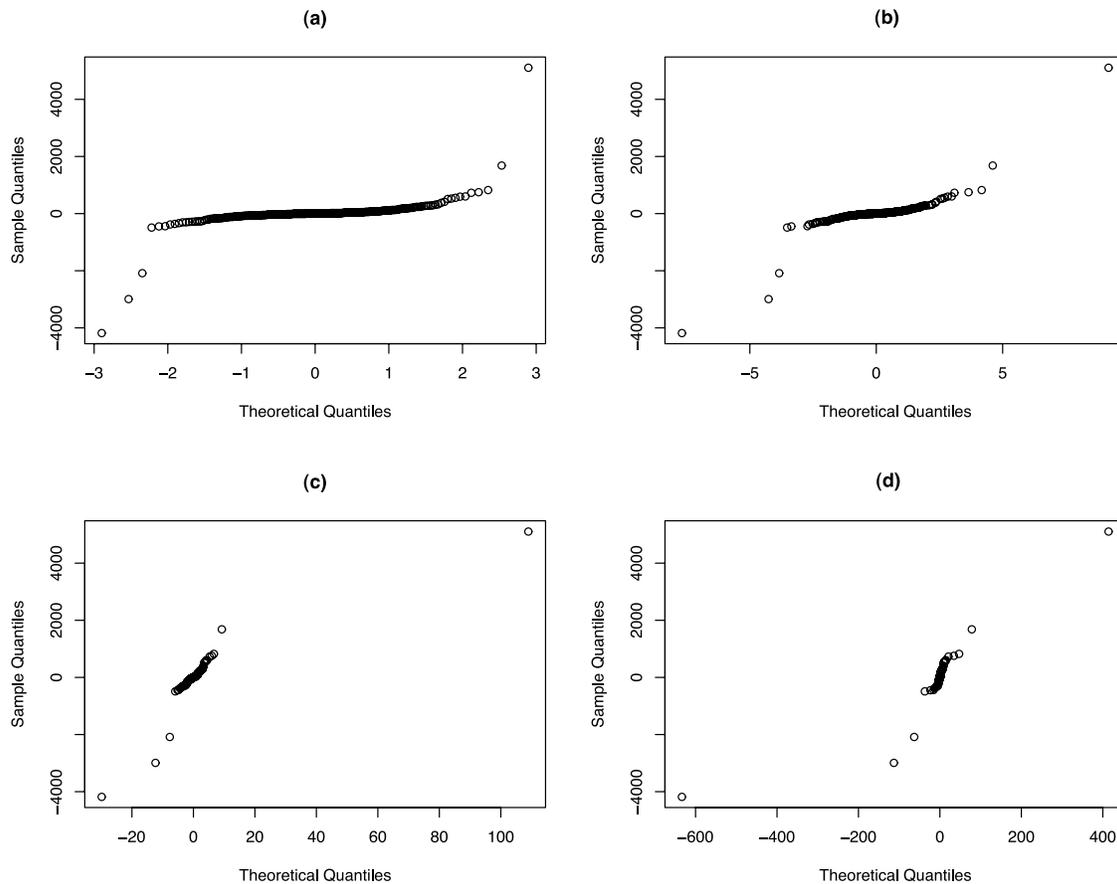


Figure 12. Plots of the randomized DWT coefficients d_4 against (a) normal, (b) t_4 , (c) t_2 , and (d) $s\alpha s$, $\alpha = 1$.

coefficients (Figure 10) to the variance plots of the known distributions introduced in section 3 (Figure 6).

[39] The wavelet coefficients d_1 seem to belong to the finite variance distribution with the tail index $\alpha \approx 2.5$ (compare Figure 6b). As discussed in section 4.1, when the variance plot shows slight jumps like the variance plot of d_2 and d_3 , Figure 10, the data might come either from the infinite variance distribution with the tail index $\alpha \in (1.5, 2)$ or from the finite variance distribution with $\alpha \in (2, 2.5)$. Finally, one can clearly see that the wavelet coefficients d_4 come from an infinite variance distribution, and the tail index α appears not to exceed 1.5.

[40] The wavelet coefficients clearly belong to the class of heavy-tailed distributions either with infinite variance, with $\alpha \in (1.5, 2)$, or with finite variance and the tail index $\alpha \in (2, 2.5)$. The tail index appears to get smaller with increasing level j . It is fairly safe to conclude that the converging variance test indicates that the index α for the coefficients d_j is between 1.5 and 2.5 for $j = 1, 2, 3$ and between 1 and 2 for $j = 4$. For more details refer to section 4.3 of Maslova [2005].

5.2. QQ Plot Analysis

[41] In section 4.2, we discussed how the QQ plots can be used to detect a heavy-tailed distribution and estimate an approximate range of the tail index. In this section the QQ plots are used to verify that the distributions of the wavelet coefficients are heavy tailed.

[42] The plots of the wavelet coefficients d_2 and d_4 are presented, respectively in Figures 11 and 12. Corresponding plots for the coefficients d_1 and d_3 are similar and are not shown to conserve space. The empirical distributions of the wavelet coefficients are compared to the standard normal distribution, t_4 and t_2 distributions, and $s\alpha s$ distribution with $\alpha = 1$.

[43] As noted in section 4.2, if the points in the plot follow a line, one can infer the reference distribution F . If the plot first curves down and then up, then the distribution has a tail heavier than F . If it first curves up, then down, the tail is lighter.

[44] The DWT coefficients at all levels $j = 1, 2, 3, 4$ are seen to have tails much heavier than normal, the tail index is, in fact, smaller than 4, compare panels a and b in Figures 11 and 12. On the other hand, the tail index is always greater than 1, see panel d in Figures 11 and 12. It is thus very safe to conclude that $1 < \alpha < 4$. The analysis of panel c in Figures 11 and 12 indicates that α is perhaps slightly greater than 2. These conclusions accord with those reached in section 5.1 except in case of the coefficients d_4 for which the converging variance test indicated infinite variance, i.e., $\alpha < 2$.

[45] We note that the main part of the plot in Figure 12d appears small because of the presence of two relatively large observations which must be accommodated by the plot. This underscores the difficulty of using plots for very heavy tailed distributions. We also note that the number of

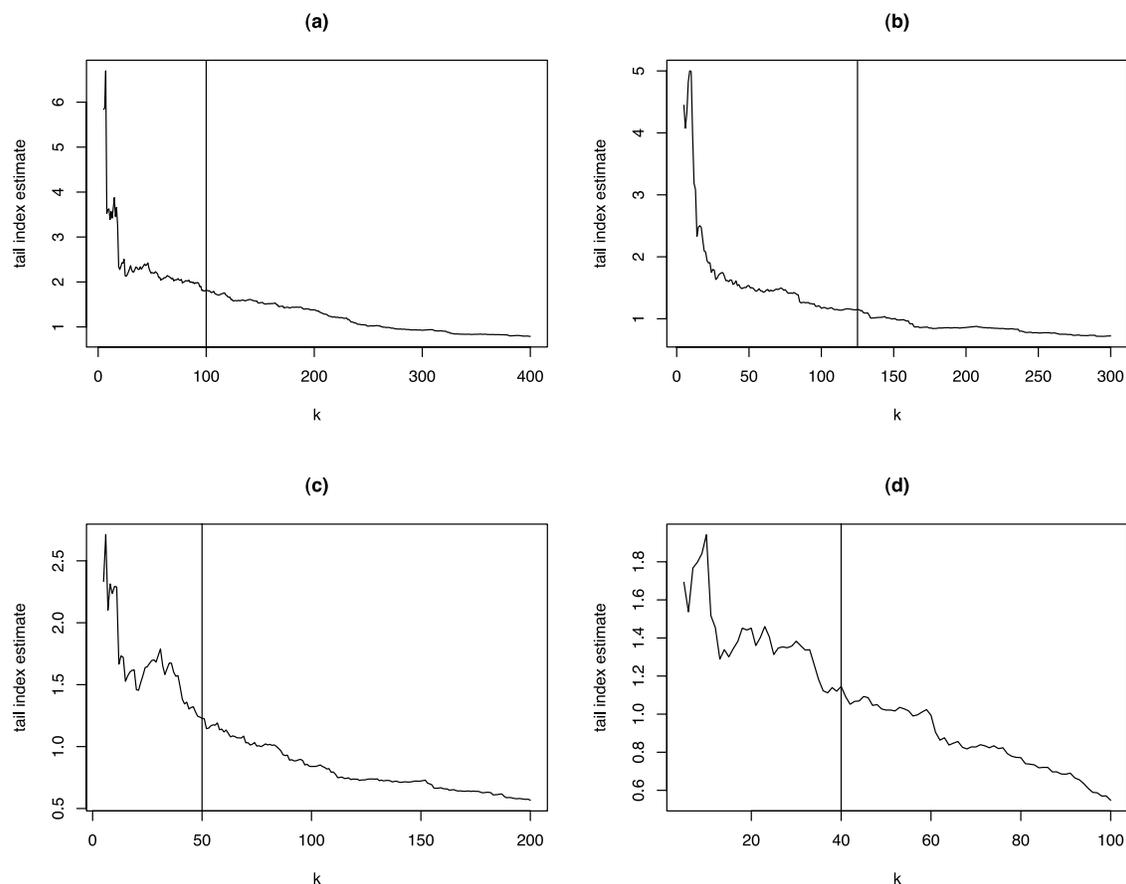


Figure 13. Hill plots of the DWT coefficients (a) d_1 , (b) d_2 , (c) d_3 , and (d) d_4 .

observations on which these plots are based decreases by two from $n = 2160$ for $j = 1$ to $n = 270$ for $j = 4$. This also partially explains the different appearance of the coefficients in Figure 9: The coefficients d_4 may look much more heavy tailed than d_1 because there are 8 times fewer points, so only a few large observations which dominate the plot are possible. We therefore apply in the next section the Hill estimator which has been calibrated in section 4.3 to the specific values of n used in our analysis.

5.3. Hill Estimates of the Tail Index

[46] As discussed in section 4, the classical Hill estimator is sensitive to the choice of the number of upper order statistics k , but we found values of k which give reasonable tail index estimates for sample sizes n considered in this study. These k values are now used to estimate the tail index of the distributions of the wavelet coefficients and are indicated by vertical lines in Figure 13 (and displayed in Table 1). The squaring method introduced in section 4 only slightly overestimates the tail index and is thus considered to be most accurate. The smoothing method reduces the variability of the Hill plot; however, it severely underestimates α .

[47] In Table 2, the classical, average and “squared” Hill estimates of the tail indexes of the wavelet coefficients are given. These estimates are consistent with the results obtained from the converging variance test, but indicate tails heavier than those inferred from the plots.

[48] The values of the estimates in Table 2 accord with the properties of the three versions of the Hill estimator. To better illustrate this point refer to Figure 14, where tail index estimates found using three methods and their asymptotic 95% confidence intervals are shown. The numbers on the x -axis correspond to the wavelet coefficient levels $j = 1, \dots, 4$. Notice that the average Hill estimator provides the smallest estimates in all cases. The “squared” Hill estimator gives the highest tail index estimates.

[49] Figure 14 shows that it is safe to conclude that the DWT coefficients at all four levels have tail indexes smaller than 2.5. Since the “squared” Hill estimator has a tendency to overestimate α , it is likely that all tail indexes are smaller than 2, in which case the DWT coefficients should be modeled using an infinite variance distribution.

6. Conclusions, Extensions, and Further Research

[50] The objective of this paper is to show that the distribution of the wavelet coefficients of the H component

Table 2. Hill Estimates of the Tail Index $\hat{\alpha}$ of the Wavelet Coefficients d_j , $j = 1, \dots, 4$

Tail Index Estimates $\hat{\alpha}$	Wavelet Coefficients, d_j			
	d_1	d_2	d_3	d_4
Hill estimates	1.8195	1.146	1.229	1.1453
Average Hill estimates	1.535	0.884	0.9937	0.8991
“Squared” Hill estimates	1.9697	1.495	1.544	1.4094

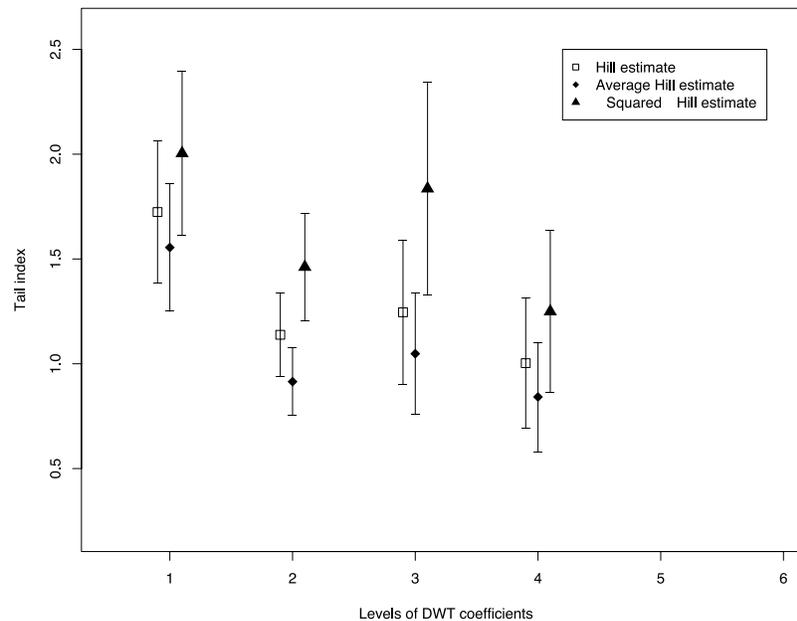


Figure 14. Hill estimates and their confidence intervals using different methods: (1) Hill estimates of d_1 , (2) d_2 , (3) d_3 , and (4) d_4 .

of magnetograms is heavy tailed and to obtain a quantitative information on the probability tails (estimate the tail index). This has been accomplished by applying two visual and three inferential statistical tools which have been calibrated for our specific application in section 4. A detailed analysis of one data set was presented in section 5. The same overall conclusion holds for other periods, other stations and other wavelet filters: The wavelet coefficients have heavy tails with the tail index in the range from 1 to 3.

[51] An example of a more extensive analysis is presented in Table 3 and Figure 15 which show “squared” Hill estimates of α for five stormy and five quiet periods in March and April 2001. We focused on the “squared” version of the Hill estimator because it has been found to be most accurate.

[52] Some interesting conjectures can be gleaned from Table 3 and Figure 15. For every DWT level $j = 1, 2, 3, 4$, the average tail index of the five stormy periods is smaller than the corresponding average of the five quiet periods. As seen in Figure 15, the range of the tail indexes for the quiet periods is somewhat shifted to the up relative to the range for stormy periods, especially for levels $j = 2$ and $j = 3$, which correspond to physical scales of 4–8 min. In fact, an application of a two-sample t test [see, e.g., Hayter, 2002, section 9.3] shows that the averages for levels $j = 2$ and $j = 3$ are statistically significantly different with P-values equal, respectively, to 0.63% and 4.08%. This means that the probability of extreme events in the wavelet domain is higher during stormy periods than during quiet periods. While such a finding is not surprising, our analysis offers a quantitative way of approaching this difference in behavior.

[53] The conjecture that the tail index decreases during stormy periods must be confirmed, or refuted, by a more extensive and systematic analysis similar to that presented by Wanliss [2005], who used a large collection of active and

quiet period in the epoch 1981–2002 to show that the self-similarity exponent of the SYM-H index increases during active periods. Such an analysis is however beyond the scope of the present paper. The main lesson of this paper is that when applying statistical techniques to the wavelet coefficients of the magnetograms one must make sure that these techniques are robust to large departures from Gaussianity and take into account the presence of heavy probability tails.

[54] In our future work, by using a much larger database, we will be able to find out whether the systematic decrease of the tail index value from quiet to storm periods can be statistically confirmed and will further explore the behavior of the tail index for wavelet coefficients at octaves $j = 2, 3$. We will try to understand why the most pronounced difference in the tail index between storm and calm periods occurs at these scales. This future research may reveal the

Table 3. Hill Estimates of the Tail Index, $\hat{\alpha}$ of the Wavelet Coefficients d_j , $j = 1, \dots, 4$

	Wavelet Coefficients, d_j			
	d_1	d_2	d_3	d_4
<i>Stormy Periods</i>				
March 2001	2.9998	2.3308	2.5429	1.7422
19–21 March 2001	2.0210	1.9221	2.6507	2.0690
30 March to 2 April 2001	1.9697	1.4950	1.5440	1.4094
5–7 April 2001	2.7775	2.3183	2.2374	2.7570
11–13 April 2001	2.2930	1.7145	1.6952	1.5947
Average	2.4122	1.9561	2.1341	1.9145
<i>Quiet Periods</i>				
8–10 March 2001	3.0642	2.9300	3.1270	3.1469
14–16 March 2001	2.0557	2.7402	2.6747	1.5555
25–27 March 2001	2.4027	2.2886	2.3275	1.9983
15–17 April 2001	3.4009	2.7007	2.5397	2.2251
23–25 April 2001	3.0805	2.5401	2.8320	2.0916
Average	2.8008	2.6399	2.7002	2.2035

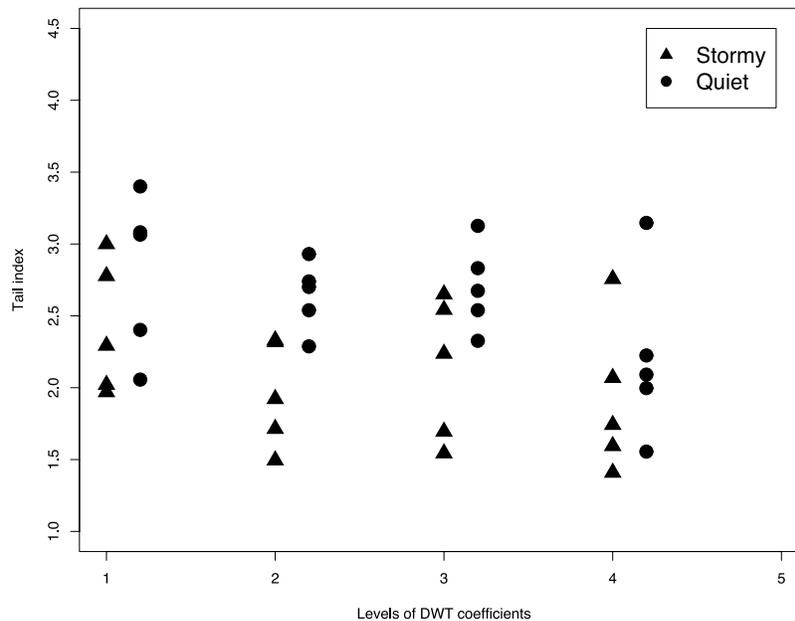


Figure 15. “Squared” Hill estimates of the wavelet coefficients during the stormy and calm periods.

underlying physics of tail index for the wavelet coefficients representing different timescales and connect them to the dynamic variations of various individual current components in the M-I system.

[55] **Acknowledgments.** Research was supported by NSF grant DMS-0413653 and NASA grant NNG05GJ48G. Data were provided by USGS via the global network of observatories INTERMAGNET.

[56] Arthur Richmond thanks Hualiang Wei and another reviewer for their assistance in evaluating the manuscript.

References

- Abry, P., P. Flandrin, M. S. Taqqu, and D. Veitch (2002), Self-similarity and long-range dependence through the wavelet lens, in *Theory and Applications of Long Range Dependence*, edited by P. Doukhan, G. Oppenheim, and M. S. Taqqu, pp. 527–556, Birkhäuser Boston, Cambridge, Mass.
- Adler, R., R. Feldman, and M. S. Taqqu (1998), *A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions*, Birkhäuser Boston, Cambridge, Mass.
- Billingsley, P. (1995), *Probability and Measure*, 3rd ed., John Wiley, Hoboken, N. J.
- Boberg, F., H. Lundstedt, J. T. Hoeksema, P. H. Scherrer, and W. Liu (2002), Solar mean magnetic field variability: A wavelet approach to Wilcox Solar Observatory and SOHO/Michelson Doppler Imager observations, *J. Geophys. Res.*, *107*(A10), 1318, doi:10.1029/2001JA009195.
- Brockwell, P. J., and R. A. Davis (2002), *Introduction to Time Series and Forecasting*, second ed., Springer, New York.
- Chan, Y. T. (1995), *Wavelet Basics*, Springer, New York.
- Craigmile, P. F., P. Guttorp, and D. B. Percival (2005), Wavelet-based parameter estimation for polynomial contaminated fractionally differenced process, *IEEE Trans. Signal Proc.*, *53*, 3151–3161.
- Daubechies, I. (1992), *Ten Lectures on Wavelets*, CBMS-NSF Ser., vol. 61, Soc. of Ind. and Appl. Math., Philadelphia, Pa.
- Deheuvels, P., E. Hausler, and D. M. Mason (1988), Almost sure convergence of the Hill estimator, *Math. Proc. Cambridge Soc.*, *104*, 371–381.
- Embrechts, P., C. Klüppelberg, and T. Mikosch (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, New York.
- Friedrich, E., G. Rostoker, M. G. Connors, and R. L. McPherron (1999), Influence of the substorm current wedge on the *Dst* index, *J. Geophys. Res.*, *104*, 4567–4576.
- Gnedenko, B. V., and A. N. Kolmogorov (1954), *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Boston, Mass.
- Haldoupis, C., D. Pancheva, and N. J. Mitchell (2004), A study of tidal and planetary wave periodicities present in midlatitude sporadic E layers, *J. Geophys. Res.*, *109*, A02302, doi:10.1029/2003JA010253.

- Hayter, A. J. (2002), *Probability and Statistics*, Duxbury, Boston, Mass.
- Jach, A., and P. Kokoszka (2005), Wavelet based confidence intervals for the self-similarity parameter, *J. Stat. Comput. Simul.*, in press.
- Krankowski, A., W. Kosek, and L. W. Baran (2005), Wavelet analysis and forecasting of VTEC obtained by GPS observation over European latitudes, *J. Atmos. Sol. Terr. Phys.*, *67*, 1147.
- Lui, A. T. Y., and A.-H. Najmi (1997), Time-frequency decomposition of signals in a current disruption event, *Geophys. Res. Lett.*, *24*, 3157–3160.
- Maslova, I. (2005), Wavelet analysis of magnetometer data, Master’s thesis, Utah State Univ., Logan.
- Mason, D. M. (1982), Laws of large numbers for sums of extreme values, *Ann. Probab.*, *10*, 756–764.
- Percival, D. B., and A. T. Walden (2000), *Wavelet Methods for Time Series Analysis*, Cambridge Univ. Press, New York.
- Percival, D. B., S. Sardy, and A. C. Davison (2000), Wavestrapping time series: adaptive wavelet based bootstrapping, in *Nonlinear and Nonstationary Signal Processing*, edited by W. J. Fitzgerald, R. L. Smith, A. T. Walden, and P. C. Young, pp. 442–471, Cambridge Univ. Press, New York.
- Resnick, S. I., and C. Stărică (1997), Smoothing the hill estimator, *Adv. Appl. Probab.*, *29*, 271–293.
- Samorodnitsky, G., and M. S. Taqqu (1994), *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York.
- Tsyganenko, N. A. (2000), Modeling the inner magnetosphere: The asymmetric ring current and region 2 Birkland current revisited, *J. Geophys. Res.*, *105*, 27,739–27,754.
- Veitch, D., and P. Abry (1999), A wavelet based joint estimator of the parameters of long-range dependence, *IEEE Trans. Inf. Theory*, *45*(3), 878–897.
- Wanliss, J. (2005), Fractal properties of SYM-H during quiet and active times, *J. Geophys. Res.*, *110*, A03202, doi:10.1029/2004JA010544.
- Wei, H. L., S. A. Billings, and M. Balikhin (2004), Prediction of the *dst* index using multiresolution wavelet models, *J. Geophys. Res.*, *109*, A07212, doi:10.1029/2003JA010332.
- Whitcher, B., S. Byers, P. Guttorp, and D. Percival (2002), Testing for homogeneity of variance in time series: long memory, wavelets and the Nile River, *Water Resour. Res.*, *38*(5), 1054, doi:10.1029/2001WR000509.

P. Kokoszka and I. Maslova, Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322-3900, USA. (piotr.kokoszka@usu.edu; inga@cc.usu.edu)
 J. Sojka and L. Zhu, Center for Atmospheric and Space Science, Utah State University, 4405 Old Main Hill, Logan, UT 84322-4405, USA. (fasojka@sojka.cass.usu.edu; zhu@cc.usu.edu)