Some Comments on the Analysis of "Big" Scientific Time Series

Article *in* Proceedings of the IEEE · September 2016

DOI: 10.1109/JPROC.2016.2598218

2 authors:

David James Thomson
Queen's University
168 PUBLICATIONS **5,923** CITATIONS
SEE PROFILE

Frank L. Vernon
University of California, San Diego
256 PUBLICATIONS **3,038** CITATIONS
SEE PROFILE

SOME COMMENTS ON THE ANALYSIS OF "BIG" SCIENTIFIC TIME SERIES

David J. Thomson[1] *and Frank L. Vernon, III*[2]

1 Dept. of Mathematics and Statistics,
Queen's University, Kingston Ontario K7L 3N6
djt@mast.queensu.ca (613) 533-2426
2 IGPP, Scripps Institution of Oceanography,
Univ. of California at San Diego, La Jolla, CA 92093
flvernon@ucsd.edu (858) 534-5537

ABSTRACT

Experience with long time series from space, climate, seismology, and engineering has demonstrated the need for even longer data series with better precision, timing, and larger instrument arrays. We find that almost all the data we have examined, including atmospheric, seismic data, and dropped calls in cellular phone networks contains evidence for solar mode oscillations that couple into Earth systems through magnetic fields, and that these are often the strongest signals present. Long series of daily climate data show some of the identical modes that have been observed in space. We show two examples suggesting that robustness has been overused and that many of the extremes in geomagnetic and space physics data may be the result of a superposition of numerous modes. We also present initial evidence that the evolution of turbulence in interplanetary space may be controlled by modes.

Returning to the theme of "big data", our experience has been that theoretical predictions that spectra would be asymptotically unbiased have turned out to be largely irrelevant with very long time series primarily showing that we simply didn't understand the problems. Data that was considered to have excessively variable spectra appear to evolve into processes with dense sets of modes. In short data blocks these modes are not resolved and as the relative phase of the modes within the estimator varies, so does the apparent power. Ideas that data series become uncorrelated at modest distances in either time or space do not seem to be true with the long duration continuous time series data we have examined.

1. INTRODUCTION

This paper is concerned with changes in the analysis of *scientific* time series data implicit in "big data" with examples from seismology, climate, space physics, and engineering.

Our experience is that, in these fields, the requirement for precision instruments has, if anything, increased. In addition to the number of samples that can be acquired in a scalar time series, the precision of the data and the dimensionality have increased as well. In the early 1970s, a continuous series of a million samples was rare, but are now routine and getting on the small side.

All of these influence how one analyzes data. In our experience, much of the analysis occurs in the frequency domain where increases in precision and sample size put *increasing* demands on spectral analysis and related methods. The hope that large enough samples would cause relatively simple estimates of spectral densities to converge to something useful, *i.e.*, [2, 3, 4] has been overwhelmed by the increasing complexities revealed by the improved and larger data. The late John W. Tukey once remarked (approximately) "Don't spend too much time on asymptotics because the only reason that anyone will go to the trouble and expense of collecting large amounts of data is because they are going to ask harder questions."

The largest single change that we have observed with "big data" is in our approach to *science*. Traditionally, one often had a specific hypothesis or question in mind, *e.g.*, "can we detect a particular mode in this data?", followed by "what is its frequency and Q?" An early foray into "big data" time series [5, 6], was given by the development of the WT4 millimeter waveguide system [7] at Bell Labs. The dominant loss mechanism in this waveguide was from mode conversion, with the loss proportional to the power spectrum of the distortion. There were many propagating modes, and the various spectra were estimated from *mechanical* measurements [8] of the geometry of individual sections of waveguide. A combination of manufacturing, shipping, and field considerations limited the length of individual waveguides to about 9 meters so a dominant concern at the time was how to obtain accurate estimates of spectra from such short samples. There were between 1 and 30 series of mechanical measurements on individual waveguide,

each between 900 and 5000 samples, and more than 1500 different waveguides. The goal was initially simple: "does this piece of waveguide meet our quality specifications?" with the major specifications bounded by limits on the spectrum of geometric distrortions. Attempts to discover *why* individual sections failed quickly converted this simple goal into more complex estimates. One notes, incidentally, that the first goal involved sampling distributions and these were built into the specifications but, once these were violated and one entered the "detective" phase, exact statistical distributions were almost irrelevant. Finally, at the field–trial stage [9], the individual sections were welded together to give a 14 km test run. This was measured mechanically at a 1 cm rate to produce series of horizontal and vertical curvature and diameter, each approximately 1.4 million samples, "big" by the standards of the 1970s.

Currently, the question is more "what exotic features are suggested by this data and do we have analysis tools to describe them?" One should remember that modern spectrum estimates, as described in most texts, are conceptually identical to what was suggested by Sir G.G Stokes in 1878, [10]. Although numerous extensions to the basic theory have been proposed, such as: coherence [11]; bispectra [12] and the families of "higher–order" spectra [13], and improvements in estimators, non–stationary series are still more art than science. With such data, more apropos estimates are often obtained by iterating the basic spectrum estimation procedure. That is, one computes a series of spectrum estimates, extracts some parameters of interest from each, then treats these as a new time series. This can usually only be done in big data situations. Some of these new tricks are described in sections 10.2 and 10.3.

This paper provides a brief review of the theoretical background for multitaper spectral analysis, coherence, and similar functions followed by examples of involving large amounts of data. "Large" can be used here to describe two distinct phenomena: first, the obvious one where there are so many observations that they become computationally or graphically cumbersome; second, the more interesting cases where the number of observations is modest, but have a relatively long duration and are often the longest such data set known. Examples include the Burgundian Pinot Noir grape harvest dates [14], §8.2, and the Uppsala daily temperature record, §10.5.

2. WHAT IS "NOISE"?

A major goal of both acquiring and analyzing data is to improve the "signal–to–noise" ratio. If you are fortunate enough to have a description of the noise in your data that starts with kT, where k is Boltzmann's constant and T is the temperature in Kelvins and analysis shows residuals within a dB or two of this level, you can skip the rest of this section. Most of us, however, are not so fortunate and characterizing "noise" is a major task. We give two specific examples:
Example one, in seismic data, "noise" on the two horizontal components is often 10 dB (or more) higher than it is on the vertical [15] although all three components may be derived from identical sensors. A possible explanation was given in [16].
Example two, in magnetotellurics and the related problem of estimating induced voltages and currents on power lines and pipelines from changes in the geomagnetic field, *improper* complex–valued data effectively generates a "noise" field with a similar magnitude as the proper data, [17].

Based on these and similar observations, a major change in our outlook is that much of what used to be considered "noise" in many different kinds of data appears to originate in normal modes of the Sun. Solar modes are systematic oscillations of the Sun's interior excited by turbulence in the convection zone and which resonate at frequencies from ~ 0 μHz to $\gtrsim 5000$ μHz. These were first noticed in the process of investigating a series of communications satellite anomalies during the active solar maximum around 1990. This was followed by a study of data from the *Ulysses* spacecraft augmented with some from *Voyager II* and other spacecraft where it was found that much of the variability in the solar wind occurred at discrete frequencies. The higher frequencies, those with periods near 5 minutes, agreed closely with those of optically measured p–modes, while at lower frequencies they were similar to those predicted for gravity, or g–modes [18]. Given the prevailing dogma that the solar wind was turbulent, this paper was considered to be heresy and was attacked. Additional evidence countering these attacks was given in [19], and simulations described in [20] showed that discrete modes and turbulence could coexist. Compounding the heresy, [21] suggested that the strongest *observational* evidence for turbulence in the solar wind, the $-5/3$ spectral index[1] of many power spectra in the solar wind was a misattribution. What was being observed was a fossil of turbulence in the solar convection zone.

During the time that the the events of the preceding paragraph were occurring in the space physics and engineering literature, three papers [22, 23, 24], appeared in the seismic literature showing that the Earth's normal modes were continuously excited. More importantly, all three of these discovery papers noted the presence of unusual modes in their data. These papers on "seismic hum" launched a small industry attempting to find a mechanism that provided sufficient energy, roughly that of a magnitude 6 earthquake *every day*, to excite the modes. Inexplicably, particularly in

[1] The spectral index is the slope of a spectrum on a log–power *vs* log–frequency plot. An index of $-5/3$, corresponding to a spectrum with the shape $c/f^{5/3}$ over a wide frequency range, is commonly taken as the "signature" of Kolmogorov turbulence.

view of the fact that normal modes of the Earth have been studied since Lord Kelvin's 1863 dispute with geologists on whether Earth's interior was primarily solid [25] or liquid [26] (see the history of this endeavour and the subsequent development of normal mode theory in Chapter 1 of [27]), the observations of the unexplained modes were largely ignored. Normal mode seismology is a mature theory with many modes having been predicted and most of the predicted ones observed, see *e.g.*, [28], and [29] for a recent update, so the discovery of new modes should have been taken far more seriously than it was.

We began studying this phenomena in 2001, attributing the strange modes to solar forcing at the AGU fall meeting in 2007, [30, 31] and, after numerous delays, published [16, 32]. The majority of these delays resulted from surprises in the data; it seemed that every time one looked at a new kind of data (or pairings of data sets) there was something new to explain. Was it a fluke, could we explain it, was it repeatible in other data sets, was there anything in the literature, and so on? Other delays were incurred by having to rewrite code that scaled poorly from a few hundred hours of data to years of data. An additional factor was the problem that these long data sets started to resolve the discrete solar modes in unexpected places. Their presence definitely complicates matters because Harvey [33] estimates that there are 10 million solar modes, so even if only the low–degree modes are reasonably resolved one must still consider many modes and one suspects that much of the "noise" between larger peaks may be mostly due to unresolved higher–degree modes. In such spectra, "smoothing" the estimate in the manner that has been common since the 1940s made an already complicated problem worse. The basic problem is that the "raw" multitaper estimates of spectra or coherence were high at modes, low between modes, and smoothing produced an estimate that was "wrong everywhere". Nonetheless, [16] shows clear evidence that "Earth's hum comes from the Sun", with coupling, most probably, through the interplanetary magnetic field. However, because the solar wind density, velocity, composition, and temperature all show strong modal features these components are almost certainly involved as well. This coupling both explains the unexpected modes noted in the three hum discovery papers (they are normal modes of the Sun) and also implies that there is more than sufficient energy to drive the observed hum. In addition to the numerous examples given in [34], work in progress shows evidence for solar modes in barometric pressure, [35, 36], ground–level ozone, and numerous other kinds of data. Daily data is not exempt because aliases of g–modes are readily detectable in solar noon 10.7 cm flux, [37]. Other low frequency solar modes can be measured both in interplanetary data [18] and daily temperatures [38, 39].

Taken together, the preceding paragraphs summarize the beginnings of a revolution in our understanding of "noise". Before this, it was common to assume that quantities such as barometric pressure, surface air temperature, ground–level ozone concentration, or seismic noise would be uncorrelated for measurements separated by modest distances in either time or space. Most statistical models of time series use low order autoregressive or moving average models (or their offshoots), all with reasonably fast mixing properties. From our viewpoint, however, one expects (and sees) correlations across continental and interplanetary scales. The departures from predictions from fast mixing in the time domain is even more extreme: at periods of a few minutes one expects to see solar p–modes. These modes are randomly excited and, depending on the mode, having damping times of a few days to months. At lower frequencies, those corresponding to periods from several days down to $\sim$ 30 minutes, one finds evidence for solar gravity modes. These are predicted [40] to have damping times between 10^5 to 10^7 *years*. Thus, instead of stationary mixing noise, one has a dense thicket of high–Q modes as a background in many applications. Very few of these modes have been identified [41], in the sense of assigning "quantum numbers" to them, so one is often faced with a bewildering forest of peaks in a spectrum. The frequencies of solar p–modes were found to vary with solar activity in [42] and the frequencies, line widths, etc. studied in numerous papers, *e.g.,* [43, 44, 45, 46, 47, 48]. The Sun's acoustic cutoff frequency also varies by ~ 100 to ~ 150 μHz with activity [49] so the number of p–modes that are observed will also vary with activity. Similarly, the frequencies of g–modes are also predicted [50] to vary with irradiance. Consequently, taking exceptionally long series will give the expected improvement in frequency resolution, but this may not correspond to an improvement in detectability because the mode's frequency may shift out of the resolution bandwidth. To conclude this section, we note that although solar modes can have exceptionally long lifetimes, they are described using essentially Newtonian mechanics and have nothing to do with "long–memory" processes [51, 52].

3. STATIONARITY

The rise of big data is changing our understanding of stationarity. The increasing availability of continuous time series creates opportunities for non-stationary processes to be observed and identified. Going forward, stationarity will play much the same role as Gaussian distributions. It is useful because stationarity, particularly when combined with the Gaussian assumption, allows one to do analysis where the resulting formulae are simple enough to understand. The importance of having such results is difficult to underestimate, but they must always be used in combination with Martin Wilk's comment "The hallmark of good science is

that it uses models and theories but never believes them."

4. SOME HISTORY

Looking back over the past few decades, a major change has occurred from simply trying to estimate a spectrum, to concentrating on descriptive statistics of the spectrum estimates. As a specific example, in the discussion of [53] Dr I. T. Jolliffe commented on "the large number of peaks that appear in the spectral estimates". However, we had gone out of our way to choose relatively simple examples, ones with no more than half–a–dozen or so peaks. The data series used in those examples were all relatively short, 1000 samples or less. In much recent data, data block of 50,000 to 500,000 samples are routine and we would be surprised if the spectrum contained less than a few hundred significant peaks and finding several thousand lines is not unusual. This implies many changes:

First, [53] was published in 1979, before the invention of the multitaper estimate in [54] in 1982, so the spectrum estimates being considered were single–taper direct estimates. Spectrum estimates made with a single data taper have a nominal χ^2_2 distribution, that is chi–squared with two degrees–of–freedom (DoF), and so have atrocious sampling properties such as large variance. If one has a sample of M χ^2_2 independent random variates, z_i, all with mean 1, their minimum is expected to be $\mathbf{E}\{z_{min}\} = 1/M$ while their maximum $\mathbf{E}\{z_{max}\} \sim \ln M + \gamma + \frac{1}{2M}$ where $\gamma = 0.57721566\ldots$ is Euler's constant. In otherwords, one expects the ratio of largest to smallest values in a random sample to be $\sim M \ln 1.78M$ so, for $M = 100,000$ the range of the spectrum estimates will be 1.2×10^6. We emphasize that this is in white noise with absolutely nothing interesting happening.

Second, if one is doing exploratory spectrum estimation and takes, for example, the 99% significance level as an initial threshold for "interesting" features one knows by definition that about 1% of the the estimates will lie above it. Single–taper estimates, however, tend to distribute this 1% in two narrow peaks per Rayleigh resolution [36] Figure 2, so our hypothetical spectrum will have ≈ 2200 spurious peaks above the 99% level.

Third, it was proven by Lord Rayleigh [55] that the periodogram is inconsistent, that is its variance does not decrease with sample size. This applies to all single–taper spectrum estimates.

Fourth, the sidelobes of many direct estimates [56] are so high that these estimates cannot cope with the range of much modern data. For example, the spectrum of the barometer data shown in Figure 1 of [36] has a range $\gtrsim 10^{13}$ so many estimates will be biased by several orders of magnitude. The variance of a direct estimate is roughly the square of its expected value, $(EV + bias)^2$, so may be too high by ridiculously large factors. These are *not* the sampling statistics one needs for scientific discovery or engineering analysis. In contrast, the sampling properties of a multitaper estimate with 20 DoF have a lower, 0.001%, point of 0.166 and an upper, 99.999%, point of 2.952. For the $M = 100,000$ sample example considered above, this reduces the range from $> 10^6$ to ~ 17.8 and the ≈ 2200 spurious peaks above the 99% level to about 730.

During approximately the same time interval, there was a similar improvement in instrumentation. As a specific example, the spectral windows plotted in Blackman and Tukey's 1958 book on spectrum estimation, covered such a small range that many could be plotted on a *linear* scale and one can see the sidelobes; see, *e.g.*, [57] Figure 15(f), 21(a). Similarly, the appendix of Bogert *et al.*'s 1967 paper introducing the cepstrum, [58], contains a comparison of hand and digital voltmeter digitizion of seismic records to three decimals. Progress, however, was rapid and, *e.g.* the gauging system for **WT4** millimeter waveguide [8] was routinely producing 25–bit data by 1975. The analysis of this high–quality data drove developments in robustness, required the efficient prolate data tapers, [5, 6], and led to the development of multitaper methods, [54].

While the impetus for the development of multitaper spectral analysis techniques was driven by the WT4 millimeter waveguide analysis, many other instrumentation, computing, and communications innovations started driving the need for these tools to analyze big data. In the context of this paper we extend the use of the term "big data" beyond a large number of bytes to include long time series spanning years of continuous data, high dynamic range data, data arrays created by thousands of sensors, and the integration of multidisciplinary sensor types.

Since this paper is focused on "Big Data" applications of multitaper spectral analysis, it is important to start in the time domain where the original data are recorded. To maximize the efficiency of any spectral technique, it is necessary to maintain a constant sampling rate. For finite duration local experiments, this has been a tractable problem for many years. However, when we shift to long term continuous measurements on a global scale, this becomes a very difficult problem. Radio time–synchronization services started in the 1960s with the introduction of time information into WWV, WWVB [59] and their international siblings introduced the possibility of distributed timing accuracies of 1 millisecond for WWV and 0.1 millisecond for WWVB. The advent of GPS–based time synchronization systems improved time accuracy for distributed systems to better than 1 microsecond [60]. Due to size and power requirements, long term continuous measurement and observing systems needing accurate timing also had an implicit requirement for continuous line power, preferably with some form of

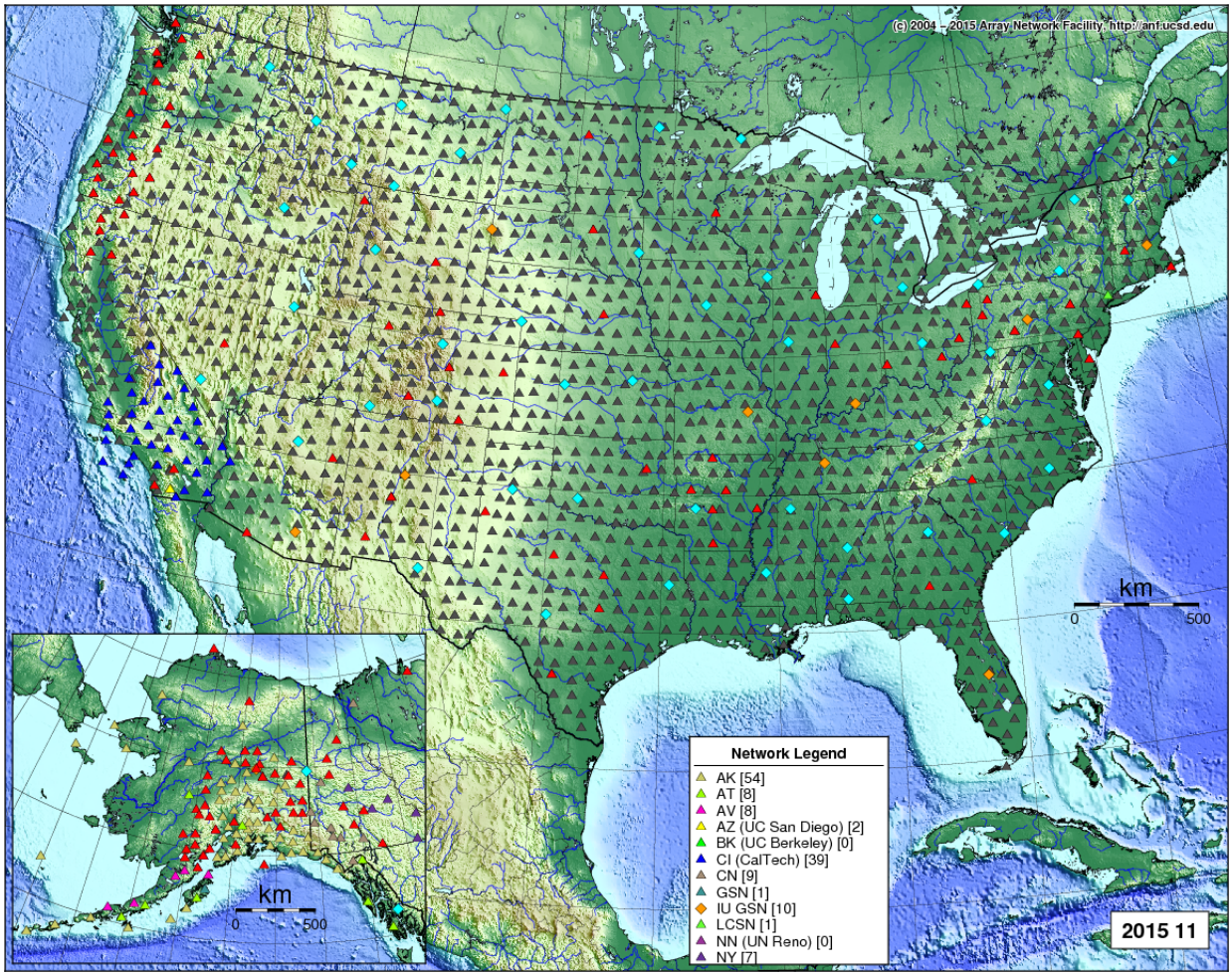

Fig. 1. Cumulative deployment of USArray Transportable Array from 2004–2015. Core TA stations are shown as red triangles (operating) or grey triangles (decommissioned). Other contributed stations from other networks are identified in the Network Legend box.

uninterruptible power supply. In the late 1990's low power GPS–based clocks started to become available that allowed for continuously disciplined clock oscillators, eliminating the need for timing corrections in systems that used them. Coupled with the rapid evolution of low power integrated circuits, we see an escalation of distributed sensor networks with high quality time resolution.

One important attribute of multitaper spectral analysis is the capability to investigate data with high dynamic range. Commercially available analog-to-digital converters (adc) based on integrated circuits, digital data acquisition systems using 8-bit adc with 48 db dynamic range started to be available in the 1970s and have evolved to the now ubiquitous 24-bit adc with 144 db dynamic range. In parallel, passive electromechanical sensors have transitioned into high–dynamic range active feedback sensors matching the increased dynamic range provided by each generation of adc.

Other factors needed to generate large datasets to be processed by multitaper–driven algorithms include improved computation, storage, and communications. Computationally, processors transitioned from typically 16-bit platforms in the 1970s to the current standard 64-bit platforms that are now able to easily handle the computational requirements. Local low-power flash storage at remote field sites is now on the order of terabytes and at data centers large numbers of terabyte to petabytes of data storage are now available. For distributed sensor networks there are now IP networks over wireless and satellite communications that provide for essentially lossless data telemetry if an appropriate retransmission protocol is used. In the wired environment, the use of fiber optics can enable effective use of cloud computing

and cloud storage for large scale computing problems.

The confluence of high quality time and the ability to record long gap–free time series with high dynamic range systems, coupled with advances in computing, storage, telemetry, communication, and power systems create the opportunity to effectively use multitaper spectral analysis techniques on large datasets. One example system is the NSF Earthscope USArray Transportable Array (TA). This system was designed to simultaneously record 400 autonomous stations with real time telemetry in continuous time series. A total of 1738 TA stations were deployed in a nominal 70 km grid starting in the western US in 2004 and rolling eastward across the contiguous 48 states finishing in 2015 (Figure 1). The TA in the continental US had a median station deployment time of one year and 315 days. The longest operational station is still currently operating for 11 years and 207 days. All stations were sampled with GPS continuously disciplined clock oscillators providing better than 1 microsecond timing accuracy across the whole sensor network. As of the time of this writing, a total of 1,290,101 station–days of continuous data have been recorded, with 26,206 unique science sensor station-channels and 36,639 unique system state-of-health channels. To date, over 17.7 terasamples of science data has been recorded using 70.8 Tbytes storage (23 TBytes using lossless compression). To date, this TA data has been used in over 100 PhD Theses and 200 refereed journal papers, so we will not even attempt to summarize it here.

Currently, in looking for "descriptive" summaries of power spectra it is becoming common to iterate the process. The first example of such a procedure that we know, the cepstrum, [58] took a Fourier transform of the logarithm of an estimated spectrum to find an echo. This was rapidly generalized, [61], and today one might use a second multitaper estimate and F–test to identify an echo. Second, taking the SVD of a log–spectrogram, [62] decomposes the matrix into left and right eigenvectors where, *e.g.* the left eigenvectors are simply new time series. Computing their spectra often shows systematic features in nonstationary data. Below, in §10.2, we show a more direct approach with spectral indices of interplanetary data.

5. INITIAL DATA SCREENING

Almost the first advice given in any text on time series analysis or multivariate statistics is "plot the data". We agree with this but, with large amounts of data, there are practical difficulties: First, both graphical devices and the human eye have finite resolution, so plotting a million-sample time series on a device with a resolution of a few thousand pixels is pointless. Second, the point of plotting the data is to spot outliers, instrument malfunctions, data gaps, and regions that appear different from the rest and, possibly excluding large outliers, this cannot be done without adequate graphical resolution. Third, this advice is given for a good reason: the number of things that can go wrong is so large that it does not seem possible to categorize them or to automate detection. The human eye and brain seems to be the best general pattern recognition device available, so plot the data. Fourth, this too has limitations; plotting data at high enough resolution to see peculiar features often requires the equivalent of several hundred plots, and looking through such a mass of data produces fatigue, so anomalies on the first frame or two will probably be detected, those on the equivalent of frame 450 may not be.

This has resulted in a variety of methods for screening data ranging from simple plots of order statistics, to running Karhunen-Loève expansions, multitaper spectrograms and coherences, which are appropriate for big data problems.

6. TECHNICAL BACKGROUND

To make the following sections comprehensible we need some notation so we begin with a short section on random processes and one on Slepian sequences. Assume one is given a sample of data, $x(t)$ for $t = 0, 1, \cdots, N-1$ from some physical process for the purpose of estimating a spectrum. The spectrum is a function of frequency, denoted by f that, for a unit sampling rate, $\delta t = 1$, has principal domain $[-\frac{1}{2}, \frac{1}{2})$. Note that f is ordinary cyclic frequency, not radian frequency. That is, if $\delta t = 1$ second, then f is in units of cycles per second, defined in the S.I. system as Hertz (Hz). Note that frequency is usually considered to be real–valued, but can be complex–valued. We now take the Fourier transform of the time series,

$$y(f) = \sum_{t=0}^{N-1} x(t)\, e^{-i2\pi ft} \,. \tag{1}$$

The Fourier transform is a trivially sufficient statistic because one may take the inverse Fourier transform

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} y(f)\, e^{i2\pi ft}\, df = \begin{cases} x(n) & 0 \le n \le N-1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

and recover the data.

We now *assume* that the sample is from a second–order (or covariance) stationary random process. In this section we also assume, (see §8.1), that the process is zero–mean, that is $\mathbf{E}\{x(t)\} = 0$: this implies that $\mathbf{E}\{dX(f)\} = 0$, where $\mathbf{E}\{\cdot\}$ denotes the expected value operator. This further implies a spectral representation [63, 64, 65]

$$x(t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi\xi t} dX(\xi) \tag{3}$$

for the process. The spectrum is a decomposition of variance into a function of frequency, i.e., the average power in

a small band $(f, f+df)$, and formally defined [64] in terms of the spectral representation as

$$S(f)df = \mathbf{E}\{\,|\,dX(f)\,|^2\,\}\,. \tag{4}$$

The stationarity assumption implies that the autocovariance function

$$R(\tau) = \mathbf{E}\{x(t+\tau)\cdot x^*(t)\,\} \tag{5}$$

is *not* a function of time, t, and this has the further implication that

$$\mathbf{E}\{dX(f)\,dX^*(\xi)\,\} = S(f)\,\delta(f-\xi)\,df\,d\xi \tag{6}$$

where $\delta()$ is the Dirac delta function. Stationarity requires that signal elements at distinct frequencies are uncorrelated. One also notes that, using the spectral representation (3) in (5), then applying (6), gives

$$R(\tau) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S(f)\,e^{i2\pi f\tau}\,df\;, \tag{7}$$

the *Einstein–Wiener–Khintchine theorem*. (It was discovered independently by Wiener [11] and Khintchine [66] in the 1930, but Einstein [67] had preceded them in a short note in 1914. The great Russian probabilist Yaglom, [68] commented that Einstein's derivation is more satisfactory than either of the later two.)

When the process is nonstationary but harmonizable, (6) must be replaced with the Loève spectrum and one has

$$\mathbf{E}\{dX(f_1)\,dX^*(f_2)\,\} = S_L(f_1,f_2)\,. \tag{8}$$

This was introduced by Loève in the 1940's (see the references in [69]) but, recalling that stationary processes were still considered difficult and obscure, they have not been used as much as they should have been. Rotating the (f_1, f_2) coordinates by $45°$ and taking a Fourier transform on the "nonstationary" frequency gives a spectrogram.

One must be cautious in using spectral representations for reasons originally noted in [70], and commented on in [13]. Also remember that the spectral representation is accurately named, it is a *representation* whose main use is as a mathematical tool for assessing properties of estimators. We would be surprised to learn of a physical process that originates in such a way.

6.1. Slepian sequences

In a famous set of papers entitled *Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty* begun in 1965 [71], Slepian, Landau, and Pollak established bounds on simultaneous time–frequency concentrations, the uncertainty principle when boundaries are imposed. Part five (V) of this series [72] described the discrete case where one has N time points at $t = 0, 1, \cdots, N-1$ and continuous frequency, $-\frac{1}{2} \le f < \frac{1}{2}$. Given a unit norm sequence, v_t, defined on $[0, N-1]$, the integral of its magnitude–squared Fourier transform, $V(f)$, over the Nyquist band $[-\frac{1}{2}, \frac{1}{2})$ is, by Parseval's theorem, also unity. That is

$$1 = \sum_{t=0}^{N-1} |v_t|^2 \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} |V(f)|^2\,df \tag{9}$$

Slepian found the sequences with the largest fraction of their energy, λ_k, in the frequency band $(-W, W)$, that is

$$\lambda = \int_{-W}^{W} |V(f)|^2\,df\,. \tag{10}$$

The optimum sequences are the *discrete prolate spheroidal sequences*, (DPSS's), now known as *Slepian sequences*. They are the real, orthonormal eigenvectors of the Toeplitz matrix eigenvalue equation

$$\lambda_k(N,W)\,v_t^{(k)}(N,W) = \sum_{n=0}^{N-1} \frac{\sin 2\pi W(t-n)}{\pi(t-n)}\,v_n^{(k)}(N,W)\,. \tag{11}$$

They are ordered by their corresponding eigenvalues $1 > \lambda_0 > \lambda_1 \cdots > \lambda_{N-1} > 0$, with λ_k giving the energy concentration of the k^{th} sequence in $(-W, W)$. Also, the k^{th} sequence has k zero crossings in $[0, N-1]$ and the corresponding Slepian functions,

$$V_k(N,W;f) = \sum_{n=0}^{N-1} v_n^{(k)}(N,W)\,e^{-i2\pi nf}\;, \tag{12}$$

have k zeroes in $(-W, W)$; see Figure 2. The functions are doubly orthogonal, first on the full frequency range

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} V_j(N,W;f)V_k^*(N,W;f)df = \delta_{j,k}\;, \tag{13}$$

and, second, on the inner (or local) band,

$$\int_{-W}^{W} V_j(N,W;f)V_k^*(N,W;f)df = \lambda_k\delta_{j,k}\,. \tag{14}$$

That is they are orthonormal on $[-\frac{1}{2},\frac{1}{2})$, and orthogonal on $(-W, W)$. We omit the explicit dependence on N and W so, e.g., the k^{th} sequence is denoted by $v_t^{(k)}$ and the eigenvalues by λ_k. Of these, the first $K \lessapprox \lfloor 2NW \rfloor$ eigenvalues are extremely close to one and, for example, with $NW = 6$, $1-\lambda_0 \approx 1.31\times 10^{-15}$ and with $NW = 10$, $1-\lambda_0 \approx 3.05\times 10^{-26}$. Section VI of [34] and Figure 3, below, shows estimates of spectra of barometer data where the range (max/min) is $\sim 10^{16}$, with the bulk of the power at very low frequencies, so ordinary spectrum estimates are dominated by bias from the low frequency range. In this

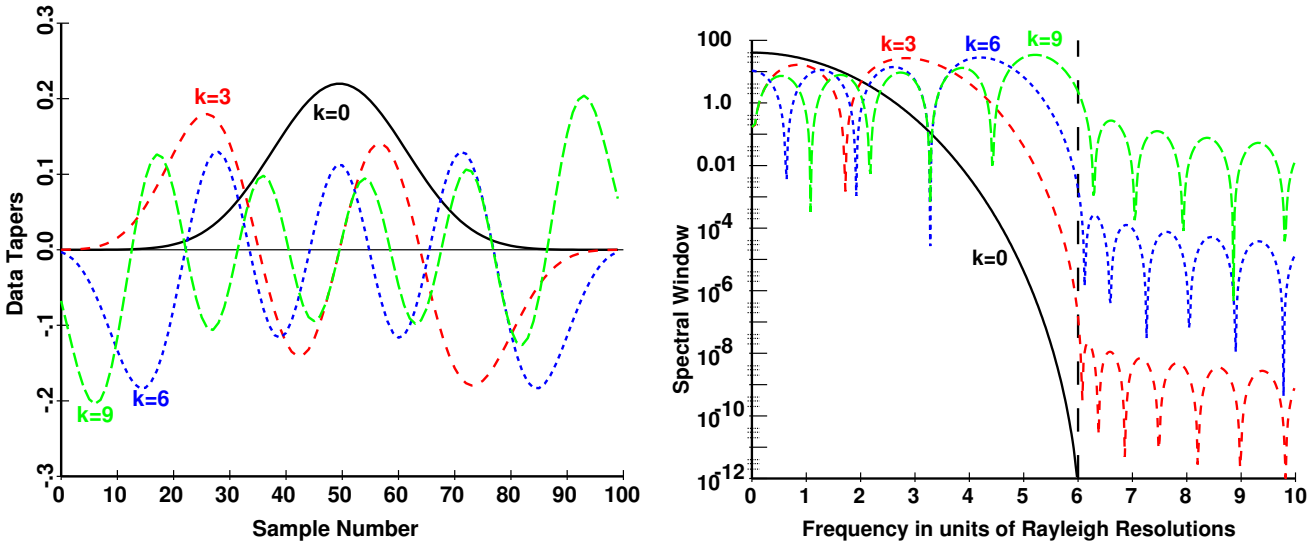

Fig. 2. The left panel shows Slepian sequences for $NW = 6$, $N = 100$, and $k = 0, 3, 6$, and 9. The right panel has the corresponding spectral windows, $|V_k(f)|^2$ with frequency expressed in units of Rayleigh resolutions, and the dashed vertical line showing the positive bandedge. The total power in the outer band for the k^{th} sequence is $1 - \lambda_k$, here 2.9×10^{-14}, 4.7×10^{-10}, 6.7×10^{-6}, and 0.011 for the sequences shown. The zeroes of the $k = 3$ and $k = 9$ functions at the origins are supressed and the sidelobe levels are clipped below 10^{-12}.

case, tapers with extremely strong out-of-band suppression are needed, and the Slepian sequences serve nicely.

Slepian's analysis defines the dimensionality of the time–frequency region $[0, N-1] \times (-W, W)$. Many of the properties of Slepian sequences depend on the *time–bandwidth product* $C_{\mathcal{R}} = NW$ as opposed to N and W individually. The $K = 2NW = 2C_{\mathcal{R}}$ lowest–order sequences are a complete bases for this subspace. In spectrum estimation problems, W is the bandwidth of the estimate and is a factor of $C_{\mathcal{R}}$, typically 2 to 10, of the Rayleigh resolution[2], $\mathcal{R} = 1/(N\Delta t)$. As N increases, properties of the Slepian sequences rapidly approach those of the continuous–time prolate spheroidal wave functions. Details are given in [72], and Appendix A of [73] gives a summary. Finally, note that while (11) is analytically convenient, it is not recommended for numerical use. To compute the sequences we use the tridiagonal form with separate calculations of the odd and even sequences given in Appendix B of [73].

The choice of $C_{\mathcal{R}}$ determines the number of tapers, $K \sim \lfloor 2C_{\mathcal{R}} \rfloor$, the degrees of freedom, $\nu \approx 2K \lesssim 4C_{\mathcal{R}}$, and bias properties of the estimate. The broad–band bias, i.e., the bias at f contributed by frequencies outside the band $(f - W, f + W)$, decreases exponentially with $C_{\mathcal{R}}$.

[2] $C_{\mathcal{R}}$ seems to be increasing in time. In [54] $C_{\mathcal{R}} = 4$ was "standard" but the spectral range of modern data often requires $C_{\mathcal{R}} \geq 8$ or 10, even with prewhitening.

7. MULTITAPER SPECTRUM ESTIMATES

Temporarily shifting the origin of the Fourier transforms in (3) and (1) to the center of the observation interval, that is, $\bar{t} = (N-1)/2$, so replacing t with $t - \frac{N-1}{2}$ in these two equations and merging them gives

$$y(f) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\sin N\pi(f-\xi)}{\sin \pi(f-\xi)} dX(\xi) \; , \qquad (15)$$

so the Fourier transform of the data can be written in terms of the spectral representation, giving what Parzen has called "the fundamental equation of spectrum estimation". It is a convolution of the Dirichlet (or diffraction) kernel and dX.

In multitaper analysis one thinks of (15) as an *integral equation* and attempts to solve it. Recognizing that a unique solution is impossible, one goes for the best one can get. This is done by local least–squares in a band $(f-W, f+W)$, with sequential bands estimated by sliding the band along in frequency. Recalling the standard method of solving Fredholm integral equations of the first kind, the eigencoefficients (the coefficients of the eigenfunctions) are defined by

$$y_k(f) = \frac{1}{\lambda_k} \int_{-W}^{W} y(f-\xi)\, V_k(\xi) d\xi \; . \qquad (16)$$

After some algebra this becomes

$$y_k(f) = \sum_{t=0}^{N-1} x(t)\; v_t^{(k)}\; e^{-i2\pi f t} \; , \qquad (17)$$

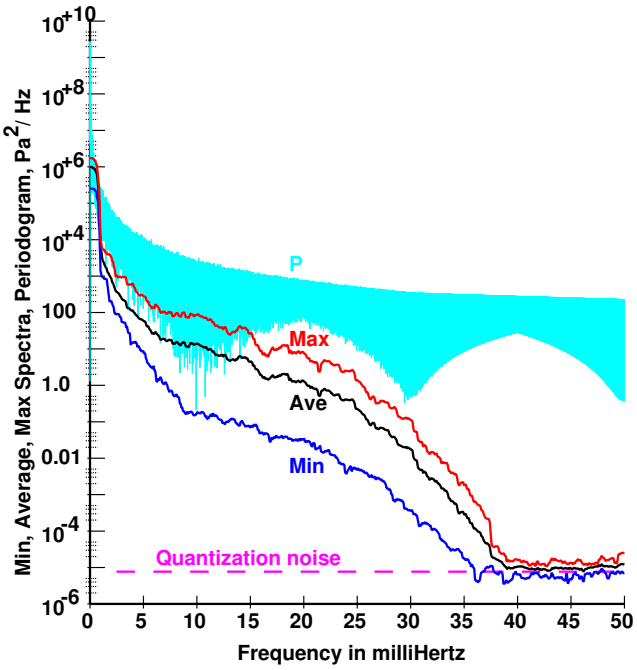

Fig. 3. Spectrum estimates for eight adjacent sections of 1000 samples each of barometric pressure data from Black Forest Observatory. The top "curve" (cyan) is the periodogram from the first section. The three lower curves are the minimum, arithmetic average, and maximum of the eight multitaper estimates. These were computed with $NW = 10$ and $K = 12$ tapers on each section. Note that above $\sim 35\,m$Hz the periodogram is biased by a factor of $> 10^7$. Any estimate based on sample autocovariances will have similar bias problems as the periodogram.

the standard definition of the eigencoefficients. The k^{th} eigencoefficient is thus the Fourier transform of the data multiplied by the k^{th} Slepian sequence. We denote its squared magnitude by

$$S_k(f) = |\, y_k(f)\,|^2 \;, \tag{18}$$

the k^{th} eigenspectrum.

In anticipation of the examples, consider a multivariate time series consisting of a set of P time series, $x_p(t)$ with $1 \le p \le P$. These series are assumed to be time–aligned with N samples, that is $t = 0, 1, \ldots, N-1$, and to represent (individually) samples of a second–order stationary process.

As above, begin by computing the multitaper eigencoefficients

$$\hat{x}_{k,p}(f) = d_{k,p}(f) \sum_{t=0}^{N-1} x_p(t)\, v_t^{(k)} e^{-i2\pi f t} \tag{19}$$

where the $d_{k,p}(f)$ are the adaptive weights (see [54] Section V for details). When the spectrum is white, the weights, $d_{k,p}(f) = \sqrt{\lambda_k}$ and so the weights are approximately unity for $k < K$.

To estimate the spectral matrix, one first collects the weighted eigencoefficients, (19), into a $K \times P$ complex matrix

$$\mathbf{X}(f) = \begin{bmatrix} \hat{x}_{0,1}(f) & & \hat{x}_{0,P}(f) \\ \vdots & \cdots & \vdots \\ \hat{x}_{K-1,1}(f) & & \hat{x}_{K-1,P}(f) \end{bmatrix} , \tag{20}$$

so the formal estimate of the $P \times P$ spectral matrix, $\widehat{\mathbf{S}}(f)$ is

$$\widehat{\mathbf{S}}(f) = \frac{1}{K} \mathbf{X}^\dagger(f) \mathbf{X}(f) \tag{21}$$

where the $\dagger$ represents a conjugate–transpose. Here, again, there has been a change from the theory (and what practice there was) of Wiener's 1930 "Generalized Harmonic Analysis" [11] through to modern texts, *e.g.*, [74, 75] where the spectral matrix plays a central role. Our experience is that one rarely computes $\widehat{\mathbf{S}}(f)$ directly but usually the diagonal elements (power spectra of the individual series), coherences and phases (often with jackknife error estimates), or takes the singular value decomposition (SVD) of $\mathbf{X}(f)$. Some examples are shown later.

Due to the superior spectral leakage properties of the multitaper spectrum, Figure 3 demonstrates why one should go to this effort. It shows the minimum, average, and maximum of eight multitaper (MT) spectra made on non–overlapping ~ 3 hour blocks with a periodogram computed on the first block. Over most of the frequency range the periodogram is biased by at least a factor of 1000, a large error. Below $\sim 35\,m$Hz, the ratio of the maximum to minimum of the eight MT estimates varies between ~ 100 to ~ 500, not the ~ 3.5 that one would expect from Hartley's "maximum F–ratio" [76] indicating reasonable non-stationarity during the day. Above $\sim 35\,m$Hz the average spectrum is close to that expected from quantization noise, $\sim 9.1 \times 10^{-6}\ Pa^2$/Hz, and near the expected range.

7.1. Multivariate Analysis and Likelihood

One should note that beginning with the matrix $\mathbf{X}(f)$, or its many variants, one can apply most of the standard methods of multivariate analysis with some minor differences. First, the usual column of "ones" is missing; second, the matrices are all functions of frequency; and third, everything is complex–valued, so "Gaussian" and "normal" are not synomonous. In compensation, there is a long history showing that Fourier transforms of "reasonable" data have an approximately Gaussian distribution, so tests such as Bartlett's "M" that are often avoided because of their sensitivity to departures from Gaussian, work quite well.

We assume the process is "locally white", that is that the true spectrum is approximately constant on $(f-W, f+W)$,

the *rows* of $\mathbf{X}(f)$ are approximately uncorrelated and $\widehat{\mathbf{S}}(f)$ will have a complex Wishart distribution with $2K$ degrees–of–freedom (DoF). For the matrix to be nonsingular, one must have $K > P$. For scalar series the spectrum estimates are distributed as χ^2_{2K}, are approximately unbiased and approximately maximum–likelihood, [77].

If one collects the eigencoefficients (17) into a vector

$$\mathbf{Y}(f) = [y_0(f), y_1(f), \cdots, y_{K-1}(f)]^T \tag{22}$$

its covariance matrix

$$\mathbf{C} = \mathbf{E}\{\mathbf{Y}(f)\mathbf{Y}^\dagger(f)\} \tag{23}$$

whose elements may be evaluated using the spectral representation

$$\mathbf{C}_{jk} = \int_{-\frac{1}{2}}^{\frac{1}{2}} S(f-\xi)\, \mathcal{V}_j(\xi)\, \mathcal{V}_k^\star(\xi) d\xi \ . \tag{24}$$

The concentration of the Slepian functions implies that one may replace the integral over $[-\frac{1}{2}, \frac{1}{2})$ with one over $(-W, W)$ with neglible error.

$$\mathbf{C}_{jk}(f) \approx \int_{-W}^{W} S_o(f-\xi)\, \mathcal{V}_j(\xi)\, \mathcal{V}_k^\star(\xi) d\xi \ . \tag{25}$$

This means that one can test for specific shapes in the spectrum, that is, $\hat{S}(f)$ is proportional to $S_o(f)$ in the range $(f-W, f+W)$, with log–likelihood

$$\mathcal{L}(f; S_o) = \ln|C| + \mathbf{Y}^\dagger(f)\mathbf{C}^{-1}\mathbf{Y} \ . \tag{26}$$

In the band–limited approximation, the eigenvalues of $\mathbf{C}(f)$ are the eigenvalues of the corresponding Karhunen–Loève expansion.

8. PERIODICITY, THE HARMONIC–F TEST AND T^2 TESTS

Generally speaking, *harmonic analysis* has come to mean the study of the line components in a spectrum without regard to whether they are at multiples of a common frequency or not. Unfortunately, since the techniques used for harmonic analysis have been virtually identical to those used for general spectrum estimation, the two names have been used almost interchangeably. The distinction, however, is crucial because it has been shown that one cannot specify both the spectrum and first–order probability density arbitrarily, [78]. In other words, the periodic components should be considered to be part of the mean–value function, not of the random part.

To make sense of harmonic analysis, it is essential to recognize that the assumption of "pure" line components is a convenient fiction, while often good for a moderate time span is rarely supportable over extended periods of time. This tends to further divide the subject by series length: in short series, *detection* and resolution of line components are major problems; for longer series the problems of interest typically concern the *structure* of the line; that is, could it be the result of a high–Q resonator driven by noise, or is it a sinusoid with time–varying amplitude or frequency or both. In cases where there is barely enough data to make a detection such distinctions are often ignored but, as the length and precision of data increases, so does the necessity of careful analysis.

In our applications we invariably assume that frequency is continuous because the frequencies that appear in data are determined by nature and do not depend on the choice of sampling interval or the size of the sample. Thus the conventional *Fourier frequencies* (those spaced $1/N$) are more–or–less irrelevant and one must zero–pad the data, use a "slow" Fourier transform, [79], or interpolate the eigencoefficients. (And even here, polynomial interpolation error formulae imply that the process is much more accurate if one zero pads by a factor of π or better before interpolating.) Experience is a hard teacher and we have learned to keep frequencies in Hertz — using a mixture of cycles/hour, cycles/day, etc. makes it all too easy to miss noting common frequencies in different data series[3].

Multitaper analysis [54] approached the problem of "mixed" spectra, that is, where line components are embedded in stationary background noise with a continuous spectrum, using a regression approach. The method described below extends this to multivariate and cyclostationary problems. With this approach, which consists simply of applying regression techniques to the eigencoefficients [54], the two problems of spectrum estimation and harmonic analysis are distinct. There is also no distinction between an ordinary constant mean and periodic components except for the trivial one that, in real–valued data, there are only half the degrees–of–freedom for the constant mean.

The distinction between the two problems is that *spectrum analysis* is the study of the second and higher moments of $dX(\nu)$ while the emphasis in *harmonic analysis* is on the first moment of $dX(\nu)$, effectively a Munk–Hasselmann [80] extension of the spectral representation. In such cases the process is usually described as having a non-zero mean value function consisting of a number of sinusoidal terms at various frequencies, plus perhaps a polynomial trend, plus a stationary random process of the type we have been considering in §6. In terms of the spectral representation this, in practice, amounts to having

$$\mathbf{E}\{dX(f)\} = \sum \mu_m \delta(f - f_m) \tag{27}$$

in place of the usual assumption $\mathbf{E}\{dX(f)\} = 0$. With

[3] We makes an exception for paleoclimates where the time step may be in millenia.

this definition the continuous portion of the spectrum is the second absolute *central* moment of $dX(f)$.

8.1. The Harmonic F-test

To show the derivation for the *harmonic F-test*, assume the simplest case of a single line component at frequency f_0 so that the eigencoefficients have a non-zero expected value

$$\mathbf{E}\{y_k(f)\} = \mu V_k(f - f_0)\ . \tag{28}$$

Again making the assumption that the continuous component of the spectrum near f_0 is slowly varying or "locally white" (25) shows that

$$\mathrm{Cov}\{y_k(f), y_j^*(f)\} \approx S(f) \cdot \delta_{j,k} \tag{29}$$

where the spectrum, $S(f)$, is the *continuous* spectrum and does not include the line power.

There are two obvious limiting methods to estimate μ; point regression at f_0, and integral regression in the neighborhood of f_0 with the obvious changes to either both coefficient weighting. The second case (integral regression) is more complex, and as it uses information from a wider bandwidth, is subject to noise from the same bandwidth. For this reason, point regression is more commonly used. In this (first) case one uses only the data at f_0 where one has the obvious relation

$$\mathbf{E}\{y_k(f_0)\} = \mu V_k(0) \tag{30}$$

and can estimate the mean, μ, by standard complex regression methods, [81],

$$\hat{\mu}(f) = \frac{\sum_{k=0}^{K-1} V_k(0) y_k(f)}{\sum_{k=0}^{K-1} V_k^2(0)} \tag{31}$$

This estimate is the high resolution estimate described in section 10 of [54] at $f = f_0$ before squaring. As the eigencoefficients are combined linearly we may write

$$\hat{\mu}(f) = \sum_{n=0}^{N-1} h_n(N, W) x_n e^{-i2\pi f(n - \frac{N-1}{2})} \tag{32}$$

where the effective harmonic analysis data window, $h_n(N, W)$, is given by

$$h_n(N, W) = \frac{\sum_{k=0}^{K-1} V_k(0)\, v_n^{(k)}(N, W)}{\sum_{k=0}^{K-1} V_k^2(0)} \tag{33}$$

The variance of the estimated mean depends on the local continuous part of the spectrum,

$$\mathbf{Var}\{\hat{\mu}(f)\} = \frac{S(f)}{\sum_{k=0}^{K-1} V_k^2(0)} \tag{34}$$

which is only slightly larger than $S(f)/N$. Subtracting their estimated means from the eigencoefficients gives an estimate of the continuous spectrum. Comparing this value of the background spectra with the power in the line component results in an F variance-ratio test, [82, 83], with 2 and $2K - 2$ DoF for the significance of the estimated line component,

$$F(f) = \frac{(K-1)|\hat{\mu}(f)|^2 \sum_{k=0}^{K-1} V_k(0)^2}{\sum_{k=0}^{K-1} |y_k(f) - \hat{\mu}(f) V_k(0)|^2}. \tag{35}$$

8.2. Comments on Significance Testing

When applied to climate data or data from some process correlated with climate, it is common for the harmonic F–test to produce an unusual number of lines at intermediate significance levels. With monthly or economic data, one must beware of "calendar frequencies" produced by the irregular lengths of the months [84]. These aside, the spacing between lines can provide useful information.

As an example, the Burgundy grape harvest series [14] beginning in 1370, records the day of the year when Pinot–Noir grapes were harvested in the vicinity of Dijon and is of importance because it is both the longest such series and particularly because it includes a long span of observations before the Maunder minimum. It was found that this interval, noted for a long span with no sunspots and "the little ice age", also marked a change in the spectrum of the grape harvest data [85] at about 1675. Taking the entire span of 634 years with $C_{\mathcal{R}} = 4$ and $K = 5$, the F–test found one frequency, 0.2412 c/y (a period of 4.146 ± 0.0022 years) with a significance of 99.9%.

The Pinot Noir series has a total of 49 frequencies in the F–test with significances above the 90% level. Ten percent of 634 suggests that these may be simply random noise but looking at their spacings one finds: 13 pairs with an average spacing of 18.63 years, close to the lunar nodal period of 18.613 years; 16 pairs spaced by 103.9 years, the Suess cycle from ^{14}C. This agrees well with a period between 103.97 and 104.14 years in [62]; 18 pairs spaced 80.02 years. The period of the Gleissberg cycle was given as 80 years [86], but estimates vary, see [87]; 23 pairs spaced an average of 53.79 years, the McKenzie cycle first identified in 1829 in English wheat prices, [88]; and 19 pairs at an average spacing of 34.98 years, possibly the second harmonic of the Suess cycle.

It seems unlikely that, if the frequencies were random, one would get these numbers of pairs of equally–spaced differences and adding the fact that they correspond to known solar and climate periods, we are inclined to the opinion that the climate is responding to a complicated forcing mechanism. However, as noted in [62], many of these periodic components do not appear as "additive periodicities", but as

modulations of the statistical structure of the process. Moreover, there are enough long periods in the climate system that shorter period terms may be modulated by the longer period components so the F–test will be supressed. The annual temperature cycle can be offset from the tropical year by precession and phase modulated by the solar cycle, see [89]. The lesson here is that, when working with long data series, do not expect to find behaviour that mimics simple models.

8.3. Common Periodic Components and the T^2 Test

Turning now to the periodic components with lower amplitudes, one suspects that these consist of a mixture of solar rotation (that varies between $\sim$ 25 days at the Sun's equator to $\sim$ 40 days at its poles) and related geometric effects, convection modes [90, 91, 34], possibly gravity (g) modes [92] or their aliases, and possibly unknown effects. Our object here is not to identify specific sources on the Sun so, for simplicity, we refer to all of these as simply "lines".

When considering a multivariate series of this type, at a line frequency the eigencoefficients will be the complex line amplitude times the value of the corresponding Slepian Function at zero frequency. Thus, returning to the derivation in §7, one expects that the columns of (20) will have a mean–value given by the zero–frequency Slepian function. That is, if series p contains a term at frequency f_m so that

$$x_p(t) = a_{p,m}\, \cos(2\pi f_m t + \theta_{p,m}) + \cdots \,, \qquad (36)$$

the corresponding eigencoefficient at a frequency f close to f_m will have an expected value

$$\mathbf{E}\{y_{k,p}(f)\} = \frac{a_{p,m}}{2} e^{i\theta_{p,m}}\, V_k(f - f_m) \,. \qquad (37)$$

Because the $V_k(f)$'s are very small for $|f| > W$ one can usually treat lines separated by more than W as being independent.

Denote by $\mathbf{V}$ the $K \times 1$ vector of $\{V_k(0)\}$'s and by $\mathbf{A}(f)$ the $P \times 1$ vector of complex amplitudes, $a/2\, \exp\{i\theta\}$ and write

$$\mathbf{X}(f) = \mathbf{V}\mathbf{A}^{\mathbf{T}}(f) + \mathbf{R}(f) \,. \qquad (38)$$

One now estimates $\mathbf{A}(f)$ by ordinary (complex) least–squares at each frequency on a fine mesh and, at each frequency, we can test the fit for significance using a complex T^2 statistic [93]. Formally, one computes the unnormalized residual spectrum

$$\mathbf{S}_R(f) = \mathbf{R}^\dagger(f)\mathbf{R}(f) \qquad (39)$$

and

$$T^2(f) = ||\mathbf{V}||^2 \frac{K-P}{P} \mathbf{A}^\dagger(f)\mathbf{S}_R^{-1}(f)\mathbf{A}(f) \qquad (40)$$

but, in practice, one reduces $\mathbf{R}(f)$ with a QR decomposition in the usual way. $T^2(f)$ has an F distribution with $2P$ and $2(K-P)$ DoF, see [94]. With a single series this test becomes the usual harmonic F-test derived in the previous section, and many of its properties are similar.

If one chooses the point, $\hat{f}$, where the harmonic F-test is maximized as an estimate of line frequency, its variance is

$$\mathbf{Var}\{\hat{f}\} = \frac{1}{(2\pi T)^2} \frac{6}{\rho(f)} \qquad (41)$$

where, as above, T is the total length of the series and $\rho(f)$ the local signal–to–noise power ratio. Other things remaining constant, the variance of the frequency estimate decreases as $1/T^3$ so the accuracy of a frequency estimate can be much better than the Rayleigh resolution. This implies that one must zero pad extensively or zero pad by π and interpolate.

To close this section on periodic components we reiterate the point [78] that one cannot specify both the spectrum and the first–order probability density independently. Experiments where one sets the significance level of the harmonic F–test, subtracts the estimated lines from the time–domain data, and plots a histogram of these residuals have favoured relatively low thresholds, that is, most of the data we have been studying has *many* lines. A related procedure is to note that multitaper estimates have a χ^2_ν distribution, where ν is the degrees of freedom at frequencies where lines are absent, but the presence of line components causes the distribution to be *non–central*, $\chi^2_{\nu,\lambda}$, where λ is the non–centrality parameter, the sum–of–squares of the periodic or modal components [36]. The signal–to–noise power ratio $\rho = \lambda/\nu$. The distinction between the central and non–central components has become increasingly important with big data sets because, once one has long enough data sets to resolve the various periodic and modal components, a spectrum estimate has a mixture distribution; non–central at modes and central between them. Examples are given in Plates 2–6 of [19] and Figures 5 and 6 of [36].

9. ROBUST ESTIMATES: HOW MUCH IS TOO MUCH?

Since Box's introduction [95] of the term "robust" into the statistics literature in 1953, the field has exploded and searching on "Robust Statistics" in Google Scholar produced a staggering 2.8 million hits. Having contributed several papers to this literature ourselves, see *e.g.*, [5, 6, 53, 96, 97, 98, 99, 100, 101, 102], we have both a vested interest but, also having been subjected to large quantities of seismic, geomagnetic, and space physics data, one begins to wonder if robustness was not "oversold". In sections §10.3 and §10.4 below we examine two examples of robust time series analysis and find: *first* that the "median spectrum"

commonly used in seismic noise studies can be seriously misleading; and *second*, the common practice of clipping extreme data may result in incorrect inferences about processes generating much data.

9.1. The Median Spectrum

To frame the discussion of the properties of the frequently used median spectrum, we will start with the procedures used to study seismic "hum" and seismic noise, *e.g.*, [22, 15]. Later, in section §10.4, the artifacts resulting from the median spectrum are demonstrated with specific data examples. The robust variant of the Welch Spectrum estimation procedure used in [22, 15] is:
1) Choose a block length from one to four or five days duration. This choice governs the frequency resolution of the estimate, and a block of duration T seconds will correspond to a Rayleigh resolution $\mathcal{R} = 1/T$. A one–day block will have a resolution $\mathcal{R} = 11.57\,\mu$Hz. The data blocks are usually offset by 50% of their duration and often skip intervals after large earthquakes. Assume that there are J blocks, and label the successive blocks by j with $1 \le j \le J$. Let N denote the number of data samples in a block. Standardize the sampling interval to $\delta t = 1$, and denote the data (possibly after removal of the local average, slope, and known effects such as tides) in the $j^{\rm th}$ block by $X_j(t)$ with $t = 0, 1, \cdots N-1$.
2) Taper the data on each block with a Hamming or better data taper $D(t)$, and compute the spectrum estimates

$$\hat{S}_j(f) = \mid \sum_{t=0}^{N-1} D(t)\, X_j(t)\, e^{-i2\pi f t} \mid^2 . \qquad (42)$$

3) At each frequency bin in the selected frequency range, sort the J estimates into increasing order

$$\hat{S}_{(1)}(f) \le \hat{S}_{(2)}(f) \le \cdots \le \hat{S}_{(J)}(f) \; , \qquad (43)$$

using the common statistical notation that a subscript (j) denotes the ordered values. Keep the median $\hat{S}_{(J/2)}(f)$ as the estimate. The motivation for this procedure is that one expects normal modes excited by large earthquakes to occur in the upper quantiles $\sim \hat{S}_{(\lesssim J)}(f)$ and the median excludes them.

There are dangers with such a procedure: *First*, for such problems the median has an efficiency of only $\sim$ 48 %, meaning that its variance is twice that of an efficient estimator, so assessing effects such as temporal dependence is unnecessarily difficult. *Second*, the estimate is biased. This does not prevent *identifying* modes, but the estimated power in the modes will be low by a factor of ≈ 1.35. This complicates the identification of forcing mechanisms. *Third*, the series may be non-stationary causing misidentification of peak frequencies in median spectrum.

To establish the baseline performance of this procedure we assume that the data is approximately a Gaussian coloured noise process as one expects from resonances driven by turbulence. We also assume that, apart from large earthquakes or other outliers occurring in some blocks, the process is stationary. The individual spectrum estimates (42) will have a central chi–square distribution with two degrees–of–freedom (DoF), that is,

$$z_j(f) = \frac{\hat{S}_j(f)}{\mathbf{E}\{\hat{S}_j(f)\}} \qquad (44)$$

will have the probability density function e^{-z}. In (44) the expected value, $\mathbf{E}\{\}$, of the estimate is

$$\mathbf{E}\{\hat{S}_j(f)\} = \int_{-\frac{1}{2}}^{\frac{1}{2}} |\tilde{D}(\xi)|^2 S(f-\xi) d\xi \qquad (45)$$

where $S(f)$ is the true spectrum of the process at frequency f and $\tilde{D}(f)$ is the Fourier transform of the data taper $D(t)$. Robust spectrum estimates from such distributions have been studied previously, [5, 6, 53], for general choices of the clipping level, and in [19], general χ^2_{2K} distributions were considered. Many of the problems occur because most of our experience with robustness is from *symmetric* distributions, and the exponential distribution is emphatically *not* symmetric.

9.2. A Robust Spectrum Estimate

Consider a set of J independent variates, z_j from a standard exponential distribution, $\exp\{-z\}$, with a population mean of one, that is $\mathbf{E}\{z_j\} = 1$. The z_j's are sorted into increasing order

$$z_{(1)} \le \cdots \le z_{(j)} \le \cdots \le z_{(J)} \; . \qquad (46)$$

Properties of the order statistics of the exponential distribution are available in Volume 2 of [103], page 500, and the expected value, or mean, of the $j^{\rm th}$ quantile is

$$\mathbf{E}\{z_{(j)}\} = \mathcal{M}_J(j) = \sum_{m=1}^{j} \frac{1}{J+1-m} \; , \qquad (47)$$

for $1 \le j \le J$. This means that if one repeats the experiment of taking J independent variates and keeps the $j^{\rm th}$ smallest sample from each experiment, they will have the mean $\mathcal{M}_J(j)$ and variance

$$\mathbf{Var}\{z_{(j)}\} = \sum_{m=1}^{j} \frac{1}{(J+1-m)^2} \; , \qquad (48)$$

For $j = 1$ (47) gives $\mathbf{E}\{z_{(1)}\} = 1/J$ and, for $j = J$ one has

$$\mathbf{E}\{z_{(J)}\} = \sum_{m=1}^{J} \frac{1}{m} \;\approx\; \ln J + \gamma + \frac{1}{2J} \qquad (49)$$

where $\gamma = 0.57721566\ldots$ is Euler's constant. That is, one expects the smallest and largest values in a random sample to be $1/J$ and $\sim \ln 1.78J$, respectively. We associate these order statistics with the probabilities

$$P_j = 1 - e^{-\mathcal{M}_J(j)} \ . \tag{50}$$

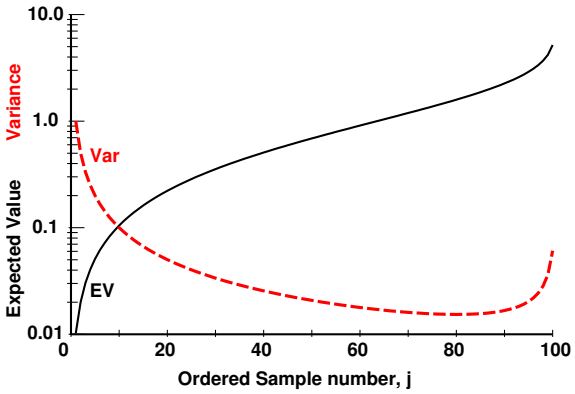

Fig. 4. The solid black line shows the expected value, $\mathcal{M}_{(j)}$ of the order statistics of a sample of size $J = 100$ independent exponential variates. The red dashed line shows the variance $\mathcal{V}_{100}(j)$ of the unbiased estimate (52).

Figure 4 shows the expected values of the order statistics from a sample of size $J = 100$ independent exponential variates and it can be seen that the lowest ten have expected values that are a factor of 10 to 100 below the mean of 1. Similarly, the top 15 are factors of 2 to 5 larger than the mean. This bias can be corrected by defining a gain factor

$$\mathcal{G}_j = \frac{1}{\mathcal{M}_J(j)} \tag{51}$$

so the estimated mean, say $\widehat{U}_{[j]}$, from the j^{th} order statistic

$$\widehat{U}_{[j]} = \mathcal{G}_j \cdot z_{(j)} \ , \tag{52}$$

is unbiased, but the $\widehat{U}_{[j]}$'s are not monotone. Within the expected statistical fluctuations, all values of j should give an unbiased estimate of the spectrum. The variance of $\widehat{U}_{[j]}$, scaling (48) by the squared gain, is

$$\mathcal{V}_{(j)} = \mathbf{Var}\{\widehat{U}_{[j]}\} = \mathcal{G}_j^2 \, \mathbf{Var}\{z_{(j)}\} \ . \tag{53}$$

The DoF of $\widehat{U}_{[j]}(f)$ can be estimated from the first two moments of a chi–squared distribution, $\hat{\nu} = 2/\mathcal{V}_{(j)}$ to set *approximate* confidence intervals. Returning to the example with $J = 100$, the red dashed line in Figure 4 shows the variance of the unbiased estimates as a function of j so both small and large quantiles should be avoided. For an *uncontaminated* sample, the minimum variance (maximum–likelihood) estimate is the arithmetic mean of the z_j's with a variance of $1/J$ or 0.01 in this case. Thus none of the single quantile estimates is particularly efficient. The median has an efficiency of $\sim 48\%$ and the minimum variance, 0.0154 (an efficiency of 65%), occurs at $j = 80$. Also note that one cannot simply average the various $\widehat{U}_{[j]}$'s and obtain a much better estimate because the sorting causes them to be correlated. Estimates that use a linear combination of all the variates, $z_{(j)}$, up to some limit, say $J' < J$ are discussed in [5, 6]. Finally, the spectra discussed in the remainder of the paper are all computed with sections overlapping by 50% and so are clearly *not* independent. For a noise–like process the correlation between these spectra on adjacent sections, see section 3.3 of [5], is only 3.338%.

10. LARGE SCALE DATA EXAMPLES

The previous three sections have discussed the multitaper spectral analysis techniques and their applications and implications for "big data" sets. In this section we will present results from signals originating in the Sun, which working out from the Sun are observed by interplanetary spacecraft, geomagnetic observatories, and by finally at terrestrial seismic stations. In addition we observe artifacts of these solar signals in long duration surface temperature records. Before doing this, however, we define four families of normal modes, two solar and two terrestrial, that are referred to in the following section. The distinction is important because all four of these families are observed in terrestrial seismic data.

10.1. Solar and Terrestrial Normal Modes

Modes are usually described by a spherical harmonic expansion and consequently have a set of "quantum numbers", l, n, and m. Here l represents the spherical harmonic degree and gives the number of nodal lines in latitude. The radial order, n, is the number of zeroes along a radial line. Finally, m gives the azimuthal order, that is the number of cycles in longitude with $-l \le m \le +l$. Except where necessary, we omit the m index.

The solar modes of interest here are the *pressure* (p–modes) and potentially *gravity* (g–modes), denoted as $p_{l,n,m}$ and $g_{l,n,m}$ respectively. Introductions to helioseismology are given in [104] and [105].

Pressure, or p–modes are acoustic standing waves in the frequency range ~ 250 μHz to $\gtrsim 5000$ μHz, so their periods range from about an hour down to a few minutes. Their frequencies are given asymptotically by

$$\nu_{l,n,m} \asymp \nu_r \cdot m \ + \ \nu_0 \left[\tfrac{1}{2}(l + \tfrac{1}{2}) + n + \alpha\right] + \ldots \tag{54}$$

where $\nu_0 \approx 135$ μHz and the rotational splitting term ν_r is approximately 440 nHz, but depends on the particular

mode. The solar p–modes are split into $2l+1$ singlets, indexed by $-l \le m \le +l$. The lowest–frequency p–mode, $p_{0,1}$, has a predicted frequency of $\sim 258\mu$Hz, but very few p–modes have been definitively observed. The frequencies of p–modes vary slightly with solar activity, [43], and this variation is larger at the upper end of the p–mode band than it is at lower frequencies.

Solar buoyancy, or gravity, or g–modes arise from a density inversion at the top of the solar core. They have frequencies $\lesssim 500$ μHz, but the consensus is that they have not been identified in the sense of assigning quantum numbers.

Solar p–modes have been detected in charged particles and the interplanetary magnetic field in space [18], in terrestrial data, [21], and others cited in §7 of [32]. Line frequencies corresponding to predictions of solar g–modes were reported in [18] and in terrestrial seismic data in §**XXI** of [34].

Terrestrial seismic S– and T–modes are well described in several texts, *e.g.*, Chapter 8 of [106] or, for details, [27]. Briefly, spheroidal modes are denoted by ${}_nS_l$, and toroidal modes by ${}_nT_l$. Tables of mode frequencies and Q's are given in [28] and further discussion in [16].

10.2. Spectral Indices of Interplanetary Magnetic Fields

The nature of turbulence and evolution of the solar wind with radial distance from the Sun has been studied since the dawn of the space age. Numerous studies have concluded that, on average, the *spectral index* denoted by G, that is, the slope of the spectrum measured on a log–power *vs* log–frequency setup is approximately $-5/3$, that is the spectrum can be described approximately as

$$S(f) \approx \frac{s_1}{f^{5/3}} \ . \tag{55}$$

In the solar wind this does not hold at either very low or very high frequencies, see Figure 1 of [107]. Strictly speaking one should use wavenumber, k, rather than frequency, f, but in most cases suitable data is not available and it is commonly assumed that wavenumber is proportional to frequency. The index $-5/3$ is generally taken as a signature of Kolmogorov turbulence and its presence in power spectra of solar wind data has been taken as evidence that the solar wind is turbulent. We emphasize that we do not question this observation (*e.g.*, Figure 1 of [21]), but rather its interpretation.

Despite the many papers on turbulence in the solar wind there are problems: *First*, almost all the papers in this field select data intervals conditioned on, for example, the speed of the solar wind [108], visual inspection [109], etc. As a result they violate the conditions stated in the quote by Airy given earlier or, in modern terms, suffer from selection bias. *Second*, Kolmogorov turbulence is a stationary process. However, given evidence for solar wind anisotropy

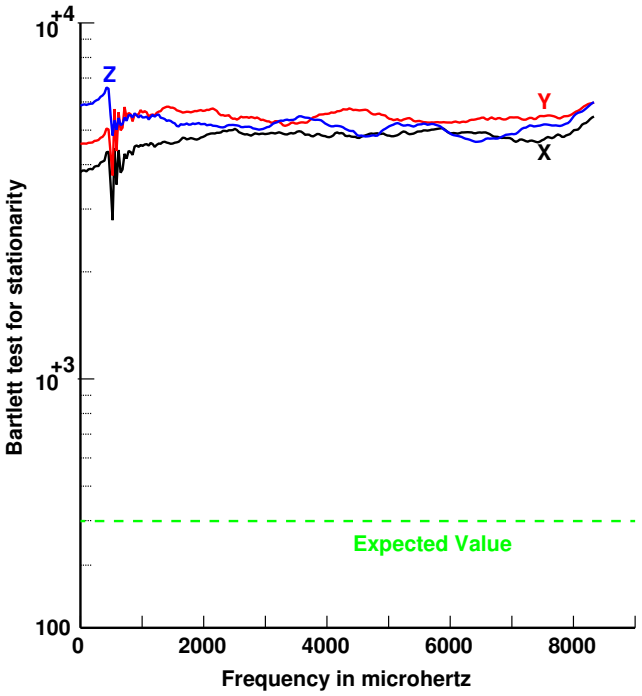

Fig. 5. The Bartlett test for stationarity applied to 400 sections of ACE IMF data. In these tests 3–hour data blocks (180 1–minute samples) offset by six hours were used. The spectrum on each block was computed using $K = 10$ Slepian tapers with a time–bandwidth of $C_{\mathcal{R}} = 5.5$. The test was performed on all three components separately with that for B_X in black, B_Y in red, and B_Z in blue, defined in GSE coordinates. For a stationary process one expects a χ^2_{399} distribution. The test levels, ~ 5000, show that stationarity is not a tenable assumption.

conditioned on the magnetic field [110], plus the implicit recognition of nonstationarity in the selection bias problem, solar wind turbulence must be questioned. Figure 5 shows the results of a simple data analysis experiment where spectra computed on 400 blocks of IMF data from the ACE spacecraft during 1999 were estimated and compared. Each block was three hours long with a three–hour gap between blocks. At an average solar wind speed of 500 km/s these blocks are separated by over 5 million km. Each spectrum was estimated with $C_{\mathcal{R}} = 5.5$ and $K = 10$. The comparison used Bartlett's M–statistic, see Table 32 of [111] or §**IV** of [6]. When the process is stationary, this test has an approximately chi–squared distribution that with 400 data blocks has ~ 399 DoF. For all three components (X, Y, and Z in geocentric solar ecliptic (GSE) coordinates) $M \sim 5000$, so stationarity must be strongly rejected. The break at ~ 250 μHz is near the transition from g– to p–modes.

Thus observed statistical properties *exclude* Kolmogorov

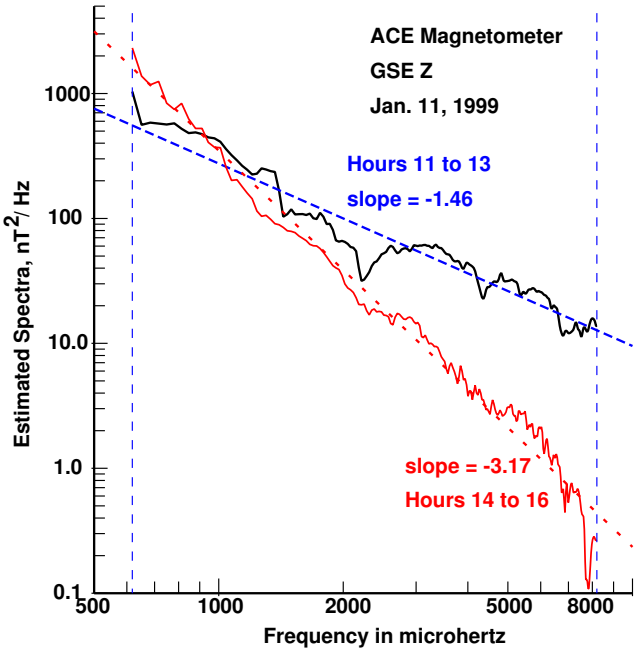

Fig. 6. Two spectra of the GSE B_Z component of the IMF at ACE for hours 11 to 13 (black) and hours 14 to 16 (red) on Jan. 11, 1999. The estimated slopes, shown by dashed lines, are -1.46 and -3.17, respectively.

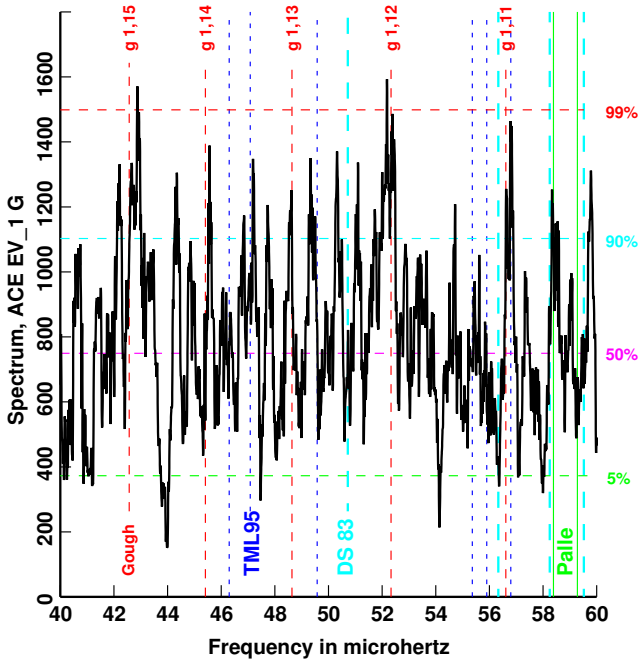

Fig. 7. Spectrum from ACE $\vec{\mathbf{B}}$: $G^{[1]}(t)$. Spectrum Days 1–365, 1999. Resolution: 32 nHz, $C_{\mathcal{R}} = 5$, $K = 8$. Proposed observations of solar g–modes : blue, Thomson & Maclennan & Lanzerotti 1995 [18]; cyan, Delache & Scherrer, 1983 [112]; green, Palle 1991 [113]. Theoretical solar g–mode frequencies (red), D. Gough (personal communication.

turbulence as an adequate explanation for fluctuations in the solar wind. This makes most studies of radial evolution of solar wind turbulence invalid because, if one does not understand the physics causing nonstationarity, one cannot select data sections from a nonstationary process and then pretend that the results have general validity.

10.2.1. Data Analysis Experiments

Following up on this, we tried the following experiment: *First*, take fixed–length blocks of interplanetary data, compute a spectrum on each block, and estimate the slope (in a log–frequency — log–power frame) on each block. Following [109], we began with 3–hour blocks of ACE IMF data. These were offset by one hour and computed every hour for a year. Figure 6 shows multitaper spectra from two adjacent, but non–overlapping, blocks, and it can be seen that the indices differ dramatically.

Second, regard the series of slopes as a new time series and compute its spectrum. This is found to have many low–frequency, high–Q, peaks. The spectra from the 3–hour blocks showed signs of being aliased and the peaks appear to correspond to solar g–modes, similar to the results in Moghtaderi et al. [37]. The frequency range of solar g–modes is limited to $\lessapprox 400\,\mu$Hz corresponding to a sampling rate of 20 minutes.

Third, we used 45 minute data blocks offset by 20 minutes. However, instead of working with a single component of the magnetic field, we took all three components. To partially compensate for an obvious nonstationarity, the magnetic field components from the Ulysses spacecraft were multiplied by its heliographic radius. We computed the eigencoefficients (here with $C_{\mathcal{R}} = 4.5$ and $K = 8$), formed matrices (20), found their singular value decomposition (SVD), and replaced the spectrum with the eigenvalues when estimating the index. We refer to the spectral index of the largest eigenvalue as $G^{[1]}(t)$.

Fourth, Figure 7 shows a power spectrum of $G^{[1]}(t)$ for 1999. The spectrum has been postwhitened and the peaks show reasonable agreement with some possible measurements of solar g–modes [112, 113, 18] and theory[4].

Fifth, the next comparison was between ACE and Ulysses. There was an aproximate radial alignment of ACE and Ulysses in 1999. Sixty days of data were used, giving a low-resolution spectrum. However, as may be seen in Figure 8 there is remarkable agreement between the two spectra. Scaling the Ulysses data by radius (≈ 5.08 to 4.99 AU during this interval) does not change the slope. There is a reasonably sharp small cutoff, $\sim$ 290 – 300 μHz in the

[4]D. Gough, personal communication

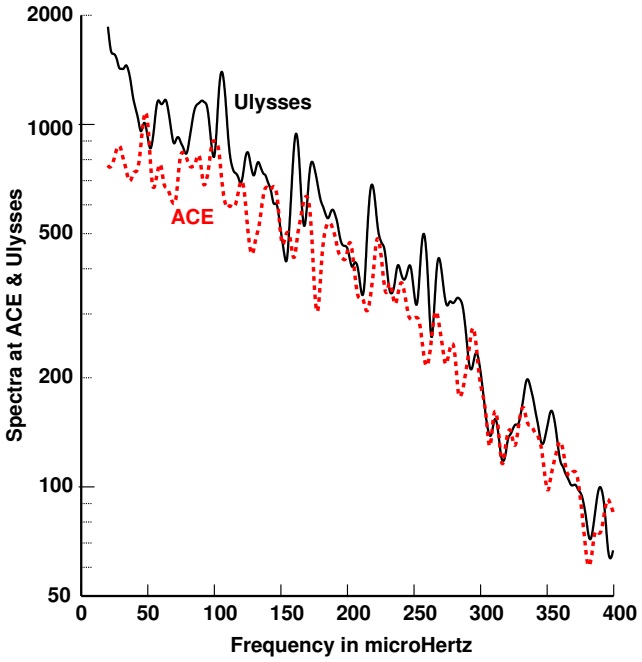

Fig. 8. Spectral index data ACE: DOY 42–102, 1999, Uly: DOY 60–110, 1999, (4320 samples, $\delta t = 20$ min.) $C_{\mathcal{R}} = 5$, $K = 8$ windows Note that there are *no* corrections for the different orbits. There are only minor changes in the spectra. Solar p–modes have a frequency range ~ 250 μHz to $\gtrsim 5000$ μHz. Solar g–modes theoretically have frequecies $\lesssim 500$ μHz.

spectrum near the edge of the g–mode band, but closer to the transition between g–modes and p–modes. Given the low resolution, the "index spectra" at ACE and Ulysses are nearly identical. This implies that, with 1-minute data and between 0.99 and 5.2 AU, the statistics of the slopes do not evolve detectably. Taking a longer 299 day block between 2003 day 271 and 2004 day 205 when Ulysses was near aphelion and between heliographic latitudes of $\pm 7.6°$ (chosen to have the same range as Earth and ACE), the average index was -2.218 at ACE and -2.203 at Ulysses. The minimum index among the 27,000 samples was below -5.1 for both spacecraft while the standard deviation increased from 0.583 for ACE to 0.775 for Ulysses.

How does one explain these observations? The histograms of the indices appear similar at both spacecrafts with only minor evidence for evolution between 0.99 and 5 AU, and the difference may depend on their necessarily different distribution of heliographic latitudes sampled. The distribution at Ulysses is nearly uniform, but U–shaped at ACE. However, spectra of the index series had a larger modal component at Ulysses than that observed at ACE.

A possible sequence of events to explain the observations is to begin in the turbulent solar convection zone. It has high densities, high temperatures, is completely ionized, has strong magnetic fields, supersonic flows, high conductivities, low damping, and generally everything that is required to generate Kolmogorov turbulence. The convection zone turbulence drives both p– and g–modes. This is a standard assumption and agrees with theory by [114] and the observation that signal–to–noise ratios of modes in the IMF is approximately constant across frequency. Proceeding outwards, both turbulence and modes drive surface phenomena and modulate the solar wind by the same mechanism.

Magnetic field lines and particles flow radially outward from the Sun as originally proposed in [115, 116]. A combination of horizontal flows and g–modes (whose surface motion is mostly horizontal) moves the footprint of the flux tubes in an apparently random manner. The seismic surface of the Sun is ~ 300 km below the photospheric surface [117], so most interface phemonena are not directly observable. However, our hypothesis is that the $f^{-5/3}$ baseline observed in spectra of interplanetary data is a fossil remnant of the convection–zone turbulence.

On longer time scales, one sees the average position of flux tubes and analyzing long series shows solar modes as the dominant phenomena. Most, $\gtrsim 99\%$ of the energy in the IMf occurs in the g–mode band, that is, at frequencies < 500 μHz, so given the observation that magnetic fields control anisotropy in the solar wind [110], evolution of the spectral index may depend on the magnetic signatures of g–modes. Given the relatively stronger modal characteristics of data at Ulysses compared to those on ACE (but lower overall power), one wonders if the presence of modes does not, in fact, supress turbulence.

Also note that these ideas should be considered preliminary and that they push both ends of the spectrum estimation problem. The short blocks cause some bias on the index estimates, whereas taking a year or so of these index estimates runs into many of the same resolution problems as other helioseismology problems, so improvements are needed in both.

10.3. Geomagnetism and Robustness

Fluctuations in Earth's geomagnetic field induces currents on electrical cables, pipelines, and similar structures and can cause overheating in power system transformers, corrosion on pipelines, and power supply problems in communications systems, *e.g.*, see [118, 119]. The problem is not new, and was first documented by Barlow in 1849, [120].

A persistent problem in attempting to analyze geomagnetically induced currents (GIC) and in the closely related field of magnetotellurics [121] is that the effective signal–to–noise ratio is inexplicably poor, outliers are common, and large amounts of data are often needed. Estimating the induced voltage on a cable from fluctuations in the geomagnetic field requires transfer functions. Estimating these

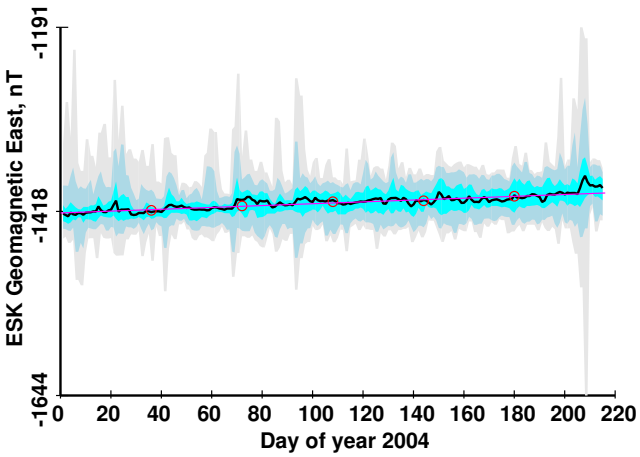

Fig. 9. Summary plot for a relatively short section, 215 days, of the East component of the geomagnetic field at Eskdalemuir. Here $\delta t = 240$ s, and summary statistics are computed on two–day blocks (720 samples), offset 50%. The bands shows minimum–to–maximum, light gray; 5% to 95%, blue–gray; and 25% to 75%, cyan, with the median in black. The red line is a linear trend.

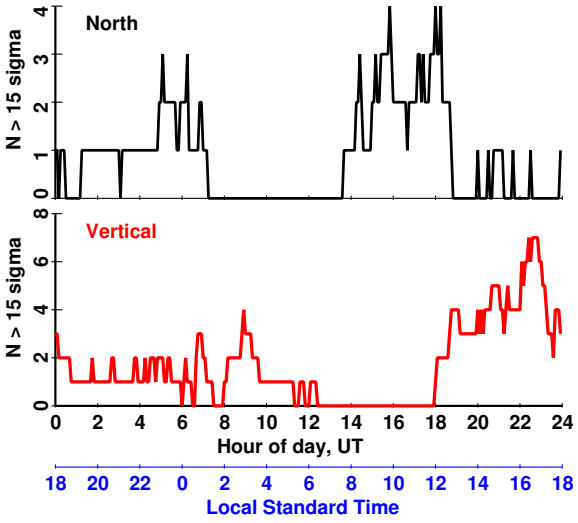

Fig. 10. Time–of–day dependency for extremes larger than 15 σ for the North (top, black) and Vertical (lower red) in Boulder geomagnetic data.

transfer functions [121] is often unsatisfactory because the estimates are excessively variable across frequency and the residuals from the frequency–domain regressions are often much larger than seems reasonable. The variation across frequency was pronounced on Transatlantic cables, [122], and a cepstrum, [58], gave a delay of 110 minutes. The source of this apparent delay is unknown. The high residuals appearing in transfer function estimates are the result of coherences between magnetic fields and GICs from cables being "improper". This means that if $E_k(f)$ denotes the eigencoefficients of the induced electric field at some frequency f and $B_k(f)$, the corresponding ones for a component of the magnetic field, the "improper" (or reverse) cross–spectrum $\sum_k E_k(f)\, B_k(f)$ is often larger in magnitude than the "proper" (or forward) cross–spectrum $\sum_k E_k^*(f)\, B_k(f)$ where the superscript $*$ denotes complex conjugate. See the discussion in [123]. Moreover, the improper terms are often distributed across the frequency spectrum and so form a "noise source" whose power is similar or larger to that in the proper components. Again, however, the source of the improper data is not well understood and they were only recognized recently, [17].

In addition to these effects, time series of geomagnetic or interplanetary magnetic fields commonly have periods of quiescence interspersed with intervals of high activity and it has been common practice to classify these into "quiet" and "active" days since at least the 1860s. Figure 9 shows a plot of moving quantiles from the East component of the geomagnetic field at Eskdalemuir, Scotland and the effect is obvious. However, it has been recognized for almost as long that this traditional classification, while convenient for many purposes, is not scientific and, in 1863 [124] G.B. Airy, the Astronomer Royal, stressed

> "First, that there is no such thing as a day really free from disturbance, and no reason in the nature of things for separating one or more days from the general series. There is abundant reason for such separation on the grounds of convenience $\cdots$. Under these circumstances, I cannot think it right that I should cut off a part of that salience, with the belief of obtaining results that can possess any philosophical value, from the part which is left."

While recognizing the propensity for outliers to occur in many kinds of data, most of the robust procedures do exactly what Airy is warning against and "cut off part of that salience". However, one must also recognize that much data is not Gaussian and that large "saliences" in data do not imply that the data has been contaminated or has incorrect samples. A few extreme cases of large impulsive transient events such as lightning strikes and the like are real physical signals. On the other hand, one of the largest effects on an estimated spectrum from an outlier in our experience is shown in Figures 12 and 18 of [6]. This was a true error, a repeated data sample, almost invisible on an ordinary data plot, but it did produce an error $\gtrsim 10^7$ across much of the frequency range. It was also *not* an extreme, so one must be very careful not to confuse extreme values with contamination. As a simple experiment we took data recorded by the Boulder, CO geomagnetic observatory from Jan. 30 2002 to Mar 15 2004, lowpass filtered and decimated to $\delta t = 5$ min.

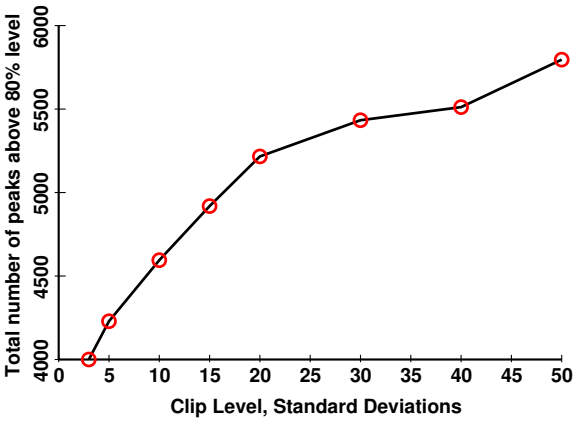

Fig. 11. Number of peaks in the harmonic F–test in the three components of Boulder, CO geomagnetic data as a function of the time–domain clip level.

We stress that there is nothing unusual about the Boulder data, indeed it appears to be a very well run observatory. Different time intervals and data from other well run observatories give similar results, see Figure 9, and our impression is that there are no indications of instrument malfunction or data contamination and the extremes are an accurate portrayal of the state of the geomagnetic field.

Given this data, we computed a robust (Winsorized) scale with the lower and upper 10% of the values trimmed. As an example, on the X (North) component the ordinary standard deviation is $\sqrt{\mu_2} = 25.165$ nT, the kurtosis, $\beta_2 = \mu_4/\mu_2^2 \approx 54.5$, and the robust scale $\hat{\sigma}_w \approx 20.846$ nT. Figure 10 shows the time–of–day dependency for extremes greater than $\pm 15\sigma$ for the North and Vertical components as a function of time–of–day over an extended interval, Jan. 1999 to May 2009. There are significant differences in patterns but both imply strongly cyclostationary behaviour. Extremes on the Eastward, or Y, component are both fewer and closer to having a uniform distribution in time–of–day and are omitted.

We then varied the clip point, c, from 3 σ to 50 σ, replacing points outside $\pm c\hat{\sigma}_w$ with a linear interpolation. Next, we computed spectra and F–tests on ten 100–day blocks with a 25–day overlap. A time–bandwidth product $C_{\mathcal{R}} = 6$ with $K = 10$ tapers were used. Finally, we computed the number of peaks in the harmonic F–test, $F(f)$ with the results shown in Fig. 11. The point of this plot is that simple clipping peaks reduces the number of potential periodic components from nearly 6000 with no clipping to just under 4000 with clipping at 3σ.

At high significance levels there are many more peaks in these spectra than one can reasonably account for by chance [36]. Randomly permuting the order of the data, see Fig. 6 and Table 2 of [36], gives crossing rates that agree with Gaussian theory. Because the original and permuted data have identical first–order probability distributions, we conclude that the large number of peaks observed depends on the temporal order of the data and not its amplitude distribution. Similar effects have been observed in charged particles and magnetic fields in interplanetary space [18, 19]; in data from the ACE spacecraft at the first Lagrange point, L_1 [125, 36]; in geomagnetic fields recorded at other observatories; and in the solid earth [16]. The only explanation that we have found for the number, frequencies, and Qs of these peaks is that they originate as normal modes of the Sun. Observed on Earth, these modes are further split by Earth's rotation and in all cases must be corrected for the orbit of the observing instrument (including the Earth) around the Sun.

The fact that *more* peaks are found as the clipping level is increased has another major implication; namely that the peaks in the time–domain data are a result of the modes, and are not just "outliers". That is to say many of the large peaks seen in such data occurs when several modes become phase aligned. This is analogous to a "Dirac comb" where a sequence of periodically spaced impulses in the time–domain has a Fourier transform that is also a periodic series of impulses in the frequency–domain. Low frequency solar modes are not equally spaced in frequency so this is not an exact analogy, but extremes are perhaps best described informally as "several modes becoming phase aligned". Similar effects are common in wideband amplifiers where phase alignment of the various subchannels causes the amplifier to saturate. The Sun does not appear to have a similar power limitation.

10.4. Seismic Spectra

The low frequency seismic noise spectrum, particularly of the horizontal components, is enigmatic. Here we discuss two contributing factors to these problems. The first is statistical, specifically, bias and other problems associated with the procedures commonly used as discussed in section §9.1. The second is physical and it appears that much of the problem in understanding horizontal seismic data is that the solar contribution is relatively strong at low frequencies. In [126] it is noted that the seismic hum on the horizontal components is *not* explained by ocean forcing. Solar modal signals are often larger than the median and often are the strongest signals present. Given that the Sun is the source of these signals, one must consider their characteristics, and on Page 183 of [105] Stix notes

> "The general appearance is that "one–third oscillates" — either of a given area at any one time, or of a given period of time at any one location on the solar surface; but this pattern changes continuously."

While obviously not a precise mathematical description of

any particular mode, such observational insight should not be ignored. Also remember that this is based on ordinary solar p–modes in the 2 to 5 mHz band and the low–frequency modes may be different. Signals that are present $\sim\frac{1}{3}$ of the time will usually be missed by a procedure that expects them to be present $\sim\frac{1}{2}$ of the time.

10.4.1. Black Forest Observatory Example

The problem with seismic data is that all of the properties predicted in section §9.1 are violated. To illustrate this, we used the North velocity from the STS–1 instrument at Black Forest Observatory (BFO) between Jan. 10 and July 4, 2004. This was globally a seismic quiet period with only two earthquakes with $M \geq 7$ and 49 with $M \geq 6$ distributed over the 157 day interval. The horizontal components of the STS–1 are known to be insensitive to magnetic fields, [127], so the solar peaks in the spectra cannot be from geomagnetic coupling to the instrument. For seismic normal mode frequencies we use [28] augmented with singlet frequencies provided by T. G. Masters. (The $2l+1$ singlets of low frequency seismic modes can occupy a band nearly 20 μHz wide, see Table 14.1 of [27].) For solar p–mode theory we used [50] and, for the few measured solar p–modes below 1 mHz, [128]. These are denoted $p_{l,n}$

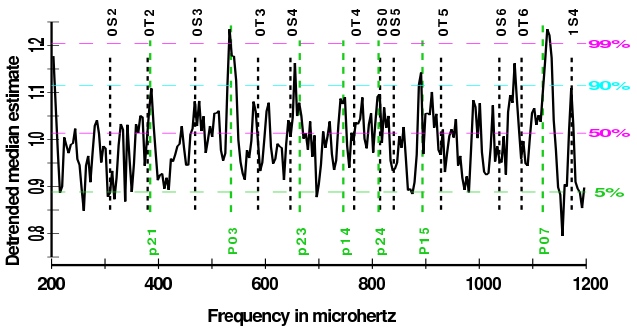

Fig. 12. Detrended median spectrum of 350 sections of BFO-North using 24-hour blocks. Horizontal dashed lines show 5%, 50%, 90%, 99% significance levels based on the chi–squared approximation. Green vertical dashed lines show observed solar p–mode frequencies. Black vertical dashed lines show observed terrestrial S and T–mode frequencies. Only two peaks exceed the 99% significance level, at 533 and 1127 μHz, both within one–quarter of a Rayleigh resolution ($\mathcal{R}/4$) of the measured p–modes $P_{0,3}$ and $P_{0,7}$, respectively.

The data was prewhitened and our initial experiment was with 350 24–hour blocks. Each section had its average and linear trend subtracted and a single $NW = 2$ Slepian taper applied, see section V.B of [34]. We then fit a quadratic to $\ln \hat{S}_{(50\%)}(f)$ over the band 150 to 1250 μHz

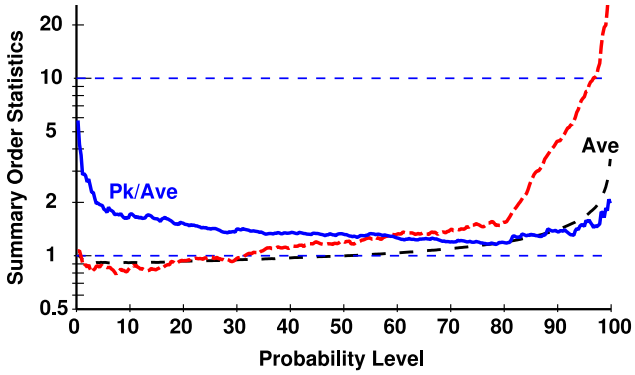

Fig. 13. Observed statistics of 350 24–hour spectra of BFO–N seismic velocity. Black dashed line shows the average order statistics. Red dashed line shows the variance divided by the expected variance. Blue solid line shows the ratio of the maximum to the average spectrum. Blue dashed horizontal lines show the 1% and 10% order statistics.

and removed the fit from all $J = 350$ order statistics. The median spectrum, Figure 12, is *less* variable (across frequency) than expected and much less so than the spectra made from either the 20% or 80% levels. Looking at statistics across the 200 to 1200 μHz band, Figure 13 shows that the average increases gradually with quantile up to about the 90% level after which it increases rapidly. Similarly, the variance divided by the expected variance (53) drops slightly initially, and again increases rapidly after about the 80% point. Finally, the ratio of the maximum to average spectrum shows a roughly U–shaped blue curve as a function of frequency. The behaviour at the upper end of these curves is *not* explained by the few $M \gtrsim 6.5$ earthquakes and both the low end and the behaviour in the 80 to 90% range is enigmatic. Exploratory studies suggest that the seismic noise has several components: the lowest quantiles, < 10 or 20%, may just be thermal noise (but see below); between 20 and $\sim$ 50% there is reasonable correlation with the planetary K_p index; above that there are earthquakes and discrete solar modes. That is, the data are stratified. Moreover, the seasonal dependence of the three lower components appears to vary in different ways, but this needs considerable further study. In particular, the median spectrum, Figure 12, has several moderate peaks that superficially appear to be at the frequencies of the Earth's normal S and T modes, but on closer examination, are not.

We estimated robust autocorrelations between the spectrum estimates on consecutive blocks for each frequency. Significant correlations were observed out to delays in excess of 35 days and we examined the largest ones carefully. Of these only one, the 24^{th} largest, at $f = 1037.60\,\mu$Hz and $\tau = 7$ days, was close to a seismic normal mode, ${}_0S_6$ at 1037.55 μHz. This is also close to the measured frequency

f_{th}	mode	f_{obs}	Sid	doy
246.4	$P_{0,1}-1$	245.7	90	133
269.5	$P_{0,1}+1$	270.1	max	33
300.0	${}_0S_{2,-2}$	300.1	90	22
391.3	$P_{0,2}-1$	391.1	max	33
464.3	${}_0S_{3,-2}$	463.4	90	20
470.9	${}_0S_{3,+1}$	470.9	90	31
583.3	${}_0T_{3,-3}$	583.4	80	*39
587.6	${}_0T_{3,0}$	587.5	90	183
644.8	${}_0S_{4,-2}$	644.4	max	*39
651.3	${}_0S_{4,+4}$	651.6	80	137
674.1	${}_1S_{2,-2}$	674.4	80	120
677.7	${}_1S_{2,-1}$	677.0	50	91
768.1	${}_0T_{4,+1}$	768.0	max	14
812.8	${}_0S_0$	812.3	80	*116
926.2	${}_0T_{5,-5}$	926.2	80	110
930.1	${}_0T_{5,+5}$	930.3	80	106
930.9	${}_0T_{5,+2}$	931.3	90	110
1034.5	${}_0S_{6,-6}$	1034.5	max	16
1079.1	${}_0T_{6,-4}$	1079.3	80	183
1130.2	$P_{0,7}+1$	1130.2	50	*116

Table 1. Matches within $\pm\frac{1}{2}\mathcal{R}$ between peaks in spectra and known modes. For solar P–modes the ±1 denotes a ±1 cycle/day offset. On seismic S and T modes the third index is m. The columns marked f_{th} and f_{obs} are the known and observed peak frequencies, "Sid" the spectrum quantile, and "doy" the day in 2004 where the maximum occurred. Those marked with a "*" have earthquakes with moment $> 10^{26}$ dyne–cm within the segment.

of the solar $P_{1,6}$ mode at $1039.65\,\mu$Hz. Solar p–modes are expected to be split into $2l+1$ singlets spaced $\sim 0.45\,\mu$Hz apart, so these frequency comparisons *cannot* be better than a few μHz.) Averaged over delays, the highest correlation occurred at $723.27\,\mu$Hz, $0.4\,\mathcal{R}$ from the predicted frequency of $p_{3,3}$ at $718.39\,\mu$Hz. Twenty of the largest 100 autocorrelations occur within $\pm\frac{1}{2}\,\mathcal{R}$ of this mode at delays from $\tau = 1.5$ to 21 days. Excluding ${}_0S_0$ (at $812.838\,\mu$Hz), seismic modes damp much faster than this. To improve frequency resolution, we computed the spectra on 83 sections each 100 hours long, offset by 50 hours. This was zero padded to 32k giving a frequency increment of $0.254\,\mu$Hz. Unbiased quantile estimates for the 20, 50, 80, 90, and 100% (the maximum) levels were computed and a quadratic fit made to $\ln\{U_{50\%}(f)\}$ over the band 150–1250 μHz. This fit was removed from all estimates. Figure 14 shows the flattened spectra over 800–1000 μHz. Several things are apparent: *First*; the difference in character between the four estimates implies that, as with the one–day blocks, the estimates are *not* stable. The median to 90% curves have over 80 DoF, Table 2, so their standard deviations are about 0.16

j	Mean $\mathcal{M}_{83}(j)$	Gain $\mathcal{G}_j$	Variance $\mathcal{V}_{83}(j)$	DoF	Prob. %	Notes
1	0.0120	83.0000	1.0000	2.0000	1.20	Min
17	0.2276	4.3929	0.0591	33.8541	20.36	20%
42	0.6991	1.4303	0.0248	80.6629	50.30	Med.
67	1.6213	0.6168	0.0185	108.1528	80.24	80%
75	2.2842	0.4378	0.0202	98.8784	89.81	90%
83	5.0021	0.1999	0.0653	30.6446	99.33	Max.

Table 2. Theoretical quantities associated with the $J = 83$ section spectra. The 20%, 80%, 90% , and Max entries are shown in Figure 14. In 83 spectrum estimates the largest (bottom line) is expected to be 5.0021 times the mean and so is scaled by 0.1999. In stationary Gaussian data this estimate will have a variance of 0.0653 times the mean spectrum squared, corresponding to 30.6446 degrees–of–freedom (DoF).

and the differences between the curves is much larger than this. *Second*; comparing the median curve with Figure 12, one goes from 2 peaks just over the 99% level to many above the 99.9% level. *Third*; the peak frequencies vary between levels. There were 9 cases where 2 of the 5 estimates were on the same FFT bin, but only 2 cases with three such agreements. *Fourth*; it appears that several of the peaks agree with seismic mode frequencies. Some of these agreements are summarized in Table 1 for those cases where **A:** the peaks exceed the 99.9% significance level and **B:** the observed peak agrees within $\pm\frac{1}{2}\mathcal{R}$ ($\pm1.39\,\mu$Hz) of an identified solar or seismic mode. We also allowed splittings of ±1 cycle/day on solar modes, [34].

Of the 20 peaks identified in Table 1, 15 are seismic. The problem is that only three of these occur when earthquakes with moment $> 10^{26}$ dyne–cm occur within or closely before the data block. Moreover, if one looks at the *maximum* spectrum from the 83 blocks, many of the solar peaks are larger than nearby peaks at seismic mode frequencies.

Figure 15 shows standardized quantiles $\widehat{U}_{[j]}(f)$ for a solar and a seismic mode that occur reasonably close in frequency, are both clearly visible in Figure 14, but have quite different characteristics. The seismic ${}_0T_{5,0}$ mode (the last subscript indicating the $m = 0$ singlet) is centered at $f = 930.551\,\mu$Hz and the spectrum is above the 99.99% significance level on the 90% quantile but is relatively small on the maximum. However, see Table 1, none of the spectral peaks in its frequency range appear to be closely associated with known earthquakes. This mode has 11 singlets distributed somewhat unevenly between 926.239 to $930.944\,\mu$Hz, a range that appears to include the solar $P_{9,3}$ mode at $930.540\,\mu$Hz, [129], so the 19 singlets of this mode will cover a range $\sim$ 926.9 to $\sim$ 934.1 μHz. Thus the

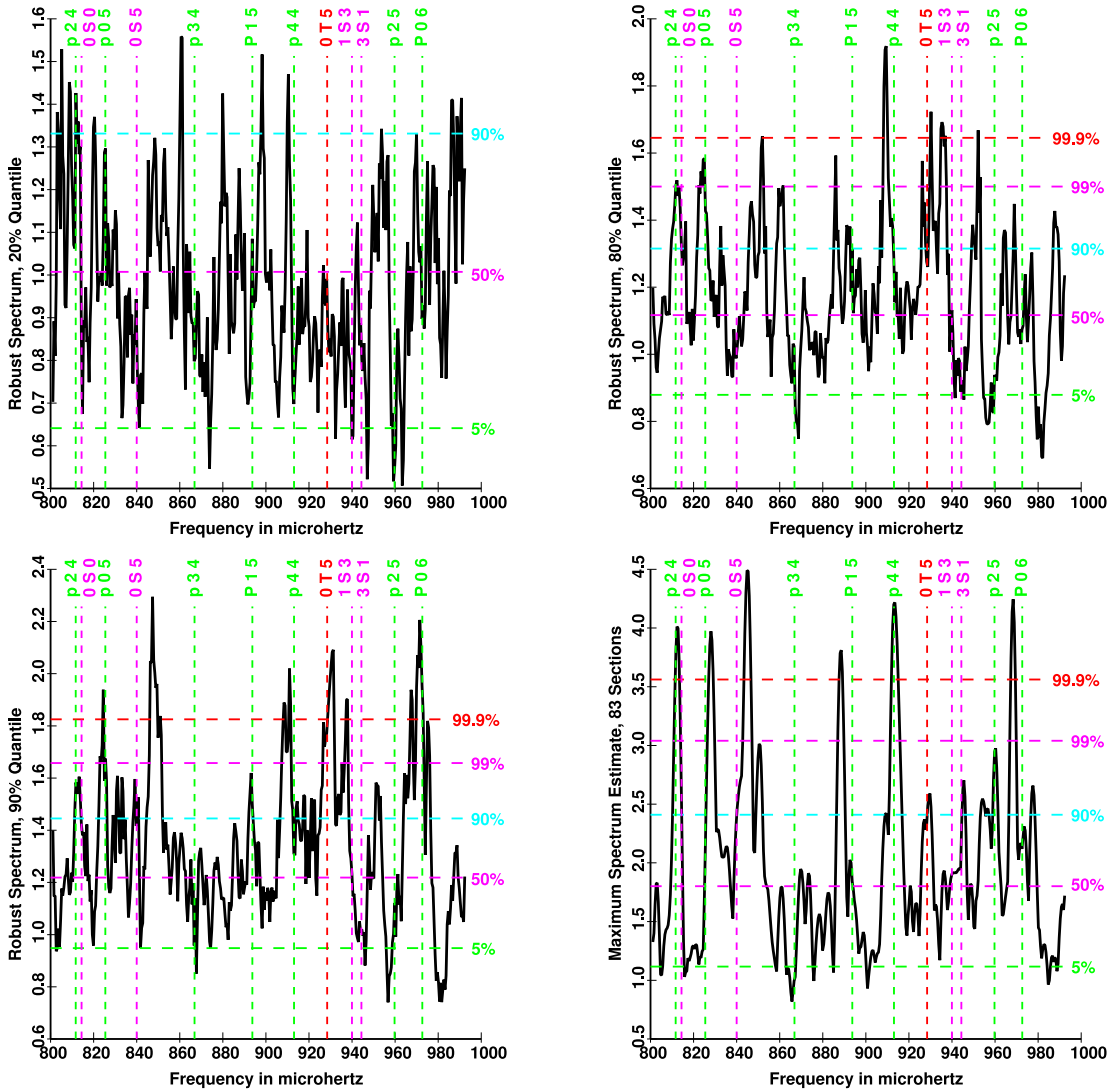

Fig. 14. Four unbiased (in the sense of (52)) and detrended spectrum estimates from 83 100–hour blocks of the BFO North data, corresponding to the 20%, 80%, 90%, and 100%, or Maximum. The horizontal dashed lines show significance levels based on the chi–squared approximation. The vertical dashed lines indicate the frequencies of solar p–modes in green, and seismic S and T modes in magenta and red respectively. Note that many of the largest peaks in the 90 and 100% estimates agree with solar, not seismic, normal modes. Many of the other peaks agree with solar mode frequencies offset by ± 1 c/d, or with higher–l modes. As shown in Figure 13, the average level creeps up with quantile. Also, the peak frequency varies with quantile, implying that these estimates do not give reliable frequency estimates.

frequency range covered by the singlets of $P_{9,3}$ overlap those of ${}_0T_5$. The solar $P_{0,6}$ mode at $f = 972.61\,\mu$Hz [130], in contrast, is largest on the maximum. The closest seismic modes are ${}_2S_{2,2}$ at $955.542\,\mu$Hz and ${}_0S_{6,-6}$ at $1014.526\,\mu$Hz well outside the bandwidth of the spectrum estimation procedure. Returning to Figure 15, one observes that for probabilities in the 15 to 55% range behaviour for both modes is similar and biased slightly low, probably indicating a need to use a more robust fitting procedure to determine the overall baseline. Below the median, ${}_0T_5$ behaves about as expected, and the general increase in level above the median may be due to the numerous $M \gtrsim 6$ earthquakes and occasional coupling to solar modes. The lower peak in the $P_{0,6}$ curve suggests that there is *always* some signal present. The sharp increase beginning around 80% appears to be a reflection of the "one–third oscillates" phenomena modulated by the Earth's orbit and the general seasonal dependence of geomagnetic activity.

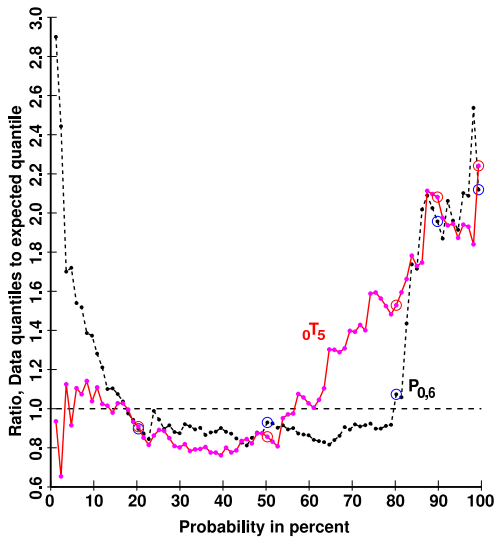

Fig. 15. Standardized quantiles, $\hat{S}_{(j)}(f)\mathcal{G}_J(j)$ *vs* P_j for the solar $P_{0,6}$ mode at $f = 972.61\,\mu$Hz (black) and the seismic ${}_0T_5$ mode at $f = 930.55\,\mu$Hz (red). These modes are reasonably close in frequency and both are prominent in Figure 14,

10.4.2. Discussion: Median Spectrum

We have shown that medians of power spectrum estimates done on short blocks are both biased and inefficient. Correcting for the bias one finds that the average varies with the quantile. Moreover, at many frequencies, the spectrum estimates are highly correlated over durations of twenty days and more. Worse, the procedure appears to miss most of the significant low–frequency features in the data. Taking 100–hour data segments we found many peaks significant above the 99.9% level in the BFO North seismic velocity. The frequencies of these peaks correspond closely to those of *solar* normal modes and not those of the normal modes of the Earth. In several cases these peaks are higher than nearby peaks at seismic mode frequencies. These spectral peaks are narrow, implying that they are being driven by a high–Q (> 1000) source. Their detectibility varies with quantile, suggesting that the data consists of a mixture of processes with approximately stratified amplitudes. This implies that this analysis method should be used cautiously at low frequencies.

Looking again at Figure 14, the differences between the estimates are almost as large as in Figure 16 of Kay and Marple's 1981 comparison of spectrum estimates [131]. Here, even though quantile estimates are "Min–Max", [132], the differences seen between the different quantile estimates are striking. Although the median estimate was chosen to exclude the effects of large earthquakes and the data interval contains a $M > 7$ earthquake, the largest peaks in the spectrum appear to be solar. Thus we stress that, when computing spectra from large quanties of data, use more than one estimate, understand the properties of the different estimation procedures, and plan for the unexpected.

10.5. Daily Temperature

Recently, several excellent long series of daily European weather records, [133], have become available and, to make the examples easier, we analyze some of them here. The series are summarized in Table 3. All of these series were preprocessed in the manner discussed at the end of §8.1, and the harmonic T^2 and F tests were computed.

City	Start	End	JD_s	N
Padua	1774	1996	2369001.0	81449
Stockholm	1756	1998	2362426.0	88754
Uppsala	1722	1998	2350019.0	101161

Table 3. Long daily European temperature series. The first and last years are given plus the initial Julian Ephemeris day number (JD_s) and the number of samples (N).

The Uppsala series of daily temperatures [134] has the distinction of being the longest one that exists. It begins on Monday, January 12, 1722 and continues to the present. In total, it is just over 100,000 samples so easily handled as a single block. Such data is cyclostationary with a period of one year. In the Uppsala data the period appears to be indistinguishable from the Gregorian calendar year, but again with long data spans this must be tested, not assumed, see [89, 135, 136].

We begin by removing the annual temperature cycle. The variance, Fig. 16 was initially computed on a 1461–day "year" to account for leap years, then averaged down for plotting. There are two estimates, the ordinary variance estimate and a robust [137] estimate. There is little difference, confirming the impression that the data quality is excellent. There is, however, a change of about 6:1 in variance over the year and, strangely, a "discontinuity" near the vernal equinox. This implies that the data will be strongly cyclostationary (or periodocally correlated). Such processes are discussed in [138, 100] and in the recent book [139].

If one now computes the dual-frequency magnitude–squared–coherence (MSC) [99, 139] between frequencies f and $f + 1$ cycle/year one finds 8735 local maxima above the 99% significance level. For this reason it does not appear to be possible to scale the data to remove the coherence between frequencies. Overall, about 25.8% of the estimates are above the nominal 99% level showing highly significant coherence. Randomly permuting the temporal order of the residuals and repeating the operation results in 533 local

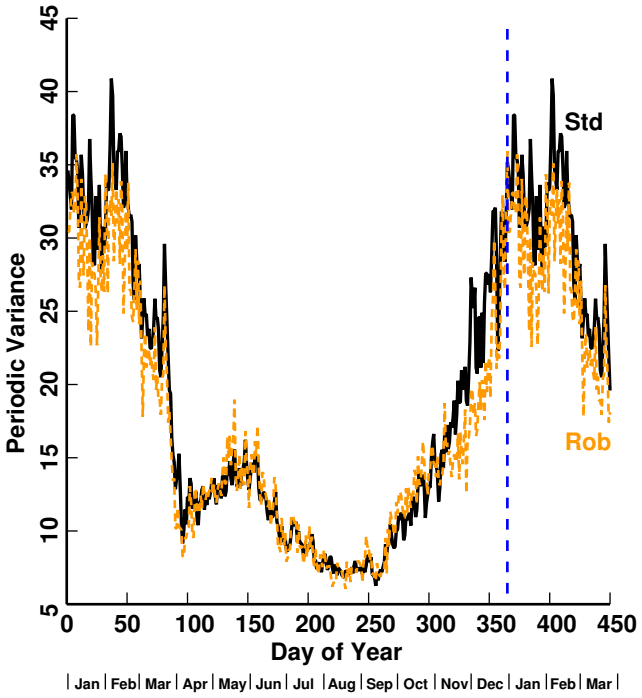

Fig. 16. Periodic Variance for the Uppsala daily temperature data. The solid black curve is ordinary variance, the dashed orange a robust estimate. Note the "discontinuity" near day 80, the vernal equinox.

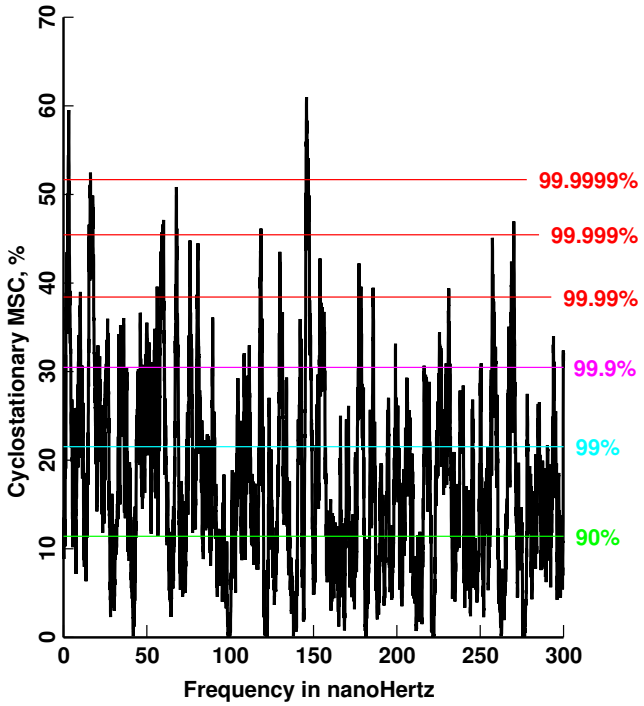

Fig. 17. Cyclostationary MSC between the Uppsala data at frequency f and $f + 1$ c/y. Horizontal colored lines delineate significance levels based on the chi–squared approximation.

maxima with 1.04% of the frequency range above the 99% level, close to expected. Figure 17 plots the lowest 300 nHz of this coherence (about 20 pages of such plots would be required to cover the Nyquist bandwidth) and some remarks are in order: *First*, the large peak near the center is the first harmonic to the Rieger periodicity (154–day, ~ 75.15 nHz), found in 0.3–100MeV solar gamma rays, [140]. *Second*, the plot demonstrates a common problem that one finds in most of these long data sets, namely that the MSC estimates are almost bimodal, high at modal frequencies, and low at noise levels between peaks. This may be seen even more clearly in the following plot where the frequency range is 3 nHz instead of 300.

Because at least some of these peaks in the real data have the same frequencies as measured in interplanetary space [18], it can be assumed that there may be a few thousand low–level solar modes contributing to the temperature variance. One assumes that, in the daily temperature data, most of these are aliases of g–modes. These modes are split by Earth's orbit and so can be detected at f and $f + 1$ cycles/year (c/y). We have begun a study of this general topic in [34], but much work remains to be done. Suitable data is usually not available and analysis methods are being developed simultaneously with applications. One problem is that of time scales, and in examples such as the dropped–call rate in cellular phone systems operating near 900 MHz [141] much of the problem is solar. Engineers are used to acquiring data, but when the fundamental periods causing the problems range from hours to decades, data spans of years may be required. This causes problems because changes in systems occur too frequently to ensure a homogeneous data stream.

In both cases the problem is that one has many low–amplitude periodic components at (usually) unknown frequencies and these will result in apparent coherent behaviour whenever f_1 and f_2 correspond to line frequencies. A Loève spectrum [54] of the Uppsala data may well contain $\sim (8735)^2$ peaks above the 99% level. This is misleading because periodic components should be considered part of the mean–value function, not the second–moments, and mixing the two leads to confusion. However, the 25.8% of the MSC's above the 99% significance level implies that either the continuum spectrum is also significantly coherent or that the spectrum is almost "all lines". The line problem can be partly solved with the complex T^2 test discussed in §8.3. A possible solution to the continuum problem is discussed in [39]. Figure 18 shows a harmonic T^2 test for common periodic terms in daily temperature data at Stockholm [142] and Padua [143]. The data begins in 1774 and ends in 1993, a span of 220 years. (The annual cycle from both series was

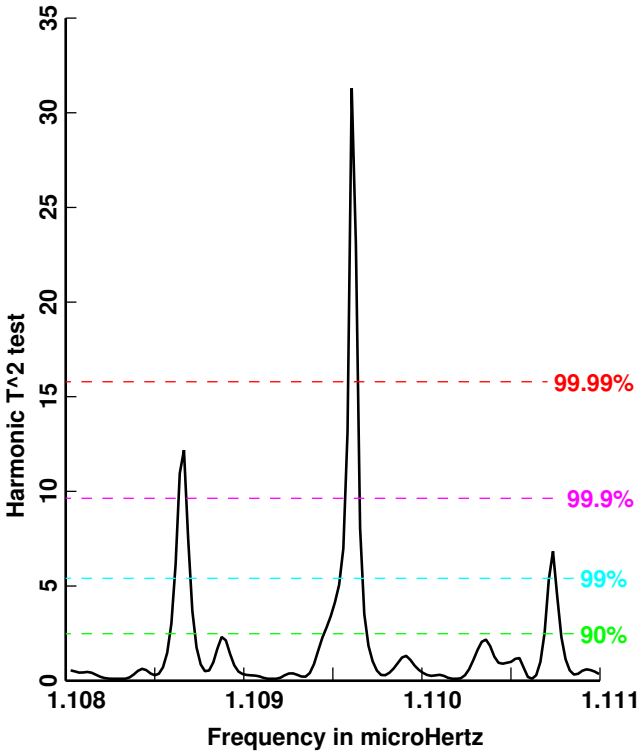

Fig. 18. Harmonic T^2 test for common periodicities in the daily temperature series at Stockholm and Padua. The record spans 1774 to 1993 giving a Rayleigh resolution of 114 pHz. In this test, the frequency was *prespecified* from Ulysses spacecraft data. The estimated common frequency is $\hat{f} \approx 1.10951\,\mu$Hz, corresponding to a synodic period of ≈ 10.4317 days.

removed before further processing and a few outliers in the Padua series from the early 1940s were interpolated.) Here a relatively high time–bandwidth product $NW = 11$ with $K = 20$ tapers was used, so spectra have ~ 40 DoF. The plotted frequency range, 1.108 to 1.111 μHz, with a total bandwidth of 3 nHz was chosen to include the 1.11 μHz line detected in interplanetary electrons by the HISCALE detector [144] on the Ulysses spacecraft [18]. This peak exceeds the 99.9999% significance level and has an extimated synodic frequency of $\approx 1.10951\,\mu$Hz. Even with $\sim 100,000$ samples, a peak at this significance level would be unusual. This agreement between sidereal and synodic frequencies implies that the mode has azimuthal order of $m = 0$. The detection of this peak in both interplanetary data and terrestrial temperature appears to answer the question posed in [145]. There are 8735 peaks where the cyclostationary MSC at f and $f + 1$ c/y exceed the 99% level, and a large fraction of these also have high values of T^2.

Again, the lessons contained in this short section is that climate data is complicated, and intimately linked to the

	Randomly Permuted		Original Data	
Prob., %	%	Peaks	%	Peaks
90.	9.641	4454	59.859	18673
99.	1.039	533	25.787	8735
99.9	0.157	95	8.918	3270
99.99	0.012	11	2.600	1016
99.999	0.001	1	0.700	314
99.9999	0.000	0	0.115	52

Table 4. Statistics of cyclostationary Magnitude–Squared–Coherence between Uppsala daily temperature at f and $f + 1$ cycle/year. Left column is significance; (Fraction of frequencies above level and number of local maxima above level) in randomly permuted and original data, $N = 101,161$

Sun. It is cylostationary, and contains *many* solar modes — there are about 1.5 lines per nHz, and at least some of them can be identified with known solar effects.

11. SUMMARY AND CONCLUSIONS

In this paper we have sketched some of the things learned in work with climate, geomagnetic, interplanetary, and seismic data. These data sets, with literally thousands of time series, are mostly long; just one set of seismic data has 122 records ranging from short ones, ~ 60 days duration with only 51 million observations to several with over 100 million samples, and an extraordinary few with unbroken seismic data extending over 400 days. The climate data considered here, in contrast, has a much smaller number of samples, only about 100,000 daily samples, but spanning nearly 300, years. The geomagnetic and space data are intermediate sized with sampling rates lower than they are for seismic data, but long, almost unbroken series are more common. The distinction is important: in seismic data some of the data breaks are to recenter the mass, which disrupts the continuity of signal and introduces artifacts. Since connecting adjacent sections without introducing low frequency artifacts can be tricky, we have mostly chosen to operate with long continuous stretches. With spacecraft data, in contrast, data gaps are often from communications problems and so are easily interpolated. We are not sure if such data count as "big" or not. However, we can testify that visually scanning a data set that extends over years or decades for problems certainly makes it seem like a lot. Even worse, examining a spectrum or coherence estimated from such a data set is a major task. As we have shown, most such estimates made from long data sets show thousands of solar modes or peaks offset from known solar modes by multiples of a cycle/day, a cycle/year, or both. Smoothing such estimates simply confuses the issue and usually just results in a plot that is "wrong everywhere" — it neither shows the

high modal peaks nor the low intervals between modes.

Very few, if any, of these series can be considered to be independent. One sees coupling between data taken in interplanetary space, in ground based magnetometers, and on to seismic data. It seems that all data measured on Earth is influenced by the normal modes of the Sun. This changes how we must think about such data. Previously, one would have considered the "noise" — generally including the things that we cannot control or do not understand — to be independent between sensors spaced by modest distances in either time or space. This is probably true for the preamplifier noise in the instruments, but not for other background fluctuations in air temperature or pressure, geomagnetic, or seismic noise and one expects the list to grow.

Compounding this problem, we have also shown that robust statistical procedures have, perhaps, been overdone and can be misleading when there is not good evidence for outliers. Other data, such as the series of spectral indices from the ACE spacecraft push spectrum estimation methods to all its limits. One used to take some comfort in the idea that spectrum estimates would be asymptotically unbiased and that, if one could just get a long enough span of data, all the problems would disappear. This has turned out to be complete nonsense; we now have very large data sets, the analysis of which has not become simpler as a result of their enormity. In fact, the scientific implications of big data such as the geomagnetic spectral index analysis here, instead of confirming our intuitive notions of turbulence, show that solar turbulence dynamics are rather poorly understood [5]. So one needs even longer data spans, faster sampling, higher precision, more accurate timing and, in addition, nothing is independent. Practically, one can also say goodbye to concepts of mixing, stationarity, proper data, and similar concepts that one uses to prove theorems.

12. ACKNOWLEDGEMENTS

We thank Prof. Gary Pavlis and Dr. Charlotte Haley for their careful and constructive reviews. We thank Maja-Lisa Thomson for coordinating the cross continental response to reviewers. We thank Dave Riegert for his outstanding sleuthing capabilities in finding the necessary files in DJT's filing system so that FLV could respond to the reviewers.

We thank Prof. Cathy Constable for needed LaTex advice, Ms Claire Boteler for proofreading a draft manuscript and Prof. Phil Jones for the European temperature data. DJT thanks NSERC, the Canada Research Chairs Program, CANSSI, and the Bonneville Power Administration (Grant TIP-290) for support.

This material is based upon work supported by the Incorporated Research Institutions for Seismology under their Cooperative Agreement Number EAR-0733069 with the National Science Foundation. The results presented in this paper also rely on data collected at magnetic observatories, particularly those at Eskdalemuir, operated by the British Geological Survey; and Boulder, operated by the United States Geological Survey; and INTERMAGNET for promoting high standards of magnetic observatory practice (`www.intermagnet.org`).

The Bartol Research Institute (BRI) of the University of Delaware, in collaboration with the Laboratory for Extraterrestrial Physics (LEP) at the Goddard Space Flight Center (GSFC), built and delivered a magnetometer instrument for the ACE mission. The Ulysses data came from ESA and NASA. Thanks for D. Gough for his numerous comments.

[5] Reading the introduction to *Mathematics 596 – An Introduction to the Frequency Analysis of Time Series* by John W. Tukey, pp 503–506 of [146] and other papers in this volume and [147] gives much insight into the departures between mathematics and data analysis.

13. REFERENCES

[1] D. Summers, I. R. Mann, D. N. Baker, and M. Schulz, Eds., *Dynamics of the Earth's Radiation Belts and Inner Magnetosphere*, AGU, Washington, 2012, Geophys. Monogr. Ser., vol 199.

[2] E. Parzen, "On asymptotically efficient consistent estimates of the spectral density function of a stationary time series," *J Roy Statist Soc.*, vol. B-20, pp. 303–322, 1958.

[3] D. R. Brillinger, "Asymptotic properties of spectral estimates of second order," *Biometrika*, vol. 56, pp. 375–390, 1969.

[4] R. H. Farrell, "Asymptotic lower bounds for the risk of estimators of the value of a spectral density function," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 49, pp. 221–234, 1979.

[5] D. J. Thomson, "Spectrum estimation techniques for characterization and development of WT4 waveguide, *Part I*," *Bell System Tech. J.*, vol. 56, pp. 1769–1815, 1977.

[6] D. J. Thomson, "Spectrum estimation techniques for characterization and development of WT4 waveguide, *Part II*," *Bell System Tech. J.*, vol. 56, pp. 1983–2005, 1977.

[7] W. D. Warters, Ed., *WT4 Millimeter Waveguide System*, vol. 56, AT&T, Murray Hill, NJ, 1977, Special Issue, *Bell System Tech. J.*

[8] D. J. Thomson, S. Harris, and P. E. Fox, "Mechanical gauging techniques for millimeter waveguide," *Bell System Technical J.*, vol. 56, pp. 2007–2023, 1977.

[9] D. J. Thomson, J. C. Anderson, J. W. Carlin, and T. J. West, "WT4 field evaluation test - transmission medium achievements," *Bell System Technical J.*, vol. 56, pp. 2157–2178, 1977.

[10] G. G. Stokes, "On a method of detecting inequalities of unknown periods in a series of observations," *Proc. R. Soc. London*, vol. 29, pp. 122–123, 1879, Comment on the Preliminary Report to the Committee on Solar Physics, In **Mathematical and Physical Papers, V** Cambridge University Press, 1905, pg 52–53.

[11] N. Wiener, "Generalized harmonic analysis," *Acta Mathematica*, vol. 55, no. 1, pp. 117–258, 1930.

[12] K. Hasselmann, W. H. Munk, and G. MacDonald, "Bispectra of ocean waves," in *Time Series Analysis*, M. Rosenblatt, Ed., pp. 125–139. Wiley, WileyA, 1963.

[13] H. He and D. J. Thomson, "The canonical bicoherence: Part I — definitions, properties, and multitaper estimates," *IEEE Trans. on Signal Processing*, vol. 57, pp. 1273–1284, 2009.

[14] I. Chuine, P. Yiou, N. Viovy, B. Seguin, V. Daux, and E. L. R. Ladurie, "Historical phenology: grape ripening as a past climate indicator," *Nature*, vol. 432, no. 7015, pp. 289–290, 2004.

[15] J. Berger, P. Davis, and G. Ekström, "Ambient Earth noise: A survey of the global seismographic network," *J Geophys. Res.*, vol. 109, pp. B11307, 2004, doi:10.1029/2004JB003408.

[16] D. J. Thomson and F. L. Vernon, "Unexpected, high–Q, low–frequency peaks in seismic spectra," *Geophys. J. Inter.*, vol. 202, pp. 1690–1710, 2015, plus Supplement.

[17] A. D. Chave, "Magnetotelluric data, stable distributions and impropriety: An existential combination," *Geophys. J. Inter.*, vol. 198, pp. 622–636, 2014.

[18] D. J. Thomson, C. G. Maclennan, and L. J. Lanzerotti, "Propagation of solar oscillations through the interplanetary medium," *Nature*, vol. 376, pp. 139–144, 1995.

[19] D. J. Thomson, L. J. Lanzerotti, and C. G. Maclennan, "The interplanetary magnetic field: Statistical properties and discrete modes," *J Geophys. Res.*, vol. 106, pp. 15941–15962, 2001.

[20] S. Ghosh, DJ Thomson, WH Matthaeus, and LJ Lanzerotti, "Coexistence of turbulence and discrete modes in the solar wind," *Journal of Geophysical Research*, vol. 114, no. A8, pp. A08106, 2009.

[21] D. J. Thomson, "Background magnetospheric variability as inferred from long time–series of GOES data," In Summers et al. [1], pp. 225–241, Geophys. Monogr. Ser., vol 199.

[22] K. Nawa, N. Suda, Y. Fukao, T. Sato, Y. Aoyama, and K. Shibuya, "Incessant excitation of the Earth's free oscillations," *Earth Planets Space*, vol. 50, pp. 3–8, 1998.

[23] N. Suda, K. Nawa, and Y. Fukao, "Earth's background free oscillations," *Science*, vol. 279, pp. 2,089–2,091, 1998.

[24] T. Tanimoto, J. Um, K. Nishida, and N. Kobayashi, "Earth's continuous oscillations observed on seismically quiet days," *Geophys. Res. L.*, vol. 25, pp. 1553–1556, 1998.

[25] Sir W. Thomson (Lord Kelvin), "On the rigidity of the Earth," *Phil. Trans. R. Soc. Lond.*, vol. 153, pp. 573–582, 1863.

[26] Sir W. Thomson (Lord Kelvin), "Dynamical problems regarding elastic spheroidal shells and spheroids of incompressible liquid," *Phil. Trans. R. Soc. Lond.*, vol. 153, pp. 583–616, 1863.

[27] F. A. Dahlen and J. Tromp, *Theoretical Global Seismology*, Princeton Univ. Press, Princeton, New Jersey, 1998.

[28] T. G. Masters and R. Widmer, "Free oscillations: Frequencies and attenuations," in *Global Earth Physics*, T. J. Ahrens, Ed., pp. 104–125. AGU, Washington, DC, 1995.

[29] A. Deuss, J. Ritsema, and H. van Heijst, "A new catalogue of normal-mode splitting function measurements up to 10 mhz," *Geophysical Journal International*, vol. 193, pp. 920–937, 2013.

[30] F. L. Vernon and D. J. Thomson, "Unexpected mode observations in the low frequency seismic spectrum," in *EOS, Trans. AGU Jt. Assem. Suppl.* 2007, vol. 88, American Geophysical Union, Abstract S34A–03.

[31] D. J. Thomson and F. L. Vernon, "Characteristics and source of unexpected modes observed in the low frequency seismic spectrum," in *AGU 2007 Fall Meeting, Supplement to EOS*. 2007, American Geophysical Union, Abstract, Session S34A-04.

[32] D. J. Thomson and F. L. Vernon, "Supplement to "unexpected, high–Q, low–frequency peaks in seismic spectra"," *Geophys. J. Inter.*, vol. 202, pp. 14, 2015.

[33] J. W. Harvey, "Helioseismology," *Physics Today*, vol. 48, pp. 32–38, Oct. 1995.

[34] D. J. Thomson, L. J. Lanzerotti, F. L. Vernon, III, M. R. Lessard, and L. T. P. Smith, "Solar modal structure of the engineering environment," *Proceedings of the IEEE*, vol. 95, pp. 1085–1132, 2007.

[35] C. A. L. Haley, *Nonparametric and Parametric Methods for Solar Oscillation Spectra*, Ph.D. thesis, Queen's University, 2014.

[36] D. J. Thomson and C. L. Haley, "Spacing and shape of random peaks in non–parametric spectrum estimates," *Proc. R. Soc. A*, 2014, 470:20140101.

[37] A. Moghtaderi, D. J. Thomson, and G. Takahara, "Unfolding of aliased line component frequencies in bivariate time series," *Canadian J. Statistics*, vol. 38, pp. 116–135, 2010.

[38] D. J. Thomson, L. J. Lanzerotti, and C. G. Maclennan, "Low frequency ($\sim$ 2.19 day period) mode in records of interplanetary and central England temperature data," in *Proc. of the SOHO 6/GONG 98 Workshop, 'Structure and Dynamics of the Sun and Sun-like Stars'*, 1998, vol. SP-418, pp. 967–971, ESA Vol. 2.

[39] D. J. Thomson, "Some problems in the analysis of possibly cyclostationary data," in *Proc. Forty–Fifth Asilomar Conf. on Signals, Systems, and Computers*. 2011, pp. 2040–2044, IEEE.

[40] J. Christensen-Dalsgaard, F. W. W. Dilke, and D. O. Gough, "The stability of a solar model to non–radial oscillations," *Monthly Not. RAS*, vol. 169, pp. 429–445, 1974.

[41] T. Appourchaux, K. Belkacem, A.–M. Broomhall, W. J. Chaplin, D. O. Gough, G. Houdek, J. Provost, F. Baudin, P. Boumier, Y. Elsworth, R. A. García, B. N. Andersen, W. Finsterle, C. Fröhlich, A. Gabriel, G. Grec, A. Jiménez, A. Kosovichev, T. Sekii, T. Toutain, and S. Turck–Chièze, "The quest for the solar g modes," *Astron. Astrophys Rev.*, vol. 18, pp. 197–277, 2010.

[42] M. F. Woodard and R. W. Noyes, "Change in solar oscillation eigenfrequencies with the solar cycle," *Nature*, vol. 318, pp. 449–450, 1985.

[43] R. Howe, R. Komm, and F. Hill, "Solar cycle changes in GONG p–mode frequencies 1995-1998," *Astrophysical J*, vol. 524, pp. 1084–1095, 1999.

[44] R. Komm, R. Howe, and F. Hill, "Solar–cycle changes in GONG p–mode widths and amplitudes 1995–1998," *Astrophysical J*, vol. 531, pp. 1094–1108, 2000.

[45] W. J. Chaplin, T. Appourchaux, Y. Elsworth, G. R. Isaac, B. A. Miller, and R. New, "Source of excitation of low–l solar p modes: characteristics and solar–cycle variations," *Monthly Not. RAS*, vol. 314, pp. 75–86, 2000.

[46] W. J. Chaplin, Y. Elsworth, G. R. Isaac, K. I. Marchenkov, B. A. Miller, and R. New, "Changes to low–l solar p–mode frequencies over the solar cycle: correlations on different time–scales," *Monthly Not. RAS*, vol. 322, pp. 22–30, 2001.

[47] T. Toutain and A.G. Kosovichev, "Study of the solar cycle dependence of low–degree p–modes with Michaelson Doppler Imager and VIRGO," *Astrophysical J*, vol. 622, pp. 1314–1319, 2005.

[48] K. Jain, S. C. Tripathy, and F. Hill, "Solar activity phases and intermediate–degree mode frequencies," *Astrophysical J*, vol. 695, pp. 1567–1576, 2009.

[49] A. Jiménez, R. A. García, and P. L. Pallé, "The acoustic cutoff frequency of the sun and the solar magnetic activity cycle," *Astrophysical J*, vol. 743, no. 2, pp. 99, 2011.

[50] J. Provost, G. Berthomieu, and P. Morel, "Low-frequency p– and g–mode solar oscillations," *Astronomy & Astrophysics*, vol. 353, pp. 775–785, 2000.

[51] J. Beran, *Statistics for Long-Memory Processes*, Chapman and Hall, New York, 1994.

[52] D. J. Thomson, "A test for "long–memory" processes," in *Proceedings of the 2003 IEEE Workshop on Statistical Signal Processing*, St. Louis, MO, 2003, pp. 541–544, IEEE Press, IEEE Cat. No 03TH8705C.

[53] B. Kleiner, R. D. Martin, and D. J. Thomson, "Robust estimation of power spectra (with discussion)," *J Roy Statist Soc, B*, vol. 41, pp. 313–351, 1979.

[54] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, vol. 70, pp. 1055–1096, 1982.

[55] L. Rayleigh, "On the spectrum of an irregular disturbance," *Philosophical Magazine*, vol. 41, pp. 238–243, 1903, (in *Scientific Papers by Lord Rayleigh*, Volume V, Article 285, pages 98-102, Dover Publications, New York, 1964).

[56] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proceedings of the IEEE*, vol. 66, pp. 51–83, 1978.

[57] R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra*, Dover, New York, 1959, Originally published in the *Bell System Tech. J.* Vol. XXXVII, 1958.

[58] R. P. Bogert, M. J. Healy, and J. W. Tukey, "The quefrency alanysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking," in *Spectral Analysis of Time Series*, B. Harris, Ed., pp. 209–243. John Wiley and Sons, New York, 1967.

[59] Michael A Lombardi and Glenn K Nelson, "WWVB: A half century of delivering accurate frequency and time by radio," *Journal of Research of the National Institute of Standards and Technology*, vol. 119, pp. 25–54, 2014.

[60] M. A. Lombardi, "Time measurement," in *Measurement, Instrumentation, and Sensors Handbook*, J. G. Webster and H. Eren, Eds., chapter 41, pp. 41–1 to 41–21. CRC Press, Boca Raton, FL, second edition, 2014, ISBN 9781439848838.

[61] D. G. Childers, D. P. Skinner, and R. C. Kemerait, "The cepstrum: A guide to processing," *Proceedings of the IEEE*, vol. 65, pp. 1428–1443, 1977, Correction, Vol. 66, 1288 (1978).

[62] D. J. Thomson, "Time series analysis of Holocene climate data," *Phil. Trans. R. Soc. Lond. A*, vol. 330, pp. 601–616, 1990.

[63] H. Cramér, "On the theory of stationary random processes," *Ann. of Math.*, vol. 41, pp. 215–230, 1940.

[64] J. L. Doob, *Stochastic Processes*, John Wiley and Sons, New York, 1953.

[65] M. B. Priestley, *Spectral Analysis and Time Series*, Academic Press, San Diego, 1981.

[66] A. Khintchine, "Korrelationstheorie der stationären stochastischen Prozesse," *Mathematische Annalen*, vol. 109, no. 1, pp. 604–615, 1934.

[67] A. Einstein, "Méthode pour la détermination de valeurs statistiques d'observations concernant des grandeurs soumises à des fluctuations irrégulières," *Archives des Sciences*, vol. 37, pp. 254–256, 1914.

[68] A. M. Yaglom, "Einstein's 1914 paper on the theory of irregularly fluctuating series of observations," *IEEE Signal Processing Magazine*, vol. 4, no. 4, pp. 7–11, 1987.

[69] M. Loève, *Probability Theory*, D. Van Nostrand, Princeton, NJ, 1963.

[70] A. N. Kolmogorov, "On the $\Phi^{(n)}$ classes of Fortet and Blanc–Lapierre," *Theory Prob. Appl*, vol. 5, pp. 337, 1960.

[71] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty -I," *Bell System Tech. J.*, vol. 40, pp. 43–64, 1961.

[72] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty V: the discrete case," *Bell System Tech. J.*, vol. 57, pp. 1371–1429, 1978.

[73] D. J. Thomson, "Quadratic-inverse spectrum estimates: applications to paleoclimatology," *Phil. Trans. R. Soc. Lond. A*, vol. 332, pp. 539–597, 1990.

[74] E. J. Hannan, *Multiple Time Series*, John Wiley and Sons, New York, 1970.

[75] M. Taniguchi and Y. Kakizawa, *Asymptotic Theory of Statistical Inference for Time Series*, Springer, New York, 2000.

[76] H. O. Hartley, "The maximum F–ratio as a short–cut test for heterogeneity of variance," *Biometrika*, vol. 37, pp. 308–312, 1950.

[77] P. Stoica and T. Sundin, "On nonparametric spectral estimation," *Circuits Systems Signal Process.*, vol. 18, pp. 169–181, 1999.

[78] M. M. Sondhi, "Random processes with specified spectral density and first–order probability density," *Bell System Tech. J.*, vol. 62, pp. 679–701, 1983.

[79] W. M. Gentleman, "An error analysis of Goertzel's (Watt's) method for computing Fourier coefficients," *Comp. J.*, vol. 12, pp. 160–164, 1969.

[80] W. H. Munk and K. Hasselmann, "Super-resolution of tides," in *Hikada Volume*, K. Yoshida, Ed., pp. 339–344. Univ. of Washington Press, Seattle, 1964, Reprinted, 1965.

[81] K. S. Miller, *Complex Stochastic Processes*, Addison Wesley, Reading, MA, 1974.

[82] H. Scheffé, *The Analysis of Variance*, John Wiley and Sons, New York, 1959.

[83] R. A. Fisher, *Statistical Methods and Scientific Inference*, Macmillan Pub Co, third edition, 1973.

[84] W. S. Cleveland and S. J. Devlin, "Calendar effects in monthly time series: Detection by spectrum analysis and graphical methods," *J American Statistical Soc.*, vol. 75, pp. 487–496, 1980.

[85] K. J. Rahim, *Applications of Multitaper Spectral Analysis to Nonstationary Data*, Ph.D. thesis, Queen's University, 2014.

[86] W. Gleissberg, "The eighty–year sunspot cycle," *J. Br. Astron. Assoc.*, vol. 68, pp. 148–152, 1958.

[87] A. N. Peristykh and P. E. Damon, "Persistence of the Gleissberg 88–year solar cycle over the last $\sim 12,000$ years: Evidence from cosmogenic isotopes," *J Geophys. Res.*, vol. 108, pp. A1, 2003.

[88] J. L. Klein, *Statistical Visions in Time. A History of Time Series Analysis 1662-1938*, Cambridge Univ. Press, 1997.

[89] D. J. Thomson, "The seasons, global temperature, and precession," *Science*, vol. 268, pp. 59–68, 1995.

[90] C. L. Wolff, "Linear r–modes below the Sun's convective envelope," *Astrophysical J*, vol. 531, pp. 591–598, 2000.

[91] N. S. Dzhalilov and J. Staude, "Eigenoscillations of the differentially rotating Sun II. Generalization of the Laplace tidal equation," *Astronomy & Astrophysics*, vol. 421, pp. 305–322, 2004.

[92] W. J. Chaplin, *The Music of the Sun: The story of Helioseismology*, Oneworld Publ., Oxford, 2006.

[93] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.

[94] N. Giri, "On complex analogues of T^2 and R^2 tests," *Annals of Math. Stat.*, vol. 36, pp. 664–670, 1965.

[95] G. E. P. Box, "Non-normality and tests on variances," *Biometrika*, vol. 40, pp. 318–335, 1953.

[96] R. D. Martin and D. J. Thomson, "Robust-resistant spectrum estimation," *Proceedings of the IEEE*, vol. 70, pp. 1097–1115, 1982.

[97] A. D. Chave, D. J. Thomson, and M. E. Ander, "On the robust estimation of power spectra, coherences, and transfer functions," *J Geophys. Res.*, vol. 92, pp. 633–648, 1987.

[98] M. E. Kappus and F. L. Vernon, "Acoustic signature of thunder from seismic records," *J Geophys. Res.*, vol. 96, pp. 10,989–11,006, 1991.

[99] R. Mellors, F. L. Vernon, and D. J. Thomson, "Detection of dispersive signals using multitaper dual–frequency coherence," *Geophys. J. Inter.*, vol. 135, pp. 146–154, 1998.

[100] K. Q. Lepage and D. J. Thomson, "Spectral analysis of cyclostationary time–series: A robust method," *Geophys. J. Inter.*, vol. 179, pp. 1199–1212, 2009.

[101] A. D. Chave and D. J. Thomson, "A bounded influence regression estimator based on the statistics of the hat matrix," *J. Roy. Stat. Soc., Ser. C (Applied Statistics)*, vol. 52, pp. 307–322, 2003.

[102] A. D. Chave and D. J. Thomson, "Bounded influence magnetotelluric response function estimation," *Geophys. J. Inter.*, vol. 157, pp. 988–1006, 2004.

[103] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, John Wiley and Sons, New York, second edition, 1995.

[104] J. Christensen-Dalsgaard, "Helioseismology," *Reviews of Modern Physics*, vol. 74, pp. 1073–1129, 2002.

[105] M. Stix, *The Sun: An Introduction*, Springer–Verlag, Berlin, second edition, 2004.

[106] P. M. Shearer, *Introduction to Seismology*, Cambridge University Press, Cambridge, 1999.

[107] M. L. Goldstein, D. A. Roberts, and W. H. Matthaeus, "Magnetohydrodynamic turbulence in the solar wind," *Annual review of astronomy and astrophysics*, vol. 33, pp. 283–326, 1995.

[108] R. T. Wicks, T. S. Horbury, C. H. K. Chen, and A. A. Schekochihin, "Power and spectral index anisotropy of the entire inertial range of turbulence in the fast solar wind," *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 407, no. 1, pp. L31–L35, 2010.

[109] B. Bavassano, D. Dobrowolny, F. Mariani, and N. F. Ness, "Radial evolution of power spectra of interplanetary Alfvénic turbulence," *J Geophys. Res.*, vol. 87, pp. 3617–3622, 1982.

[110] T. S. Horbury, M. Forman, and S. Oughton, "Anisotropic scaling of magnetohydrodynamic turbulence," *Physical Review Letters*, vol. 101, no. 17, pp. 175005, 2008.

[111] E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Cambridge Univ. Press, 1970.

[112] P. Delache and P. H. Scherrer, "Detection of solar gravity mode oscillations," *Nature*, vol. 306, pp. 651–653, 1983.

[113] P. L. Pallé, "The search for solar gravity modes," *Adv. Space Res.*, vol. 11, no. 4, pp. 29–38, 1991.

[114] P. Kumar, E. J. Quataert, and J. N. Bahcall, "Observational searches for solar g-modes: some theoretical considerations," *Astrophysical Journal, Letters to the Editor*, vol. 458, pp. L83–L85, 1996.

[115] J. T. Nolte and E. C. Roelof, "Large–scale structure of the interplanetary medium I: High coronal source longitude of the quiet–time solar wind," *Solar Physics*, vol. 33, pp. 241–257, 1973.

[116] J. T. Nolte and E. C. Roelof, "Large–scale structure of the interplanetary medium II: Evolving magnetic configurations deduced from multi–spacecraft observations," *Solar Physics*, vol. 33, pp. 483–504, 1973.

[117] J. Schou, A. G. Kosovichev, P. R. Goode, and W. A. Dziembowski, "Determination of the Sun's seismic radius from the SOHO Michelson Doppler Imager," *Astrophysical J*, vol. 489, pp. L197–L200, 1997.

[118] L. J. Lanzerotti, D. J. Thomson, and C. G. Maclennan, "Engineering issues in space weather," in *Modern Radio Science 1999*, M. A. Stuchly, Ed., pp. 25–50. Oxford University Press, Oxford, 1999.

[119] D. H. Boteler, R. J. Pirjola, and H. Nevanlinna, "The effects of geomagnetic disturbances on electrical systems at the Earth's surface," *Adv. Space Res.*, vol. 22, pp. 17–27, 1998.

[120] W. H. Barlow, "On the spontaneous electrical currents observed in the wires of the electric telegraph," *Phil. Trans. R. Soc. Lond.*, vol. 139, pp. 61–72, 1849.

[121] A. D. Chave and A. G. Jones, Eds., *The Magnetotelluric Method*, Cambridge University Press, Cambridge, 2012.

[122] L. J. Lanzerotti, D. J. Thomson, A. Meloni, L. V. Medford, and C. G. Maclennan, "Electromagnetic study of the Atlantic continental margin using a section of a transatlantic cable," *J Geophys. Res.*, vol. 91, pp. 7417–7427, 1986.

[123] S. Haykin, D. J. Thomson, and J. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, pp. 849–877, 2009.

[124] G. B. Airy, "First analysis of one hundred and seventy–seven magnetic storms, registered by the magnetic instruments in the Royal Observatory, Greenwich, from 1841 to 1857.," *Phil. Transactions of the Royal Society of London*, vol. 153, pp. 617–648, 1863.

[125] D. J. Thomson, L. J. Lanzerotti, and C. G. Maclennan, "Coherent frequency variations in electron fluxes at 1 and 5 AU in the inner heliosphere," in *Acceleration and Transport of Energetic Particles Observed in the Heliosphere, Proc. ACE2000 Symp.* 2000, vol. 528, pp. 278–281, AIP.

[126] S. C. Webb, "The Earth's hum: the excitation of Earth normal modes by ocean waves," *Geophys. J. Inter.*, vol. 174, pp. 542–566, 2008.

[127] T. Forbriger, "Reducing magnetic field induced noise in broad–band seismic recordings," *Geophys. J. Inter.*, vol. 169, pp. 240–258, 2007.

[128] R. A. García, C. Régulo, S. Turck-Chièze, L. Bertello, A. G. Kosovichev, A. S. Brun, S. Couvidat, C. J. Henney, M. Lazrek, R. K. Ulrich, and F. Varadi, "Low-degree low-order solar p modes as seen by GOLF on board SOHO," *Solar Physics*, vol. 200, pp. 361–379, May 2001.

[129] D. Salabert, J. Leibacker, T. Appourchaux, and F. Hill, "Measurement of low signal–to–noise ratio solar p–modes in spatially resolved helioseismic data," *Astrophysical J*, vol. 696, pp. 653–667, 2009.

[130] A.–M. Broomhall, W. J. Chaplin, G. R. Davies, Y. Elsworth, S. T. Fletcher, S. J. Hale, B. Miller, and R. New, "Definitive Sun–as–a–star p–mode frequencies: 23 years of BiSON observations," *Monthly Not. RAS*, vol. 396, pp. L100–L104, 2009.

[131] S. M. Kay and S. L. Marple, "Spectrum analysis - A modern perspective," *Proceedings of the IEEE*, vol. 69, pp. 1380–1419, 1981.

[132] R. D. Martin and R. H. Zamar, "Asymptotically Min–Max bias robust M-Estimates of scale for positive random variables," *J American Statistical Soc.*, vol. 84, pp. 494–501, 1989.

[133] D. Camuffo and P. D. Jones, Eds., *Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources*, Kluwer Academic Publishers, Dordrecht, 2002.

[134] H. Bergström and A. Moberg, "Daily air temperature and pressure series for Uppsala (1722–1998)," *Climatic Change*, vol. 53, pp. 213–252, 2002.

[135] T. R. Karl, P. D. Jones, R. W. Knight, O. R. White, W. Mende, J. Beer, and D. J. Thomson, "Testing for

bias in the climate record," *Science*, vol. 271, pp. 1879–1883, 1996.

[136] D. J. Thomson, "Dependence of global temperatures on atmospheric CO_2 and solar irradiance," *Proc. Natl. Acad. Sci. USA*, vol. 94, pp. 8370–8377, 1997.

[137] P. J. Rousseeuw, "Least median of squares regression," *J. Am. Stat. Assoc.*, vol. 79, pp. 871–880, 1984.

[138] H. L. Hurd and A. Miamee, *Periodically Correlated Random Sequences*, John Wiley and Sons, New York, 2007.

[139] A. Napolitano, *Generalizations of Cyclostationary Signal Processing*, IEEE Press; Wiley, Chichester, West Sussex, UK, 2012.

[140] E. Rieger, G. H. Share, D. J. Forrest, G. Kanbach, C. Reppin, and EL Chupp, "A 154–day periodicity in the occurrence of hard solar flares?," *Nature*, vol. 312, pp. 623–625, 1984.

[141] D. J. Thomson, "Jackknifing multitaper spectrum estimates," *IEEE Signal Processing Magazine*, vol. 24, no. 7, pp. 20–30, July 2007.

[142] A. Moberg, H. Bergström, J. R. Krigsman, and O. Svanered, "Daily air temperature and pressure series for Stockholm (1756–1998)," *Climatic Change*, vol. 53, pp. 171–212, 2002.

[143] D. Camuffo, "History of the long series of daily air temperature in Padova (1725–1998)," *Climatic Change*, vol. 53, pp. 7–75, 2002.

[144] L. J. Lanzerotti, R. E. Gold, K. A. Anderson, T. P. Armstrong, R. P. Lin, S. M. Krimigis, M. Pick, E. C. Roelof, E. T. Sarris, G. M. Simnett, and W. E. Frain, "Heliosphere instrument for spectra, composition, and anisotropy at low energies," *Astronomy and Astrophysics Suppl. Ser.*, vol. 92, pp. 349–363, 1992.

[145] B. Stewart and W. Dodgson, "Preliminary report to the committee on solar physics on the evidence in favour of the existence of certain short periods common to solar and terrestrial phenomena," *Proc. R. Soc. Lond.*, vol. XXIX, pp. 303–324, 1879.

[146] D. R. Brillinger, *The Collected Works of John W. Tukey: Volume I Time Series: 1949-1964*, Wadsworth, Belmont, CA, 1984.

[147] D. R. Brillinger, *The Collected Works of John W. Tukey: Volume II Time Series: 1965-1984*, Wadsworth, Belmont, CA, 1984.