Wavelet based estimation of local Kolmogorov turbulence

G. Papanicolaou* K. Sølna†

July 6, 1999

Abstract

We present a new approach for analyzing locally stationary processes with power law spectral densities. We divide the data into segments over which the process is essentially stationary and then use the wavelet scale spectrum to estimate the parameters of the power law, which are the scale factor and the exponent. These parameters vary from segment to segment, with part of the variation due to the nonstationarity of the data and part due to estimation errors that depend on the length of the segments. In the approach we introduce here, segmentation effects due to estimation errors are removed by filtering. We also estimate an effective local inertial range, that is, the set of scales over which the process can be modeled by a power law. We apply our estimation method to atmospheric temperature data that are expected to have Kolmogorov power law spectra. We find that there are significant fluctuations about the Kolmogorov law and analyze them in detail.

Contents

*Department of Mathematics, Stanford University, Stanford CA 94305; papanico@math.stanford.edu
Supported by AFOSR grant F49620-98-1-0211, by ONR grant N00014-96-1-0719 and by NSF grant DMS-9709320,
†Department of Mathematics, University of Utah, Salt Lake City, Utah 84112; solna@math.utah.edu
Supported by the Norwegian Research Council.
Keywords: Wavelets, scale spectra, locally stationary time series, Fractional Brownian motion, data segmentation, atmospheric turbulence.

1 Introduction

Stochastic processes that are approximately stationary and have approximately power law spectral densities arise frequently in modeling atmospheric turbulence, financial data, geophysical data, etc. How can we estimate the variable power law behavior of the spectral densities from data? The analysis will depend on how we segment the data and on how we choose the range of scales, or frequencies, over which we look for a power law fit. In this paper we address these issues using fractional Brownian motion as the undelying stochastic model whose parameters are estimated locally by wavelet scale spectra. We then apply the theory to atmospheric turbulence data. This data was first analyzed in [35] and subsequently in [30, 19, 24]. We provide here a general framework for estimating local power law processes and a large part of the mathematical background needed for its justification.

In order to deal with approximate stationarity and approximate power law spectra we need a good understanding of the estimation issues for stationary, power law processes, like fractional Brownian motion. Previous work on the statistics of wavelet scale spectra of fractional Brownian motion can be found [2, 7, 34, 25, 36, 37]. In Section 6 we provide a brief but complete analysis of these statistics, that is based on a general formula for the covariance of the Haar wavelet coefficients. A central limit theorem for the estimators of the power law parameters also follows readily from the covariance formula. When Fourier spectra are used the statistical analysis of global power law processes is given in [32] and an analysis of estimators for the exponent of the power law is given in [11, 16]. We use wavelet scale spectra because they provide a time-scale decomposition of the data that is well suited to power law processes, whether they are stationary (have stationary increments) or not. A comparison of power law estimation based on wavelet scale spectra and on Fourier spectra, in the stationary case, is given in [2].

In Section 2 we introduce Fractional Brownian motion as a model for turbulence and discuss the wavelet based scale spectrum that can be used for spectral estimation of such processes.

In Section 3 we introduce the atmospheric data, and carry out a data-analysis using the Haar-wavelet basis. This analysis shows that non stationary effects are indeed important for this data set. The range over which to the process can be modeled as a power law, the inertial range, as it is called in turbulence theory, vary with spatial location. The power law parameters associated with the inertial range also depend on location.

In Section 4.2 we address the issue of choosing the range over which to fit the power law by regression. We do this by using the fractional Brownian motion as a model locally. Since the process is only locally stationary power law parameters must be estimated based on relatively short spatial segments. In Section 4.3 we show how to remove the variability of the estimated power law parameters that is due to the finiteness of the segments. The filtering that we do here is used frequently in geostatistics [3, 29].

In Section 5 we return to the atmospheric data set and use the framework for estimation introduced in Section 4 for estimation of its spectrum. We segment the data into intervals over which they are approximately stationary without resorting to a full segmentation search as in the method of [17] that is based on the local cosine transform. This requires a rough estimate of the size of the intervals of stationarity, which can be obtained from a variogram analysis of the wavelet coefficients, for example, as we do with the aerothermal data of Section 3.2. Another case where the intervals of stationarity are known approximately is analyzed in [1]. In Section 5.3 we show some simulations from the estimated model that assess the overall relevance of our analysis. The main point of our analysis is that we are able to identify the local variations of the power law parameters, the Kolmogorov turbulence law, which arise from large scale atmospheric phenomena. The analysis should be applicable to other data sets, financial data for example, where departure from stationarity needs to be quantified.

2 Scale spectrum of Fractional Brownian motion

2.1 Fractional Brownian motion

We shall model 'pure' power law processes by fractional Brownian motion (fBm), $\{B_H(x); x \geq 0\}$, introduced by Mandelbrot and Van-Ness [20]. It is a Gaussian process with mean zero, stationary increments and covariance

$$E[B_H(x)B_H(y)] = \frac{\sigma^2}{2}(|x|^{2H} + |y|^{2H} - |x-y|^{2H}) \tag{2.1}$$

with $0 \leq H \leq 1$ and σ parameters. Its structure function is

$$E[(B_H(x) - B_H(x - \Delta x))^2] = \sigma^2 |\Delta x|^{2H},$$

and it is conditioned to be zero at the origin: $B_H(0) = 0$. The so called Hurst exponent H determines the correlation of the increments. The covariance of future increments with past ones is

$$\rho_H(\Delta x) = E[(B_H(x) - B_H(x-\Delta x))(B_H(x+\Delta x) - B_H(x))] = \sigma^2(2^{2H-1} - 1)|\Delta x|^{2H},$$

which is independent of x. When $H > 1/2$ this quantity is positive so if the past increment is positive, then on average the future increment will be positive. Feder [5] calls this persistence. When $H < 1/2$ we have an antipersistent process with a positive increment in the past making a positive increment in the future less likely. Ordinary Brownian motion corresponds to $H = 1/2$. In this case future and past increments are independent. Of special interest here is $H = 1/3$ which corresponds to the Kolmogorov scaling law [21].

Fractional Brownian motion has stationary increments but is not itself stationary. However, as shown in [2] for example, it is possible to assign a pseudo-spectrum to it by cutting off the low frequencies. Let

$$X = B_H * \psi$$

with star denoting convolution and ψ a function that integrates to zero so that its Fourier transform Ψ vanishes at zero frequency. The process X is stationary and its power spectrum is

$$P_X \quad \propto \quad \sigma^2 |f|^{-(2H+1)} |\Psi(f)|^2.$$

Since we usually observe power law processes through a filter that cuts off very low frequencies, we can associate with B_H the power law spectrum

$$\widetilde{P}_{B_H} \quad \propto \quad \sigma^2 |f|^{-(2H+1)} \ .$$

In the Kolmogorov case $H = 1/3$ the spectrum is $\sigma^2 |f|^{-5/3}$ over some range of frequencies, called the inertial range.

Fractional Brownian motion is self-similar since $B_H(x)$ and $a^H B_H(x/a)$ have the same finite dimensional distributions for all a. We next discuss estimation of fBm, or a pure power law process. In Section 4 we generalize the estimation procedure to locally stationary power law processes.

2.2 Haar wavelets and scale spectrum

We want to carry out a spectral analysis of the process, given as a finite set of data, and to fit the estimated spectrum to a power law. For this purpose **scale** spectra, rather than Fourier spectra, will be used. Scale spectra are a natural and flexible tool for self-similar processes [2, 14]. The scale spectrum is defined in terms of the coefficients of the data in a wavelet basis. We will use the Haar wavelet basis although other bases could have been used as well [10]. Haar wavelet based estimators are sometimes related to classical ones, as is in particular the Allan variance estimator for the slope of the log of the spectral density [27].

Let Y denote the process for which we want to compute the Haar wavelet coefficients. Denote the *approximation* coefficients at level zero by $X = (a_0(1), a_0(2), ..., a_0(2^M))$. These are defined by

$$a_0(n) \quad = \quad \int_{n-1}^{n} Y(x) dx \ .$$

Given this level zero representation, the data, we construct successively its wavelet coefficients with respect to the Haar basis as follows. Let

$$\begin{aligned} a_1(n) &= \frac{1}{\sqrt{2}}(a_0(2n) + a_0(2n-1)) \\ d_1(n) &= \frac{1}{\sqrt{2}}(a_0(2n) - a_0(2n-1)) \ , \quad \text{for } n = 1, 2..., 2^{M-1} \end{aligned} \tag{2.2}$$

be the smoothed signal and its fluctuation, or detail, at the finest scale. Note that the *detail* vector d_1 contains every other successive difference of the data. This process of averaging and differencing can be continued by defining

$$\begin{aligned} a_2(n) &= \frac{1}{\sqrt{2}}(a_1(2n) + a_1(2n-1)) \\ d_2(n) &= \frac{1}{\sqrt{2}}(a_1(2n) - a_1(2n-1)) \ , \quad \text{for } n = 1, 2..., 2^{M-2} \end{aligned}$$

and in general

$$\begin{aligned} a_j(n) &= \frac{1}{\sqrt{2}}(a_{j-1}(2n) + a_{j-1}(2n-1)) \\ d_j(n) &= \frac{1}{\sqrt{2}}(a_{j-1}(2n) - a_{j-1}(2n-1)) \ , \quad \text{for } n = 1, 2..., 2^{M-j} \end{aligned}$$

for $j = 1, ..., M$. The data vector X can then be reconstructed from $a_M, d_M, d_{M-1}, ..., d_1$ since from equations (2.2) we have

$$\begin{aligned} a_0(2n) &= \frac{1}{\sqrt{2}}(a_1(n) + d_1(n)) \\ a_0(2n-1) &= \frac{1}{\sqrt{2}}(a_1(n) - d_1(n)) \ , \quad \text{for } n = 1, 2..., 2^{M-1} \end{aligned}$$

and now a_1 can be replaced by sums and differences of a_2 and d_2, etc.

The detail coefficients at level j can alternatively be expressed as

$$d_j(n) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} \psi(x/2^j - n) Y(x) dx$$

with the mother wavelet defined by

$$\psi(x) = \begin{cases} -1 & \text{if } -1 \leq x < -1/2 \\ 1 & \text{if } -1/2 \leq x < 0 \\ 0 & \text{otherwise} \end{cases} .$$

The difference coefficients correspond to probing the process at different scales and locations, with n representing location and j scale. From the self-similarity of fractional Brownian motion it follows that for such a process $E[d_j(n)^2] \propto 2^{j(2H+1)}$.

The scale spectrum of X relative to the Haar wavelet basis is the sequence S_j defined by

$$S_j = \frac{1}{2^{M-j}} \sum_{n=1}^{2^{M-j}} (d_j(n))^2 \ , \quad j = 1, 2, ..., M. \tag{2.3}$$

It is easily verified from the definitions above that the l^2 norm of the data vector X can be written as

$$\sum_{n=1}^{2^M} (a_0(n))^2 = (a_M)^2 + \sum_{j=1}^{M} 2^{M-j} S_j \ ,$$

which is a way of expressing the orthogonality of the decomposition of X into a_M and the d_j, $j = 1, ..., M$.

The scale spectral point S_j is the mean square of the detail coefficients at scale j. The spectrum can therefore be interpreted as the energy of the signal in the different scales relative to the Haar wavelet basis. Consider data containing information only at the finest scale: $X = \{1,\ -1,\ 1,\ -1,\ \cdots\}$. Then only the $d_1(n)$ coefficients are non-zero. Hence, S_j roughly corresponds to the energy at $2^{(j-1)}$ times the finest scale.

Note that the spatial support of the integrals of the process defining the wavelet coefficients at a certain scale is adapted to the particular scale. The short scales, that provide high frequency information, are defined in terms of consecutive integrals of narrow support. A plot of these difference coefficients reveals information about how the high frequency content of the data change with location.

3 The aerothermal data

3.1 Temperature and index of refraction fluctuations

Our objective is to estimate the spectrum of temperature data taken from the upper atmosphere. One motivation for our modeling and estimation is to be able to generate synthetic random media for the simulation of wave propagation in a turbulent atmosphere. Before we analyze the aerothermal data we explain briefly how velocity, temperature and index of refraction fluctuations are related to each other and how the Kolmogorov scaling theory enters in their description [21, 31].

The Kolomogorov theory is a phenomenological statistical description of the *velocity* field in the atmosphere. Based on a scaling argument, the mean-square velocity differences are described in a universal manner over a rather broad range of spatial scales, the inertial range. If we assume homogeneity, isotropy and incompressibility the result is that the structure function of the velocity has the form

$$E[(\mathbf{v_x}(\mathbf{x_0}+\mathbf{x})-\mathbf{v_x}(\mathbf{x_0}))^2] \quad = \quad C_{\mathbf{v}}^2|\mathbf{x}|^{2/3} \tag{3.1}$$

Here $\mathbf{v_x}$ is the velocity in the direction of the displacement $\mathbf{x}$. The parameter $C_{\mathbf{v}}^2$ is the velocity structure constant, a measure of the amount of energy in the inertial range, which is typically confined between an inner scale l_0 and an outer scale L_0.

The connection to temperature is through the theory of convection-diffusion of a passive scalar. It turns out [21] that temperature fluctuations have also a structure function with Kolmogorov 2/3 scaling.

At optical frequencies the variations in the index of refraction δn are approximately

$$\delta n = -79P\frac{\delta T}{T^2}\times 10^{-6}$$

with P atmospheric pressure in millibars and T temperature in degrees Kelvin. Thus, in order to describe the spectrum of fluctuations in the refractive index we need only the spectrum of the fluctuations in the temperature. Wavelet scale spectra are used to analyze turbulence data in [13].

The above model for refractive index fluctuations is used extensively to model atmospheric wave propagation. Variations in the turbulent character of the medium are typically modeled through variations in $C_{\mathbf{v}}^2$ only and not in the exponent, which is fixed at 2/3.

3.2 Data analysis of aerothermal data

We will analyze temperature data obtained by the Air Force high-altitude laser propagation and turbulence data collection effort. For a detailed discussion of recording procedures and analysis see [30, 24, 35]. A stong effort was made to provide high quality data that could be used to characterize the turbulent atmosphere. This unique data set cannot be maid generally available, but the above references provide together a fairly detailed account for the data. Here we take the data as our starting point and examine what analysis reveals about their structure. We are interested in particular in accounting properly for nonstationary effects. All calculations are carried out in MATLAB on a Silicon Graphics workstation.

In Figure 3.1 we show the temperature data, which has approximately 4.2×10^6 points. The spatial resolution is approximately 2cm. The data is, of course, quite noisy and it is part of our task to remove spurious noise effects in the estimation process.

In Figure 3.2 we show the first $21km$ of the temperature data along with $d_1, d_3, d_5, d_7, d_9, d_{11}, a_{11}$. The detail coefficients d_j carry information about the data on larger scales as j increases. For

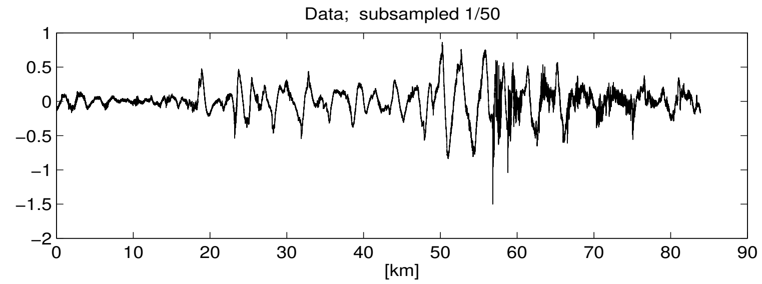

Figure 3.1: Top figure: the raw temperature data. The spatial resolution of the data is approximately 2cm.

example, d_7 shows successive differences of the temperature over distances of $2.56m$, after averaging over successive segments of length 1.28. Thus, as j increases the data are *lowpass* filtered with a filter of decreasing bandwidth and then *subsampled* before the differences are formed. It is clear that the visible periodic components seen in the d_1 coefficients cannot be attributed to turbulence or larger coherent structures in the atmosphere. In the estimation of atmospheric turbulence parameters these spurious features must be suppressed. In $d_5, \cdots$ the high frequency periodic noise component seen above has been effectively suppressed by the lowpass filtering.

Our main objective is now to examine whether the data can be modeled well by a 'power law' model over a subrange of scales. Recall that for such a process, the scale spectrum S_j is linear in a log-log plot, when the record is very long. In Figure 3.3 we show log-log plots of scale spectra over nonoverlapping segments of the data, with each segment having length $1.3km$. The top plot corresponds to the first half of the data whereas the bottom plot corresponds to the second half. Each plot contains scale spectral points over 16 scales, $1 \leq j \leq 16$. This corresponds to length scales from $l_1 = .04m$ to $l_{16} = 1.3km$. We take as abscissa the spatial frequency K_j

$$K_j \;=\; \frac{\pi}{\Delta x} 2^{1-j} \;[\mathrm{rad/m}] \;=\; 100\pi 2^{-j} \;[\mathrm{rad/m}]. \tag{3.2}$$

Note that the scale spectra show a distinct departure from power law behavior for scales below one meter ($K = 6$). For larger scales, power law behavior may be considered in the range between $2.5m$ to $80m$ ($.08 \leq K \leq 2.5$) approximately, which corresponds to the detail coefficients d_7 to d_{12}. The departure from power law behavior for the shortest scales is partly due to the influence of measurement noise. The first half of the data is less energetic than the second half and hence more noisy. The estimated intercept and slope for the log scale spectra depend on the particular segment of data used. We want to identify the part of this variability that is due to the nonstationarity of the process and minimize variability due to estimation errors.

For fractional Brownian motion the wavelet coefficients will be normally distributed. In Figure 3.4 we show a histogram of the wavelet coefficients at scale 8, normalized by a local estimate of the variance. The dashed line corresponds to a Gaussian distribution. In the bottom plot we show the autocorrelation of these wavelet coefficients, normalized by their variance. We plot the correlation in terms of the empirical variogram computed, as in (5.1) below. The dashed line is the theoretical correlation for fBm with Kolmogorov scaling, $H = 1/3$, as defined in (6.2). In Figure 3.5 we show an estimate of the structure function for $\log(d_j(n)^2)$; $7 \leq j \leq 9$. If the process had a pure power law spectrum over these scales we would see only a very small correlation. However, these quantities are correlated on the order of km, indicating that we have a local power law structure.

Before we continue with the scale spectral analysis we note the following.

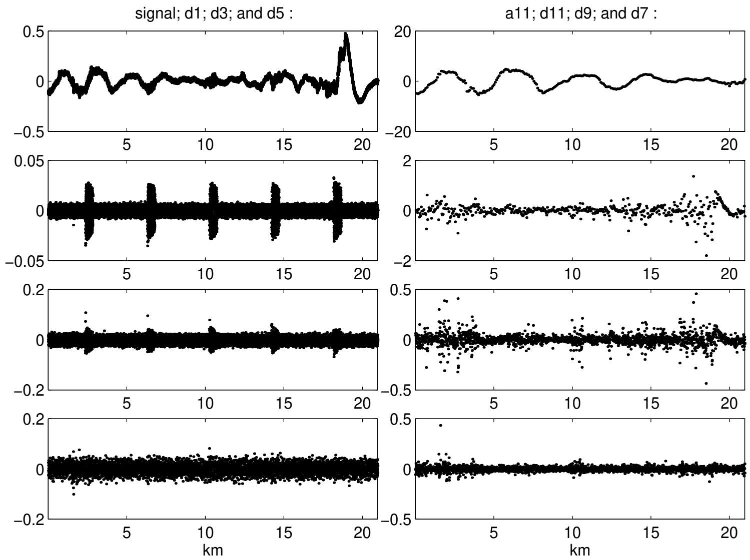

Figure 3.2: Haar wavelet coefficients for the first 21km of the data. The top left figure is the temperature data. The one below it is the d_1 detail coefficients. The third from the top is the d_3 and the bottom the d_5 detail coefficients. Note that the 4km burst seen in d_1 is not visible any more at the d_5 level, after four successive averagings of the data. The top figure on the right is the a_{11} coefficients and below it d_{11}, d_9, d_7.

- Noise bursts in the data enter only in some of the detail coefficients, as can be seen from Figure 3.2. The criterion for selecting of the inertial range, implemented in the next section, will automatically restrict the scale range.

- The temperature data are not a stationary time series and they do not have stationary increments. Computing scale spectra over *long* segments, in which the process cannot be taken as stationary, gives a quantity that is hard to interpret. Even though the average slope of the log scale spectra over several segments is close to $-5/3$, as the theory of turbulence predicts, there is a lot of variability. A **local** power law model of the kind discussed in Section 4 is likely to fit the data better than the idealized power law model, with stationary increments.

We will estimate the scale spectrum of the data when we model it as a local power law, using the method described in Section 4. To start the estimation we must have rough estimates for the intervals of stationarity and for the exponent of the power law. In view of Figures 3.3, 3.4 and 3.5 we choose these estimates to be

- $(\hat{\varepsilon})^{-1} = 2km^{-1}$

- $H_0 = 1/3$ (Kolmogorov scaling) .

In the next section we develop a framework for estimation of local power law processes with varying inertial range and power law parameters. In Sections 5.1 and 5.2 we carry out the estimation. First we estimate the inertial range and then the power law parameters.

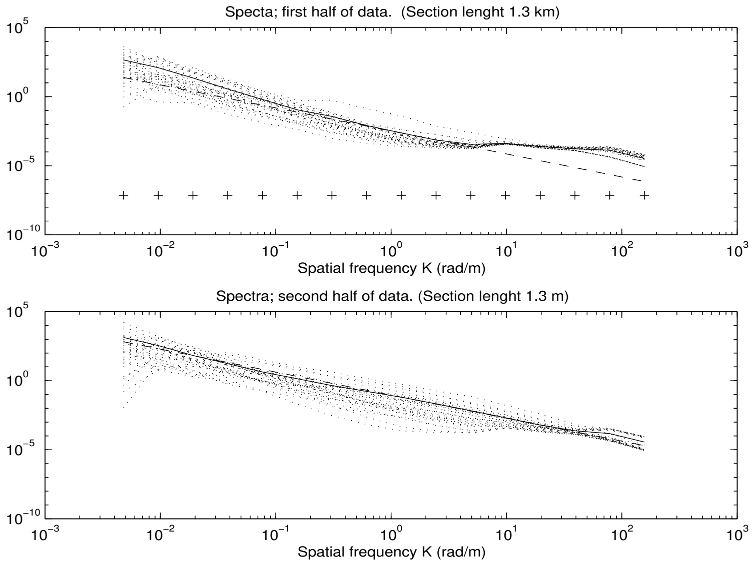

Figure 3.3: Scale spectra of $1.3km$ nonoverlapping segments of the data (2^{16} points per segment) obtained from the Haar wavelet decomposition. The top plot corresponds to the first half of the data whereas the bottom plot corresponds to the second half. The dashed line has slope $-5/3$ as in Kolmogorov spectra. The solid line is the average over the scale spectra of the different segments. After the averaging is done the scale spectra are plotted in log-log format.

4 Estimation of local power law processes

4.1 Modeling and segmentation

The fractional Brownian motion described in Section 2 is an idealization of a process with power law spectrum. It has stationary increments and if the power law is truncated at low frequencies then it is itself stationary. In most physical or financial applications where power law behavior is expected locally, the power law parameters will vary so the process cannot be stationary in the large. The aerothermal data provide an example. We would like to estimate the power law parameters over segments that are short enough so that they can be taken as constant but long enough so that their statistical estimates are stable. How are we to decide what is a good segmentation of the data in this vague sense? This is a very difficult problem that is rarely addressed in theoretical or applied studies of spectral estimation. In [17] we introduced a class of locally stationary processes and developed an algorithm for finding intervals of approximate stationarity. This method is not designed to estimate a roughly constant interval of stationarity associated with only a subset of scales and do not take into account the fact that we are dealing with turbulent data. Thus here we want to follow a somewhat different approach where we can take advantage of prior information about intervals of stationarity, as is the case of the aerothermal data that we consider in Section 5.

To fix ideas we will consider estimation of parameters for a multifractional Brownian motion (mBm). This is a generalization of fractional Brownian motion where the parameters vary with location. In the time domain they have the representation, [9, 25, 26]

$$B_\varepsilon(x) \quad = \quad \frac{\sigma_\varepsilon(x)}{\Gamma(H_\varepsilon(x)+1/2)} \int_{-\infty}^{0} [(x-s)^{H_\varepsilon(x)-1/2} - (-s)^{H_\varepsilon(x)-1/2}] dW(s) \tag{4.1}$$

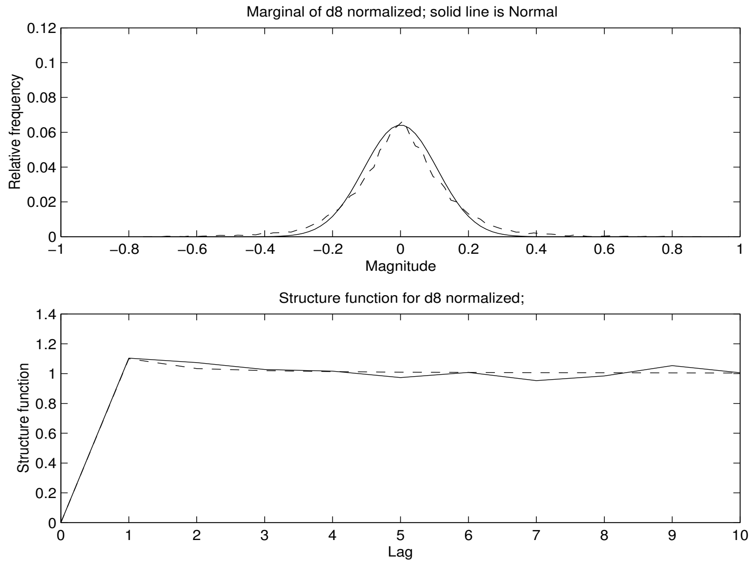

Figure 3.4: The top plot shows the empirical distribution of d_8. The dashed line is a Gaussian distribution. The bottom plot shows the correlation of these wavelet coefficients. The dashed line is the theoretical correlation for fBm with $H = 1/3$.

$$+\int_0^x (x-s)^{H_\varepsilon(x)-1/2} dW(s) \tag{4.2}$$

with W the standard Brownian motion and

$$H_\varepsilon(x) = H(\varepsilon x) \tag{4.3}$$

$$\sigma_\varepsilon(x) = \sigma(\varepsilon x) \tag{4.4}$$

where $H(\cdot)$ & $\sigma(\cdot)$ can be deterministic or random. For example, they can be stationary stochastic processes with smooth paths and decaying correlation functions, independent of the Brownian path $W(\cdot)$. When H and σ are constants then (4.1) is a time domain representation of the usual fractional Brownian motion.

The parameter ε^{-1} is a measure of the interval of stationarity in the sense that, as in [9],

$$E[(B_\varepsilon(x) - B_\varepsilon(x - \Delta x))^2] \approx \sigma_\varepsilon^2(x)|\Delta x|^{2H_\varepsilon(x)}$$

for $\varepsilon\Delta x$ small. This means that for scales that are small compared to the interval of stationarity, the processes $B_\varepsilon(x)$ behaves locally like a fractional Brownian motion with parameters frozen at x.

In Section 4.3 we will describe a method that removes dependence of the estimated parameters on a prior estimate of the interval of stationarity. The basic idea is that we know roughly what $1/\varepsilon$ should be and how the estimation errors behave if the underlying process is fractional Brownian motion, given the segment size. We then filter out these estimation errors and get estimates that do not depend sensitively on the prior choice of the size of the interval of stationarity.

Another modification that is necessary in the nonstationary case is the identification of the inertial range over each interval of approximate stationarity. The error in the power law fit to the scale spectra, over a segment of data, depends on the scale range that is used. How do we select a range of scales for which the error in the fit is acceptable? In the next section we introduce a criterion for selecting the range based on comparison with an ideal fractional Brownian motion.

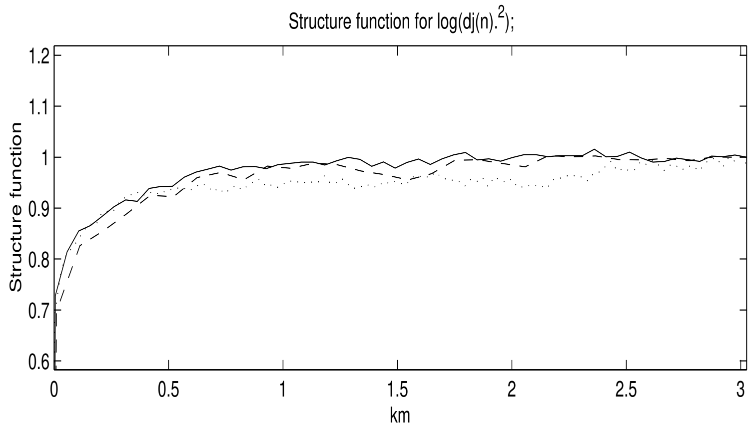

Figure 3.5: The figure shows an estimate of the structure function of $\log(d_j(n)^2)$ as function of $n\Delta x$ and with $7 \leq j \leq 9$. The magnitude of these quantities determine the scale spectrum and they decorrelate over a scale on the order of km. The dotted, dashed and solid lines corresponds respectively to scales 7,8 and 9.

4.2 Estimation of the inertial range

The power law behavior that we want to identify in the data is necessarily restricted to a finite set of scales, the inertial range, in each segment of stationarity.

Let 2^M denote the length of the data vector over the segment under consideration. For an ideal fractional Brownian motion with Hurst exponent H let $S_i,\ i = 1, \ldots, M$ be its wavelet scale spectrum. Then, for each subrange $\{i_1, \cdots, i_2\}$ the measure of misfit

$$r(i_1, i_2) = \sum_{i=i_1}^{i_2} [\log_2(S_i) - \log_2(\hat{S}_i]^2, \tag{4.5}$$

with $\hat{S}_i$ the power law estimated by weighted least squares, is a random variable whose law can be computed analytically in principle or numerically. It depends weakly on H so we fix it to equal some rough estimate $H = H_0$. For the aerothermal data of Section 3 we take $H_0 = 1/3$, the Kolmogorov exponent. Let $R(i_1, i_2)$ be the 90th percentile of the distribution of $r(i_1, i_2)$ with the value of the Hurst parameter equal to H_0. The scale range is now chosen as the *largest* range $\{i_1, \cdots, i_2\}$ for which $\tilde{r}(i_1, i_2) < R(i_1, i_2)$. Here $\tilde{r}(i_1, i_2)$ is the value of the error (4.5) obtained from the actual data.

In order to obtain stable power law parameter estimates in the regression we need a minimum number of scale range points. In the application to the aerothermal data introduced in Section 3 we choose to model in terms of a power law only over segments for which there are 5 or more points in the estimated scale range.

4.3 Segmentation independent power law estimation

Given a fixed segmentation of the data into approximate intervals of stationarity that is based on some prior information, we first calculate the inertial range as described in the previous section and then do the power law fit. The power law parameters are obtained by weighted linear regression of the log scale spectrum, see Section 6.4. The estimated power law parameters depend on the segmentation. We will now describe a method with which this dependence can be removed.

The idea is to do a filtering of the parameter estimates in order to remove the variability that is segmentation dependent. From the theory of the power law estimators for fractional Brownian motion we know that the slope estimator has the form

$$\hat{p}^i \quad = \quad p^i + w^i \tag{4.6}$$

where p^i is the slope for the i'th segment and w^i the fluctuation due to the finiteness of the segment. We cannot take large segments, that reduce this error, because we are limited by the nonstationarity. The errors w^i are essentially uncorrelated over different segments and we have to construct a filter that will predict p^i from the estimates $\hat{p}^i$ by removing the noise w^i, to the extent possible. We assume that p^i is itself a stationary stochastic process, independent of the Brownian motion that generates the fractional Brownian motion. The p^i are the intrinsic variation of the power in the locally stationary fractional Brownian motion. The correlation length of the p^i must be longer than the segments used in estimating them in order to have approximate stationarity relative to the segmentation.

It can be shown using the results from Section 6 that w^i is close to being a white process. This is important because it allows us to estimate the autocovariance of p^i from the estimates $\hat{p}^i$, using the variogram, for example. Given estimates of the autocovariance of p^i and the variance of w^i we can design a minimum mean square error predictor or smoothing filter for $\hat{p}^i$.

The minimum variance unbiased filter for prediction of the parameter processes is as follows. We describe it in the context of the slope parameter, but the filtering of the log intercept is completely analogous. Let $\hat{\mathbf{P}} = (\hat{p}_i)$ be the vector of estimates, $\mathbf{P} = (p_i)$ the realization of the slope process and $\bar{\mathbf{P}} = (\bar{p})$ the constant mean. Then the filter Γ is a matrix that transforms $\hat{\mathbf{P}}$ into $\Gamma\hat{\mathbf{P}}$ in such a way that

$$E[||\Gamma\hat{\mathbf{P}} - \mathbf{P}||^2]$$

is minimized over all matrices Γ that also preserve the mean $\bar{\mathbf{P}}$ of $\mathbf{P}$, that is $\Gamma\bar{\mathbf{P}} = \bar{\mathbf{P}}$. Let C_p be the covariance matrix of $\mathbf{P}$. Then it easily follows that

$$\Gamma \quad = \quad (C_p + C_w)^{-1}[C_p + \mathbf{u}^T \otimes \bar{\mathbf{P}}]$$

where the vector $\mathbf{u} = (u_i)$ is given by

$$u_i \quad = \quad \frac{\bar{\mathbf{P}}_i - \bar{\mathbf{P}}^T(C_p + C_w)^{-1}C_{p,i}}{\bar{\mathbf{P}}^T(C_p + C_w)^{-1}\bar{\mathbf{P}}}$$

and C_w is the diagonal covariance matrix of the estimation errors w^i. Here $C_{p,i}$ is the $i-$th column of the matrix C_p and the superscript T stands for transpose. The slope and log intercept processes are filtered separately. Filtering of this kind is discussed in [3, 29], for example.

Since the effect of the sample noise w^i in (4.6) is largely removed by the filtering procedure, the estimates of the parameter processes will be essentially independent of the prior choice of segmentation.

5 Application to aerothermal data

5.1 Estimation of inertial range

We estimate the set of scales where the process can be modeled as a 'power law'. Figure 3.3 shows that it varies with location. That is, the set of scales where the scale spectrum is approximately linear is location dependent. There are two main reasons for this. First, the scale range where the physical process has power law spectrum varies and depends on the local intensity of the turbulence. Second, the set of scales that are affected by measurement noise varies depending also on the intensity of the turbulence.

We use the scheme described in Section 4.2 to estimate the inertial range. First, we choose a segmentation that is short relative to the prior estimate of the interval of stationarity. We choose segments of length 2^{15} points. Then we apply the algorithm of Section 4.2. For all scale ranges we measure the difference between the scale spectrum computed from the data and the fitted power law and choose the estimated scale range as the largest one for which this measure is within the 90th percentile of the corresponding measure for a realization of fractional Brownian motion.

The estimated 'effective inertial ranges' are shown in Figure 5.1 by the vertical solid lines. For each segment there is one vertical line showing the scales included in the corresponding effective inertial range. The segments consist of 2^{15} points and the maximum scale considered is therefore 14. Note that the more energetic second half of the data displays a more consistent power law behavior.

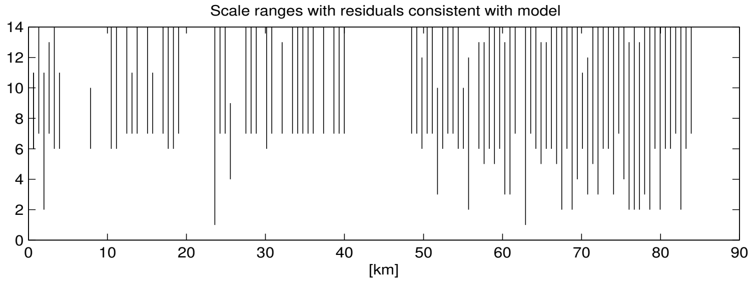

Figure 5.1: The vertical lines in the figure are the estimates of the location dependent inertial ranges. That is, the scales over which the observed process is approximately a power law. For each segment of length 2^{15} points there is one vertical line showing the scales included in the inertial range estimate.

The spectra computed with the estimated location dependent inertial ranges are shown in Figure 5.2, the top plot. They are plotted in a log-log format and shifted vertically to coincide at a center scale. We include only spectra that have at least 5 points. Note the fluctuations in the computed slopes. In Figure 5.2, bottom plot, we show the average of the spectra. The averaging is carried out *before* the log transformation. Note the excellent match with the Kolmogorov scaling law shown by the solid line. There is a slight deviation at the $5th$ scale. This corresponds approximately to the onset of the measurement noise caused by the airplane, see Figure 3.3.

5.2 Estimation of power law parameters

In the previous Section we estimated the inertial range for the scale spectrum, relative to a fixed segmentation. We will now estimate the power law parameters over this set of constrained scales. We use the method described in Section 4.3 for this purpose. Our main objective is to capture the intrinsic variation in the parameters. If we choose a coarse segmentation the variability of the estimates will be small. However, the estimated power law based on the scale spectrum may in this case be an average of power law parameters that vary within the section. This is illustrated in Figure 5.2. The average power law over the whole data set is very close to the Kolmogorov

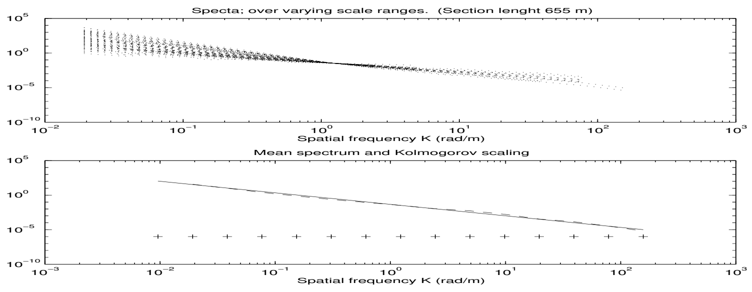

Figure 5.2: The top plot shows the computed scale spectra in each subsection, dotted lines. The estimates are based on the scale ranges shown in Figure 5.1. The bottom plot are the average of the (untransformed) spectra. Note the good match with the Kolmogorov scaling law shown by the solid line.

scaling law. However, within each section (see top plot) such a model must clearly be rejected. Therefore we must choose a segmentation that is not much larger than the interval of stationarity. The estimates should be independent of segmentation. That is, if we shorten the segments this should not lead to an increase in the variability of the estimated parameters.

To show that this is achieved by filtering we choose three segmentation lengths in addition to the one used in the previous section. We choose the following segmentations

$$160m\ (2^{13}\text{ points}),\ 327m\ (2^{14}),\ 655m, (2^{15}),\ 1.31km\ (2^{16}).$$

There are 512, 256, 128, 64 nonoverlapping segments, respectively, in each case. For each segmentation we estimate inertial ranges for the scale spectrum as above. In Figure 5.1 we show the estimated ranges corresponding to segments of length $655m$.

We use the linear least squares regression

$$\log S_j^i \approx c^i + p^i \log_2(\frac{2^j}{50})$$

for each segment i, with j the scale in suitable units. The results of the estimation are shown in Figure 5.3. If the inertial range has fewer than 5 spectral points we do not estimate power law parameters and leave a gap in the figure.

We see from the figure that the estimated slopes, p, and the log intercepts, c, vary considerably over the 80 kilometer data set. They also depend on the segmentation, with the finer ones having larger fluctuations (dotted lines). Note also that the parameter estimates corresponding to the coarsest segmentation differ qualitatively from the others. This segment length is so long that intrinsic variation in the parameters are not captured appropriately. The difference between the parameter estimates for the finest segmentations is mainly due to a white noise estimation residual. For the log intercept process this residual is small and we have essentially obtained what we sought, a parameter estimate that is stable with respect to a shortening of the segmentation. However, for the slope estimate this residual is large. This leads us to the second step of the estimation. In this step we carry out a filtering procedure in order to remove the sampling variability that is segmentation dependent and particularly strong for the slope process p. We use the filtering described in Section 4.3. From the first step we obtained the estimates (in the case of the slopes)

$$\hat{p}^i \quad = \quad p^i + w^i \ .$$

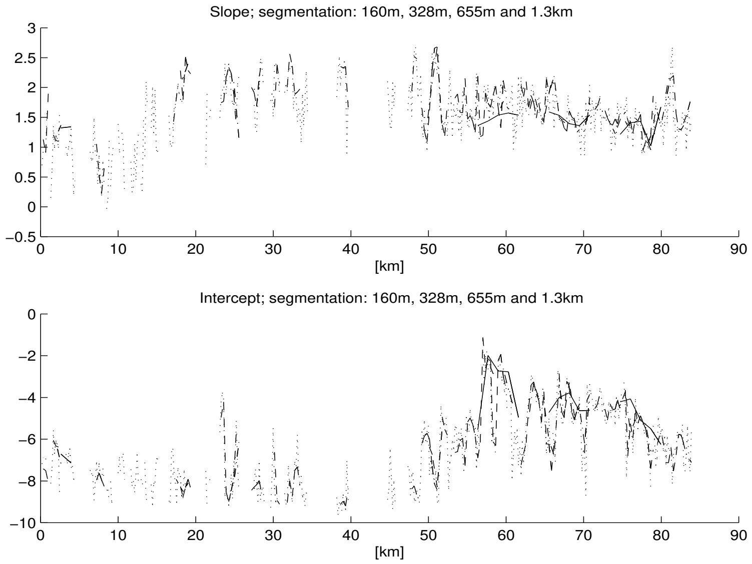

Figure 5.3: Top figure: slopes of log-log scale spectra from wavelet decompositions based on four different segment lengths. At the finest resolution the scale spectra are calculated over nonoverlapping segments 160m long. This is the dotted line that has the largest variations. At the next resolution the nonoverlapping segments are 327m long (plotted with a dashed line). The other two resolutions are 655m and 1.31km and are shown with dash-dotted and solid lines, respectively. The bottom figure is the same as the top but now for the log intercept of the scale spectra. The gaps correspond to inertial ranges with fewer than five spectral points.

The filter computes the minimum variance unbiased predictor of p^i given $\hat{p}^i$.

To construct the minimum variance unbiased filter we need the some additional estimates that we get from the process $\hat{p}^i$. We assume that the slope process is exponentially correlated. We need, therefore, its correlation length l_p and its variance σ_p^2

$$E[(p_i - \bar{p})(p_j - \bar{p})] \quad = \quad \sigma_p^2 \exp(-L|i-j|/l_p)$$

with L the length of the segments. Note that this model is intrinsic to the process and does not depend on the segmentation. As discussed above we shall model the sample noise process $\{w_i\}$ as white and we only need its level σ_w^2. We estimate the parameters σ_p^2, l_p and σ_w^2 using the empirical variogram.

For a time series $\mathbf{X} = (X_i)$ of size N the empirical variogram with lag j is defined by

$$V(j) = \frac{1}{2(N-j)} \sum_{k=1}^{N-j} (X_{k+j} - X_k)^2 \tag{5.1}$$

with dependence on the length of the data vector N not shown. Note that the variogram is essentially unaffected by a relatively slow variation in the mean $\bar{\mathbf{P}}$ of the process. The mean of the empirical variogram for the slope process $\hat{\mathbf{P}}$ is

$$E[V(j)] \quad = \quad \sigma_p^2(1 - \exp(-L|j|/l_p) + \sigma_w^2$$

and from this we obtain the parameter estimates by fitting to the empirical variogram. In particular, σ_w^2 is given by the vertical intercept since the process $\{w_i\}$ is white.

The estimated parameters are as follows:

- **For the slope**: Vertical intercept $\sigma_w^2 = 0.07$, horizontal asymptote level $\sigma_w^2 + \sigma_p^2 = 0.16$, and correlation length $800m$ (five segment lengths).
- **For the log intercept**: Vertical intercept $\tilde{\sigma}_w^2 = 0.1$, horizontal asymptote level $\tilde{\sigma}_w^2 + \tilde{\sigma}_c^2 = 3.0$, and correlation length $800m$ (five segment lengths) are the corresponding parameters.

These estimates are obtained using the finest segmentation, $160m$, and only the data in the last half section, the high turbulence section. The estimates were based on slope and log intercept processes obtained by regression as above, but for the *fixed* set of scales 7 to 12. In this case the magnitude of the estimation error does not depent on the segment. We show the correlation structure in Figure 5.4. The solid line is the fitted model. There are two parts to it, the sampling noise part (the intercept) and the exponential part corresponding to intrinsic variation in the parameters. The intercept part is seen to match well with that obtained from the asymptotic theory which predicts its value to be 0.07 for the slope and 0.05 for the intercept.

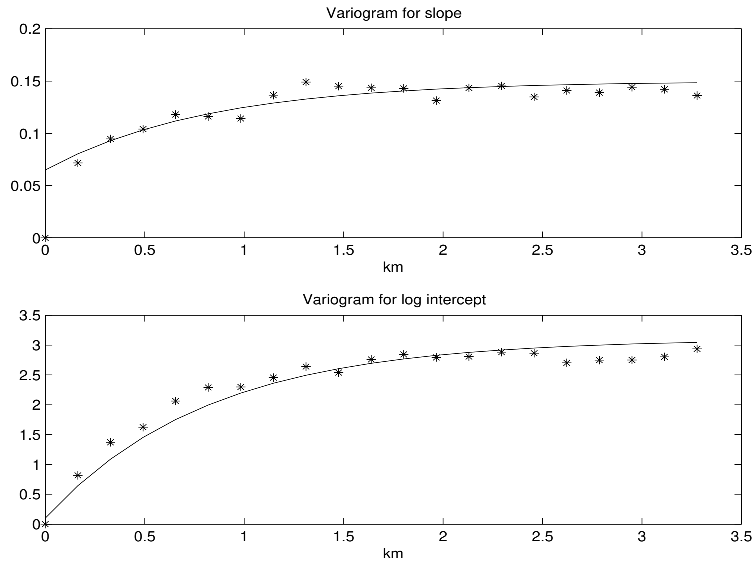

Figure 5.4: Variograms for the fluctuations of the slope (top) and log intercept (bottom) processes obtained from the temperature data. The stars are the variograms. The solid lines are fitted exponentials. We use slope and log intercept estimates from the data with the finest, $160m$, resolution.

Using the estimated model parameters we carry out the filtering to get a better estimate for the slope and log intercept processes. In Figure 5.5 we show these filtered slope and intercept processes. Note that after the filtering all three segmentations give essentially the same result; as expected. We do not include the coarsest segmentation since its segment length is long relative to the interval of stationarity and intrinsic variations in the parameters are being smoothed out.

5.3 Synthetics

We next simulate synthetic temperature profiles using the parameters estimated above. Our objective is to examine the modeling assumptions and also the accuracy of the estimates.

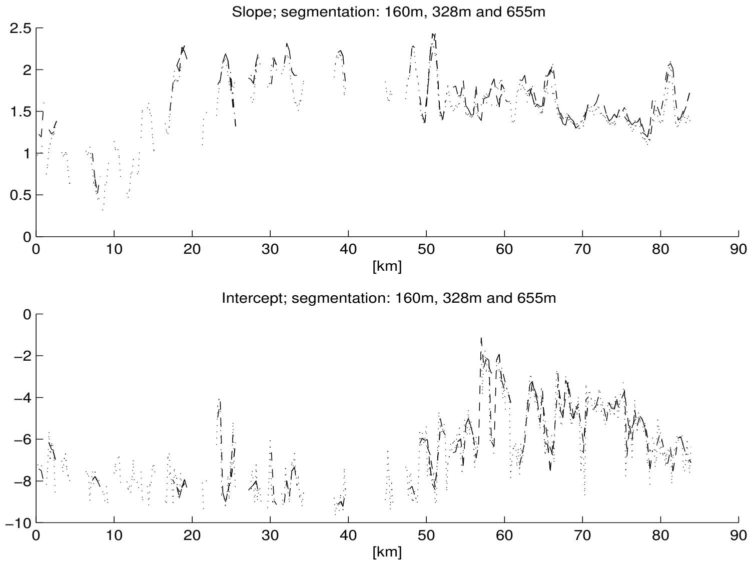

Figure 5.5: Filtered slope (top) and log intercept (bottom) processes of the actual data for the three finest segmentations. Note that the filtered processes for all three segmentations are essentially the same. The dotted line corresponds to the two finest resolutions. The dash-dotted and dashed curves correspond to the coarsest segmentations. The filtering has eliminated the differences in the slope and log intercept processes that are due to the segmentation (as in Figure 5.3) by removing the white noise component of the slope and log intercept fluctuations that is due to sampling. Note the closeness of the slope estimate to the Kolmogorov value 1.66 in the energetic, second half of the data. Note also the increased value of the log intercept, and its fluctuations, in the energetic second half. This is what is expected from physical considerations [32].

We can simulate or generate synthetic realizations in two different ways. First, the realizations can be *conditioned* on the estimated slope and log intercept processes, shown in Figure 5.5. Second, we can simulate from the full model, that is with slope and log intercept processes drawn from the stochastic model for the slope and log intercept. In Figure 5.6 the top plot corresponds to conditional simulation and the bottom two plots to two realizations of the full simulation.

The realization in the top plot of Figure 5.6 has similar features as the data shown in Figure 3.1. However, the noise in the measurements makes a direct comparison difficult. Recall that we only aim at modeling the data on a subset of scales, the scales corresponding to the effective inertial range. A better way to compare the processes is when both have been *scale* filtered so that only the coefficients in the Haar wavelet basis corresponding to these scales contribute. Figure 5.7 shows such a comparison. The similarity between the processes is now striking. The regions without a significant set of scales in the inertial range correspond to the intervals with no data shown in the bottom plot. By comparison with the top plot we see that these sections in particular corresponds to low intensity of the heterogeneities. In Figure 5.8 we show a scatter between the number of scales in the estimated effective inertial range and the estimated intensity that further illustrates this. The sections with less than 3 scales in the calculated inertial range are lumped together and plotted in the leftmost column. The parameters for these where approximated by linear regression over scales 7 to 12.

Consider the simulated data in Figure 5.6. We can analyze them in the same way as the

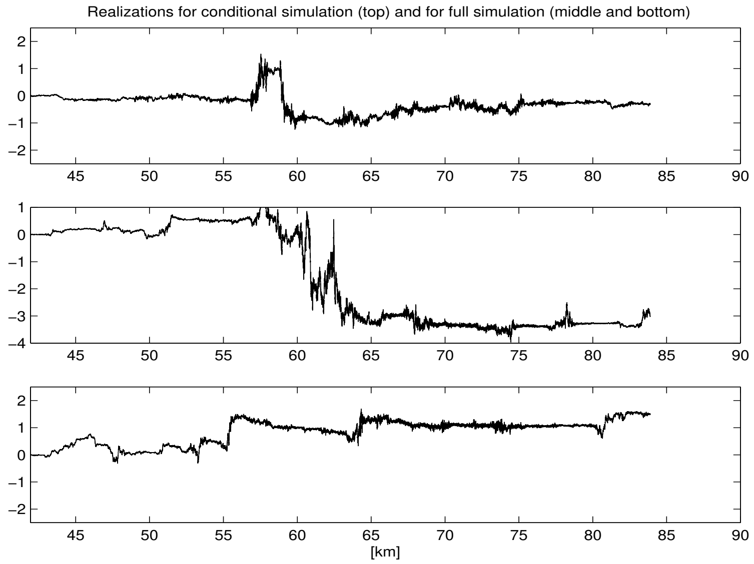

Figure 5.6: Three different realizations of simulated temperature data over the range 40 to 85 km, which is the more energetic range. The top figure is simulated local fractional Brownian motion with slope and log intercept given by the filtered estimates. The two bottom figures are local fractional Brownian motion simulations with slope and log intercept taken to be conditionally independent of the undelying fm. The slope and the log intercept have jontly stationary Gaussian fluctuations with exponential correlation. We use an estimated slowly varying background obtained from a three level matching pursuit [15].

measured data and get estimates of the slope and log intercepts of their scale spectra. They match closely the parameters used to simulate them. We assess the accuracy of the parameter estimates by simulation. We found that the standard deviation between the specified and estimated parameters for the local slope and log-intercept estimates to be .15 and .3 respectively.

6 On estimation of fractional Brownian motion

6.1 Statistics of wavelet coefficients for fBm

The scale spectrum is based on the wavelet coefficients. We now consider the statistics of the Haar wavelet coefficients for fBm. Next we will analyze the scale spectrum and power law parameter estimates based on these statistics. We assume that the data given to us is the wavelet approximation coefficients at level zero. That is, we assume that

$$a_0(n) \;=\; \int_{n-1}^{n} B_H(x)dx \; . \tag{6.1}$$

The wavelet coefficients are therefore normally distributed random variables with

$$\begin{aligned} E[d_j(n)] &= 0 \\ Var[d_j(n)] &= \sigma^2 \frac{(1-2^{-2H})}{(2H+2)(2H+1)} 2^{j(2H+1)} \end{aligned}$$

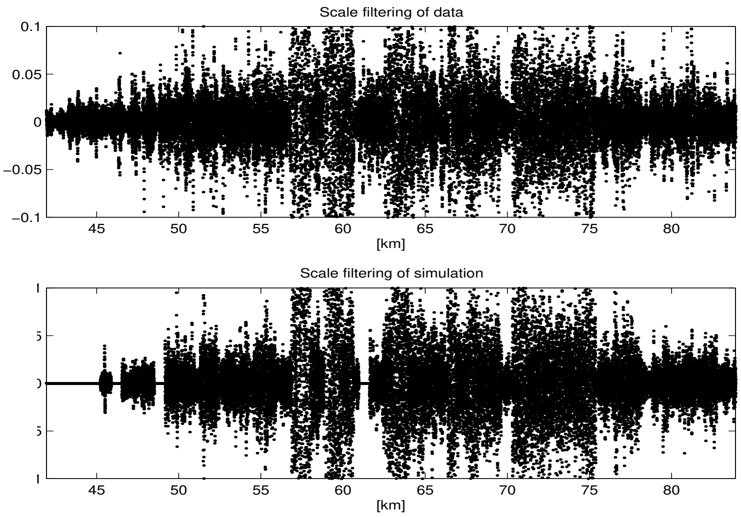

Figure 5.7: The *actual* temperature data set (top figure) obtained by synthesizing it from the detail coefficients in the range 6 to 11 of its Haar decomposition, corresponding to about 1 to 40 meters. Information from this range of scales enters into the estimates of the slope and log intercept processes. The bottom figure is the same as the top but for the *simulated* data. Note the striking similarity between the two figures.

$$Cor[d_{j_1}(n_1)d_{j_2}(n_2)] \;=\; \frac{1}{8(2^{2H}-1)}|D/\sqrt{l}|^{2H+2}\{\delta^2_{l/D}\delta^2_{1/D}|x|^{2H+2}\}|_{x=1}\ . \tag{6.2}$$

Here

$$l = 2^{j_2-j_1}, \quad j_1 \leq j_2 \tag{6.3}$$

is the relative scale,

$$D \;=\; |(2n_2-1)l-(2n_1-1)| \tag{6.4}$$

is the relative location and

$$\delta^2_d f(x) = f(x+d) - 2f(x) + f(x-d) \tag{6.5}$$

is the second order symmetric difference. These results are derived using (2.1) and self-similarity of fBm. Note that the expression for the correlation of the wavelet coefficients takes on a universal form depending only on the relative displacement in space and scale. The derivation of the precision of the power law estimators in Section 6.3 is based on (6.2), which was only known in special cases before [7, 34].

Consider the correlation of wavelet-coefficients, when

$$\frac{D}{l} \;=\; |2^{j_2}(2n_2-1) - 2^{j_1}(2n_1-1)|2^{-j_2} \;\rightarrow\; +\infty$$

then

$$\frac{\delta^2_{l/D}}{(l/D)^2}\frac{\delta^2_{1/D}}{(1/D)^2}$$

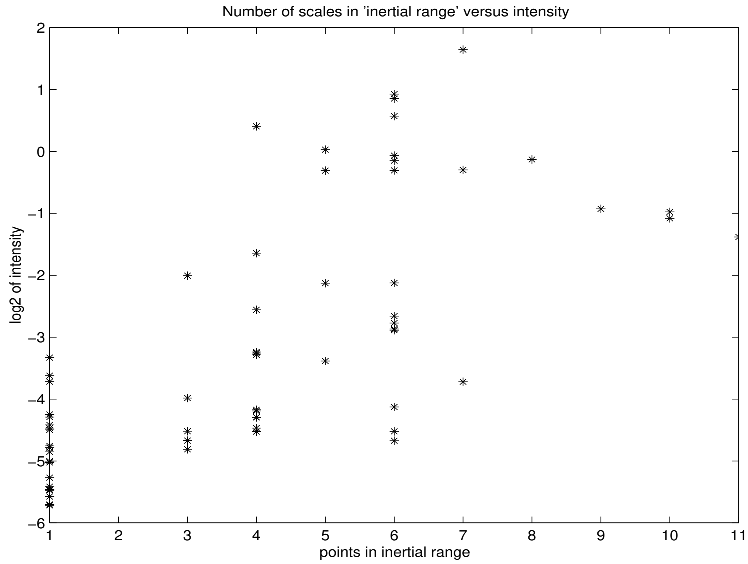

Figure 5.8: Scatter plot between number of scales in the inertial range and the intensity. The plot illustrates that the high intensity turbulence is associated with a relative broader inertial range.

is a forth order central difference operator and its expansion gives the asymptotic result

$$Cor[d_{j_1}^{n_1} d_{j_2}^{n_2}] \;=\; \frac{(H+1)(2H+1)H(2H-1)}{2(2^{2H}-1)} \left[\left(\frac{\sqrt{l}}{D}\right)^{2-2H} + O\left(\frac{\sqrt{l}}{D}\right)^{4-2H} \right].$$

If the scales j_1 and j_2 are given, then for $n_2 \to \infty$ and n_1 fixed we have that $\frac{\sqrt{l}}{D} \approx \frac{1}{2n_2\sqrt{l}}$, whereas for $n_1 \to \infty$ and n_2 fixed we have $\frac{\sqrt{l}}{D} \approx \frac{\sqrt{l}}{2n_1}$. The correlation within a single scale is in particular

$$Cor[d_j^{n_1} d_j^{n_2}] \;=\; \frac{(H+1)(2H+1)H(2H-1)}{2(2^{2H}-1)} \left[\left(\frac{1}{2|n_2-n_1|}\right)^{2-2H} + O\left(\frac{1}{|n_2-n_1|}\right)^{4-2H} \right] \tag{6.6}$$

For general wavelets the following asymptotic result is presented in [7, 34]

$$\begin{aligned} E[d_{j_1}^{n_1} d_{j_2}^{n_2}] &\sim O(|2^{j_2}n_2 - 2^{j_1}n_1|^{2(H-R)}) \\ &\text{for } \; |2^{j_2}n_2 - 2^{j_1}n_1| \;\to\; +\infty \end{aligned}$$

with R being the number of vanishing moments for the wavelet ($R = 1$ for Haar wavelets).

6.2 Statistics of the scale spectrum

Using the above results for the wavelet coefficients we get a characterization of the statistics of the scale spectral points. The mean is

$$\begin{aligned} E[S_j] \;&=\; E[d_j(n)^2] \;=\; \sigma^2 h(H) 2^{j(2H+1)} \\ &\propto \int_{-\infty}^{\infty} \sigma^2 |f|^{-(2H+1)} \frac{\sin^4(\pi 2^{j-1} f)}{(\pi 2^{j-1} f)^2} 2^j df \end{aligned} \tag{6.7}$$

where

$$h(H) = \frac{(1-2^{-2H})}{(2H+2)(2H+1)} \ . \tag{6.8}$$

The normalized variance is

$$\frac{Var[S_j]}{E[S_j]^2} \;=\; 2\frac{\sum_{n,m}(C^j_{nm})^2}{N_j^2 E[S_j]^2} \;=\; \frac{1}{N_j}\left\{\frac{4}{N_j}\sum_{k=0}^{N_j-1}(N_j-k)\rho_H^2(k) - 2\right\} \tag{6.9}$$

with $N_j = 2^{M-j}$, the number of detail coefficients at level j and 2^M the total length of the data. Note that here and in the sequel we suppress the dependence on M. The matrix $C^j = (C^j_{nm})$ in (6.9) is the covariance matrix of the wavelet coefficients at scale j, determined by (6.2), and $\rho_H(k)$ is the corresponding correlation coefficient, which depends only on $k = |n_1 - n_2|$ and H. Denote

$$g(H) \;=\; \lim_{N_j\to\infty}\left\{\frac{4}{N_j}\sum_{k=0}^{N_j-1}(N_j-k)\rho_H^2(k) - 2\right\} \tag{6.10}$$

then

$$\frac{Var[S_j]}{E[S_j]^2} \;\approx\; \frac{g(H)}{N_j} \tag{6.11}$$

for N_j large. The function $g(H)$ is computed numerically and is shown in Figure 6.1. It depends only weakly on the value of H, especially when $H \le 1/2$ which is the case of interest to us. The asymptotic behavior (6.11) holds only when the correlation $\rho_H^2(k)$ decays to zero in an integrable way. This means that the Hurst exponent must be restricted to $H < 3/4$. In the Kolmogorov case $H = 1/3$, the correlation ρ_H decays like $k^{-4/3}$.

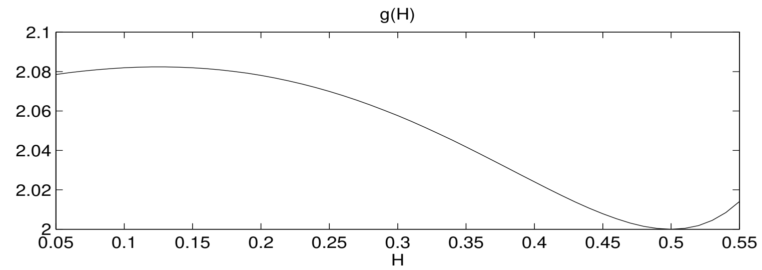

Figure 6.1: Plot of $g(H)$ showing how the relative variance depends on H.

Similarly, if $H < 3/4$, the normalized covariance can be expressed, for *fixed* scale separation, as

$$D_{j_1 j_2} \;\equiv\; \frac{Cov[S_{j_1}S_{j_2}]}{E[S_{j_1}]E[S_{j_2}]} \;=\; 2\frac{\sum_{n,m}(C^{j_1,j_2}_{nm})^2}{N_{j_1}N_{j_2}E[S_{j_1}]E[S_{j_2}]} \tag{6.12}$$

with C^{j_1,j_2} being the cross covariance between the wavelet coefficients at scale j_1 and j_2, determined by (6.2). The scale spectral points S_j are not uncorrelated for different j, as is sometimes assumed. For the power law parameter estimation we note that for fixed j_1 and j_2

$$D_{j_1 j_2} \;\equiv\; \frac{Cov[S_{j_1}, S_{j_2}]}{E[S_{j_1}]E[S_{j_2}]} \sim \frac{1}{\sqrt{N_{j_1}N_{j_2}}} \tag{6.13}$$

as N_{j_1} and N_{j_2} go to infinity. These results can be derived using (6.2).

6.3 Fluctuation theory for the scale spectra

In the least squares fit of the power law we will need the statistical properties of the logarithm of the scale spectra. We summarize the relevant facts in this section, with the details given in Appendix A.

In the large N_j limit, the *distribution* of the scale spectral points is normal

$$\begin{aligned} S_j &= \bar{S}_j(1+\frac{v_j}{\sqrt{N_j}}) \\ v_j &\sim \mathcal{N}(0,g(H)) \text{ as } N_j \to \infty \end{aligned}$$

with $\bar{S}_j \equiv E[S_j] = \sigma^2 h(H) 2^{j(2H+1)}$, and where h and g are defined by (6.10) and (6.8), respectively, and $\mathcal{N}(\mu, s^2)$ is the normal distribution with mean μ and variance s^2. The derivation of these results requires that $H \leq 1/2$. The covariance of the fluctuations is

$$Cov[\frac{v_j}{\sqrt{N_j}}, \frac{v_i}{\sqrt{N_i}}] = D_{ji} \tag{6.14}$$

with D_{ji} given by (6.13).

The fluctuations in S_j are small and in Appendix A we show that this leads to the asymptotic estimate

$$\log_2(S_j) \approx \log_2(\bar{S}_j) + \frac{v_j}{\sqrt{N_j}\ln(2)}. \tag{6.15}$$

In addition to showing that v_j is asymptotically normal, we also show in Appendix A that for any $j_1 < j_2$ fixed the random variables v_j, $j_1 \leq j \leq j_2$ are asymptotically jointly normal, as $N_j \to \infty$.

6.4 Estimators for the power law

Write (6.15) as

$$\begin{aligned} \log_2(S_j) &\approx \log_2(\sigma^2 h(H)) + j(2H+1) + \frac{v_j}{\sqrt{N_j}\ln(2)} \qquad (6.16) \\ &= c + jp + \frac{v_j}{\sqrt{N_j}\ln(2)} \quad j_1 \leq j \leq j_2 \end{aligned}$$

with the slope p and log intercept c to be estimated from data, and $v_j/\sqrt{N_j}$ a fluctuation term that is characterized in the large N_j limit by the central limit theorem, as discussed in the previous section. In view of the above we can then use generalized least squares to estimate $\mathbf{b} = [c,p]^T$ as

$$\hat{\mathbf{b}} = (X^T D^{-1} X)^{-1} X^T D^{-1} \mathbf{Y} \tag{6.17}$$

with $\mathbf{Y} = [\log_2(S_{j_1}), \cdots, \log_2(S_{j_2})]^T$, and where $j_1, \cdots, j_2$ are the scales in the range under consideration, the inertial range. The dependence on j_1, j_2 & M has been suppressed. For a discussion of generalized least squares see [12]. The design matrix X is defined as

$$X = \begin{bmatrix} 1 & j_1 \\ 1 & j_1+1 \\ \vdots & \vdots \\ 1 & j_2 \end{bmatrix}.$$

The matrix D is the normalized covariance matrix of the spectral points defined in (6.12). This matrix depends on the value of H. However this dependence is weak so we can use some rough estimate of H that is usually available, like $H = 1/3$ for the aerothermal data of Section 5. In view of (6.16) we find that the estimates of $\mathbf{b}$ have means

$$\begin{aligned} E[\hat{c}] &= \log_2(\sigma^2 h(H)) \\ E[\hat{p}] &= 2H+1 \end{aligned} \tag{6.18}$$

and covariance

$$\begin{aligned} C_{\mathbf{b}} &= (X^T D^{-1} X \ln(2)^2)^{-1} \\ &= \frac{\ln(2)^{-2}}{[\sum\sum D_{ij}^{-1} \sum\sum D_{ij}^{-1}\ i\, j - (\sum\sum D_{ij}^{-1})^2]} \begin{bmatrix} \sum\sum D_{ij}^{-1} & -\sum\sum D_{ij}^{-1}\ i \\ -\sum\sum D_{ij}^{-1}\ i & \sum\sum D_{ij}^{-1}\ i\, j \end{bmatrix} \sim \frac{1}{N_{j_1}} \end{aligned} \tag{6.19}$$

for large N_{j_1}. ¿From the above we arrive at the sought after parameter estimates

$$\hat{H} = (\hat{p}-1)/2 \tag{6.20}$$

$$\widehat{\log_2(\sigma^2)} = \hat{c} - \log_2(h(\hat{H}))\ . \tag{6.21}$$

We next analyze the precision of these estimators. We find that the estimates $\hat{H}$ and $\widehat{\log_2(\sigma^2)}$ are normally distributed with variances

$$Var[\hat{H}] \approx C_{\mathbf{b}}(2,2)/4 \tag{6.22}$$

$$Var[\widehat{\log_2(\sigma^2)}] \approx C_{\mathbf{b}}(1,1) - 2C_{\mathbf{b}}(1,2)\ \log_2(h(H))' + C_{\mathbf{b}}(2,2)\ [\log_2(h(H))']^2 \tag{6.23}$$

for N_{j_1} large with $C_{\mathbf{b}}$ defined in (6.19) and h in (6.8). The variance of the estimators is of order $1/N_{j_1}$.

Long-memory processes can also be modeled via state space models. Maximum likelihood estimators (MLE) for such models are analyzed in [4, 8].

6.5 Illustration of precision

The above result on the distribution of the estimators may be validated by numerical simulation. We generate synthetic realizations of fractional Brownian motion with known parameters. Then we use the above algorithm to estimate these parameters and compare the precision of the estimates with the predicted precision. Actually, we do not simulate fractional Brownian motion but rather the observed sequence (6.1), the approximation coefficients at level zero. The synthetic realizations are generated by the simulation algorithm described in Appendix B, which is a fast, $O(N)$ method, with N the length of the data.

In Figure 6.2 we compare the asymptotic law of the estimators with simulations. It is clear that the theoretical normal distribution predicts accurately the distribution of the estimators.

We show how the variance of the slope estimate $\hat{p}$ depends on the number of scales in the linear fit and the total number of data-points in Figure 6.3. The dotted and solid lines correspond to using respectively scales $1, 2, 3$ and scales $1, 2, 3, 4, 5$ and Hurst parameter $H = 1/3$. The crosses give the Cramer-Rao lower bound [2, 38]: $1/(2^{M+2} \ln^2(2))$ for the case that the wavelet coefficients are uncorrelated. It is seen that only a few scales are needed for the variance of the slope estimator to be close to the bound.

Note that our objective is estimation of H when $H \approx 1/3$ and that our asymptotic result is valid only when $H < 1/2$. Several authors [22] have observed that slope estimators degrade when H become large ($H \approx 1$). This is because in that case the wavelet coefficients exhibit long range correlations.

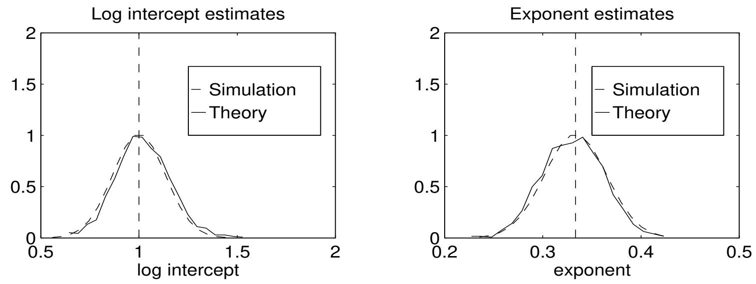

Figure 6.2: Distribution of power law parameter estimates. The parameters were estimated based on realizations of length 2^{10} with $H = 1/3$ and $\sigma^2 = 2$. The power law fitting is done over scales $1 \cdots 5$. The solid line is the empirical distribution associated with 1000 realizations, the dashed line is the theoretical distribution and the vertical lines are the specified parameters.

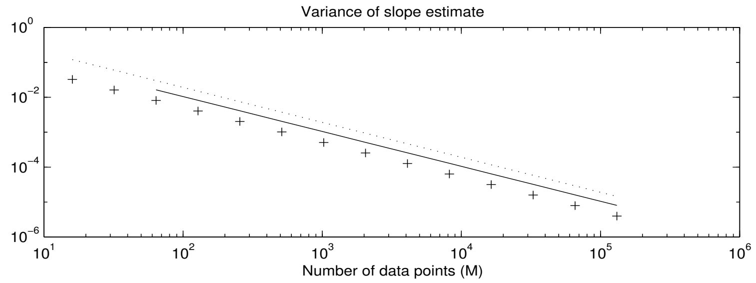

Figure 6.3: The variance of the slope estimate, computed by (6.22) with $H = 1/3$, is plotted as a function of the number of data points 2^M. The dotted line is for an estimate with three scales, the solid line with five scales. The crosses correspond to the Cramer-Rao lower bound. Note that even when the number of data points is not very large the estimate of the slope does not depend sensitively on the number of scales used.

In the application of the estimation of the power law parameters to the aerothermal data in Section 5, which we model by a local fractional Brownian motion, we have to account for spurious noise effects. This noise affects the small scales mostly. We can make the power law parameter estimates more robust by letting the weight matrix D in (6.17) depend on an additive noise parameter which must be adjusted appropriately from rough prior data analysis. This is done in Section 5.2 without a detailed discussion.

7 Summary and conclusions

We have analyzed in detail estimation of fractional Brownian motion based on the scale spectrum. Moreover, we have generalized the estimation procedure to a local power law process. For such processes the power law itself (the exponent or slope) and the multiplicative constant (log intercept of the scale spectrum) are not constants but vary slowly. We estimate the slope and log intercept of the scale spectrum by appropriately segmenting the data and then removing segmentation effects by a filtering procedure. The slope and log intercepts themselves are modeled as stochastic processes.

The estimation of the power law includes the identification of a location dependent inertial range, that is, the scale range where the process can be modeled as a power law.

We applied the estimation procedure to an important set of aerothermal data. We found that the *average* of the estimated slope process is close to the value predicted by the Kolmogorov scaling theory. To get a faithful representation of the heterogeneity in the data it is however important to incorporate the local fluctuations around this value.

We generate synthetic temperatures in two different ways. First, the local power law process is simulated conditioned on the particular parameters estimated from the data (slope and log intercept). Then we get a faithful replication of the temperature data. Second, the parameters are taken as random, that is, sampled from their model. Then the simulation is sample-wise very different from the actual data, but the statistics are faithfully reproduced. Such realizations can be used to define synthetic media for wave propagation codes. This is the application that motivated our investigation. A complete study of the effects of the heterogeneity in local power law model on wave propagation has yet to be carried out.

An important aspect of the model that we use is separation of scales in the variation of the estimated parameters (slope and log intercept) from the scale at which we sample the process.

A Central limit theorem for scale spectra

In this appendix we derive a central limit theorem for the scale spectral points, S_j, and also the log transformed spectral points $\log_2(S_j)$, where

$$S_j = \frac{1}{N_j}\sum_{n=1}^{N_j}(d_j(n))^2$$

with $(d_j(n))_{n=1}^{N_j}$ the Haar wavelet coefficients of fractional Brownian motion at scale j with $H \leq 1/2$. The total number of data points is 2^M and $N_j = 2^{M-j}$. We will consider the limit $N_j \to \infty$. This means that M is large and that j is not too close to M, that is, the central limit theorem will not be valid for the large-scale spectral points S_j with $j \approx M$.

We will use the following version of the Berry-Esseen theorem [6].

Let the y_i be independent random variables such that

$$E[y_i] = 0,\ E[y_i^2] = \sigma_i^2,\ E[|y_i^3|] = \rho_i$$

and define

$$s_n^2 = \sum_{i=1}^{n}\sigma_i^2,\ r_n = \sum_{i=1}^{n}\rho_i\ .$$

Let F_n be the distribution of the normalized sum $\sum_{i=1}^{n} y_i/s_n$.

Then for all x and n

$$|F_n(x) - \mathsf{N}(x)| \quad \leq \quad 6\frac{r_n}{s_n^3} \tag{A.1}$$

where N stands for the mean zero, unit variance normal distribution.

A.1 Central limit theorem for S_j

To use this theorem we will first transform the sum of squares of wavelet coefficients to a sum of independent random variables. Denote the vector of wavelet coefficients at scale j by $d^j = [d_j(1), \cdots, d_j(N_j)]^T$ and the covariance matrix associated with d^j by C^j. Then d^j has the same law as

$$(C^j)^{1/2}[\eta_1, \cdots, \eta_{N_j}]^T$$

and in law

$$S_j \;=\; \frac{\bar{S}_j}{N_j}\sum_{i=1}^{N_j}\lambda_i^j\eta_i^2 \;=\; \bar{S}_j\left(1+\frac{1}{\sqrt{N_j}}[\frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j}\lambda_i^j(\eta_i^2-1)]\right) \tag{A.2}$$

where η_i are independent zero mean and unit variance normal random variables, $\bar{S}_j = E[S_j]$ and λ_i^j are the eigenvalues of $C^j/\bar{S}_j$, that is the correlation matrix of d^j.

Let

$$Y_j \;=\; \frac{1}{\sqrt{N_j}}\sum_{i=1}^{N_j} y_i^j$$

with $y_i^j = \lambda_i^j(\eta_i^2 - 1)$. Note that the y_i^j are independent random variables such that

$$E[y_i^j] = 0,\;\; E[(y_i^j)^2] = 2(\lambda_i^j)^2, \quad E[|(y_i^j)^3|] \leq 28(\lambda_i^j)^3 \;.$$

According to (A.1) we need only show that

$$J_j \;\equiv\; \frac{\sum_{i=1}^{N_j}(\lambda_i^j)^3}{[\sum_{i=1}^{N_j}(\lambda_i^j)^2]^{3/2}} \tag{A.3}$$

is small for N_j large.

Using (6.9) we find that

$$\sum_{i=1}^{N_j}(\lambda_i^j)^2 \;=\; \sum_{n,m}(C_{nm}^j/\bar{S}_j)^2 \;=\; N_j^2 Var[S_j]/(2\bar{S}_j^2) \;\approx\; N_j\; g(H)/2\;. \tag{A.4}$$

Below we show that

$$\lambda_i^j \;\leq\; K \tag{A.5}$$

for some constant $K \geq 1$ independent of N_j. Hence

$$\sum_{i=1}^{N_j}(\lambda_i^j)^3 \;\leq\; N_j K^3\;. \tag{A.6}$$

¿From (A.4) and (A.6) we find that

$$J_j \;\approx\; \frac{(2K^2/g(H))^{3/2}}{\sqrt{N_j}}$$

and goes to zero as $N_j \to \infty$. ¿From the Berry-Esseen theorem we conclude that the distribution of

$$Y_j/\sqrt{g(H)} = \frac{\sum_{i=1}^{N_j} \lambda_i^j (\eta_i^2 - 1)]}{\sqrt{g(H) N_j}}$$

tends to the standard normal distribution as N_j becomes large.

Thus, for N_j large

$$S_j \approx \bar{S}_j (1 + \epsilon_j \sqrt{\frac{g(H)}{N_j}})$$

in law, with ϵ_j a standard normal random variable.

Let us now show that the estimate (A.5) is valid. The diagonal entries of the correlation matrix are all equal to one. By the Gersgorin disc theorem, $|\lambda_i^j - 1| \leq \sum_{n \neq i} |C_{in}|/C_{11}$. But the sum remains finite as $N_j \to \infty$ because we have assumed that $H \leq 1/2$ and so we can use (6.6). Rewriting this inequality with λ_i^j on the left side we get (A.5).

A.2 Central limit theorem for $\log_2(S_j)$

We consider next a central limit theorem for $\log_2[S_j]$. As in (A.2) we write the scale spectrum in terms of centered random variables Y_j

$$S_j = \bar{S}_j (1 + \frac{Y_j}{\sqrt{N_j}}) \ .$$

Taking logs and expanding in a Taylor expansion with remainder we get

$$\log_2(S_j) = \log_2(\bar{S}_j) + \frac{Y_j}{\sqrt{N_j} \ln 2} - \frac{Y_j^2}{2\xi(Y_j)^2 N_j \ln 2}$$

with

$$\xi(Y_j) \in \begin{cases} < 1, 1 + Y_j/\sqrt{N_j} > & \text{when } Y_j \geq 0, \\ < 1 + Y_j/\sqrt{N_j}, 1 > & \text{when } Y_j < 0. \end{cases}$$

We showed above that the central limit theorem holds for Y_j. We show below that $Y_j^2/(2\xi(Y_j)^2)$ is $O(1)$ as $N_j \to \infty$. We therefore, have a central limit theorem for $\log_2(S_j)$ as well.

By the Cauchy-Schwarz inequality

$$J_j \equiv E\left[|\frac{Y_j^2}{\xi(Y_j)^2}| \right] \leq \sqrt{E[\xi(Y_j)^{-4}] E[Y_j^4]} \ .$$

We show that $J_j = O(1)$ as $N_j \to \infty$. Consider first $E[Y_j^4]$. Using (A.5) we find

$$\begin{aligned} E[Y_j^4] &= E[(\frac{\sum_{i=1}^{N_j} \lambda_i^j (\eta_i^2 - 1)]}{\sqrt{N_j}})^4] \\ &\leq K^4 (N_j E[(\eta_i^2 - 1)^4] + N_j (N_j - 1) 6 E^2[(\eta_i^2 - 1)^2])/N_j^2 \ = \ c_1 + c_2/N_j \end{aligned}$$

with c_i constants. Next consider

$$\begin{aligned} I_j &\equiv E[\xi(Y_j)^{-4}] \ \leq \ E[\xi(Y_j)^{-4} I_{(Y_j \leq 0)}] + 1 \\ &\leq E[(1 + Y_j/\sqrt{N_j})^{-4} I_{(Y_j \leq 0)}] + 1 \end{aligned}$$

with I_A the indicator function of the set A. In order to bound I_j we replace $1+Y_j/\sqrt{N_j}$ by Z_j such that w.p.1 $Z_j \leq (1+Y_j/\sqrt{N_j})$.

Note that

$$1+Y_j/\sqrt{N_j} \;=\; \frac{\sum_{i=1}^{N_j} \lambda_i^j \eta_i^2}{N_j}$$

and that $\sum_{i=1}^{N_j} \lambda_i^j = N_j$, which is the trace of the correlation matrix for d^j. Let q be the number of eigenvalues that exceed $1/2$. In view of (A.5) we find

$$q \;\geq\; \tilde{N}_j \;=\; N_j(2K-1)^{-1} \;\equiv\; N_j\tilde{K} \; .$$

Thus, we can define Z_j as

$$Z_j \;=\; \frac{1}{2}\frac{\sum_{i=1}^{\tilde{N}_j} \tilde{\eta}_i^2}{N_j} \;=\; \frac{\tilde{K}}{2}\frac{\sum_{i=1}^{\tilde{N}_j} \tilde{\eta}_i^2}{\tilde{N}_j}$$

Hence

$$I_j \;\leq\; E[Z_j^{-4} I_{(Y_j\leq 0)}]+1 \;\leq\; (2/\tilde{K})^{-4}\tilde{N}_j^4 E[\tilde{Z}^{-4}]+1$$

with

$$\tilde{Z}_j \;=\; \sum_{i=1}^{\tilde{N}_j} \eta_i^2$$

in law. The random variable $\tilde{Z}_j$ has law given by the Gamma density

$$\gamma_{(\alpha,\nu)}(x) \;=\; \frac{(\alpha)^\nu}{\Gamma(\nu)}\, x^{\nu-1}\, \exp(-\alpha x)$$

with parameters $\alpha = 1/2$ and $\nu = \tilde{N}_j/2$. It follows that for $\tilde{N}_j \geq 10$

$$E[\tilde{Z}^{-4}] \;=\; \frac{(\alpha)^\nu \Gamma(\nu-4)}{(\alpha)^{\nu-4}\Gamma(\nu)} \;=\; \frac{1}{(\tilde{N}_j-2)(\tilde{N}_j-4)(\tilde{N}_j-6)(\tilde{N}_j-8)} \; .$$

Hence, $I_j \leq K''$ and

$$J_j \;\leq\; \sqrt{K''(c_1/N_j+c_2)} \;\leq\; \bar{K}$$

for $N_j \geq 10/\tilde{K}$.

¿From the above we conclude that for N_j large

$$\log_2(S_j) \;\approx\; \log_2(\bar{S}_j) + \epsilon_j \sqrt{\frac{g(H)}{N_j \ln(2)}}$$

in law, with ϵ_j a standard normal random variable.

A.3 Joint density of spectral points

We show that the spectral points are asymptotically, as $N_j \to \infty$, *jointly* normal. ¿From this follows that the central limit theorem holds also for the power law parameter estimates. Consider the distribution of $Y = a_1 S_{j_1} + a_2 S_{j_2}$ with $a_1 \neq 0$ & $a_2 \neq 0$ and $j_1 < j_2$ fixed parameters. The central limit theorem for this quantity can be shown essentially as in Section A.1. Let d^{j_1} and d^{j_2} be the vectors of wavelet coefficients at the two scales. Define

$$d = \begin{bmatrix} d^{j_1} \\ d^{j_2} \end{bmatrix}$$

and $\tilde{d} = Dd$ with D a diagonal matrix whose first N_{j_1} elements equal a_1 and last N_{j_2} elements equal $a_2 2^{j_2 - j_1}$. Then

$$Y = \tilde{d}^T \tilde{d} / N_{j_1} \ .$$

By a transformation to independent variables and an application of the Berry-Esseen theorem we find, as in (A.3), that we need to show

$$J \equiv \frac{\sum_{i=1}^{N_{j_1}+N_{j_2}} (\lambda_i)^3}{[\sum_{i=1}^{N_{j_1}+N_{j_2}} (\lambda_i)^2]^{3/2}} \tag{A.7}$$

is small for N_{j_1} large. Here λ_i are the eigenvalues of $Cov(\tilde{d}) = DCD$ with C the covariance matrix of d. Note first that from (6.9) it follows that

$$\sum_{i=1}^{N_{j_1}+N_{j_2}} \lambda_i^2 \geq (a_1)^2 \sum_{n,m} (C^{j_1}_{nm})^2 \approx (a_1)^2 N_{j_1} (\bar{S}_{j_1})^2 g(H)/2 \ .$$

Hence, (A.7) follows if we can bound λ_i uniformly with a bound independent of N_{j_1}. We find, using (6.2) and the Gersgorin disc theorem, that for some constant L and $1 \leq j \leq N_{j_1}$

$$|\lambda_j - a_1 \bar{S}_{j_1}| \leq L \bar{S}_{j_1}$$

since the rows of C are absolutely and uniformly summable with a bound that is independent of N_{j_1}. Hence $\lambda_i \leq \tilde{L} \bar{S}_{j_1}$ for some constant $\tilde{L}$. A similar argument holds for $N_{j_1} < j \leq N_{j_1} + N_{j_2}$. Therefore, (A.7) and the central limit theorem for Y follows.

B Simulation of fractional Brownian motion

In this appendix we show how we generate synthetic realizations of the turbulence data.

Our main objective is a fast method that simulates realizations of the observed process (6.1). Note that the correlation structure of this process is slightly different from that of discrete time sampled fractional Brownian motion. The simulation algorithm must be able to faithfully reproduce correlations, on all scales, without being too expensive computationally as the data record becomes very long. A detailed comparison of simulation methods and of estimators are given in [33, 28].

The main idea behind the simulation algorithm is the following. Pure Brownian motion can be simulated very efficiently, in $O(N)$ steps with N the length of the realization. We will construct a modification of this algorithm that works well for processes whose correlation structure is similar to that of pure Brownian motion. In [23] this concept was discussed in the context of more general processes including wide sense stationary processes and processes defined in two and three spatial dimensions.

We start by a discussion of simulation of pure Brownian motion in Section B.1. In Section B.2 we generalize the method and explain how we constructed the realizations of the *local* power law process. A MATLAB script is included in Section B.3.

B.1 Simulation of Brownian motion

In this section we describe a well known method for generating sample paths of Brownian motion. The method is described in [15], for instance.

The input parameter is m with 2^m being the dimension of the simulation grid. The realization, B, is sampled on the simulation grid $\{x_i = i; 1 \leq i \leq 2^m\}$ conditioned on $B(0) = 0$. The process is normalized so that $Var(B(2^m)) = 1$.

The algorithm proceeds in a top-down fashion, iteratively filling in the finer scale information. The simulation scheme starts out by initializing the boundary nodes. Then proceeds by drawing the samples on dyadic subgrids, creating a process of successively finer resolution. When the grid is refined, the new samples (one at each midpoint between samples at the previous level) need only be conditioned on the *two nearest neighbors*, both being simulated at the previous level.

Let us first describe the initial step. The value $B(2^m)$ is simulated *conditioned* on $B(0) = 0$. Since we use a scaling of the process such that this sample has unit variance we obtain the sample as

$$B(2^m) \quad \leftarrow \quad N(0,1)$$

with $N(0,1)$ being a sample from the zero mean unit variance normal distribution.

Having defined the boundary nodes we start the refinement process. At each refinement step the nodes in the midpoint in between the nodes defined at the previous level are drawn. In the first step $B(2^{m-1})$ is simulated conditioned on the nodes $B(0)$ and $B(2^m)$. Using the correlation structure of Brownian motion we find

$$B(2^{m-1}) \quad \leftarrow \quad N((B(0) + B(2^m))/2, 1/\sqrt{2}).$$

The conditional mean is just the average of the neighbor nodes. The conditional variance at this level is $1/\sqrt{2}$. This completes the first refinement step.

In the next refinement step the nodes $B(2^{m-2})$ and $B(3 * 2^{m-2})$ are simulated. Consider first $B(2^{m-2})$. We should now condition on the nodes simulated so far, that is $B(0)$, $B(2^{m-1})$ and $B(2^m)$. However, for Brownian motion we need only condition on the two nearest neighbors and

$$B(2^{m-2}) \quad \leftarrow \quad N((B(0) + B(2^{m-1}))/2, 2^{-1}).$$

Similarly

$$B(3 * 2^{m-2}) \quad \leftarrow \quad N((B(2^{m-1}) + B(2^m))/2, 2^{-1}).$$

This process terminates after m refinement steps. As in the first steps we always condition only on the nearest neighbors (both being defined at the previous refinement step). The conditional mean is simply the average of these values. At the n'th refinement step the conditional variance is $2^{-n/2}$.

Note two main aspects of the above algorithm. First, in order to simulate a process with long range interactions it is 'filled in' on dyadic subsets of the data. Thus, correlation at different scales can be incorporated in an efficient way. Second, with this simulation pattern the refinement step becomes trivial since each new data point is conditionally independent of all other data points given the values of the two neighbor points.

B.2 Simulation of processes with long range interactions

We describe the generalization of the above simulation algorithm that we use to simulate observed fractional Brownian motion. The main idea behind the generalization is that we use the same refinement pattern when simulating the process. That is, we fill in the process on dyadic subsets. However, for more general correlation structures than pure Brownian motion the above mentioned conditional independence is not valid. We therefore extend the set of points on which we condition during the refinement beyond the two nearest neighbors. With this extended conditioning set the new point will be *approximately* independent of the other, previously simulated, points. The computational cost is therefore of $O(N)$ with N number of data points.

Input parameters are H and m, with H being Hurst exponent and 2^m the dimension of the simulation grid. The realization, X, is sampled on the simulation grid $\{x_i = i; 1 \leq i \leq 2^m\}$ conditioned on $X(0) = 0$. The process is normalized so that $Var(X(2^m)) = 1$.

We decompose the simulation in two steps. In the first step we sample the process on a uniform supergrid by explicitly decomposing the covariance matrix for the samples on this grid (given $X(0)$). The conditional covariance is derived from (2.1) and (6.1). Then we continue by refining just as above. The only difference is that when a new point is sampled it is conditioned on more points than just the two nearest neighbors.

A MATLAB script with the implementation in the simplest case, where we only condition on the two nearest neighbors and use the correlation structure of the sampled fBm, is included below. Thus, the realizations will be slightly biased for the short scales. In the actual simulations we conditioned on the 10 nearest neighbors and used a super grid of size 2^{10}. The script for the general case can be found at http://cartan.stanford.edu/mgss.

We simulate realizations of a *local* power law in the following way. We divide the data vector into sections of length 2^{13} points each. Within each section the parameters, the exponent and the intensity, are kept constant. Then the process is simulated sequentially in each section starting with section one. When simulating the realization within a section the realization is conditioned on the last point of the previous subsection.

B.3 MATLAB script for simulation algorithm

```
     function    X=ssim(d,H)
%    Input   : d - dyadic dimension
%            : H - Hurst exponent
%    Output  : X - Simulated sequence

     n=2^d+1;
     r=randn(n,1);
     y=[zeros(n-1,1);r(n)];
     for i=1:d
       s=((2-2^(2*H-1))^.5)*2^(-i*H);
       j=2^(d-i);
       X(1+j:2*j:n-j)=s*r(1+j:2*j:n-j)+.5*(X(1:2*j:n-2*j)+X(1+2*j:2*j:n));
     end
```

Acknowledgement

We thank V. Kruger, C. Rino and D. Washburn with whom we worked closely on this estimation project.

References

[1] M. Asch, W. Kohler, G. Papanicolaou, M. Postel and B. White, Frequency content of randomly scattered signals, Siam Review, 33, pp. 519-626, 1991.

[2] P. Abry, P. Goncalves and P. Flandrin, Wavelets, spectrum analysis and $1/f$ processes, in 'Wavelets and Statistics', A. Antoniadis and G. Oppenheim editors, Springer Lecture Notes in Statistics 103, Springer Verlag, pp. 15-29, 1995.

[3] N. Cressie, Statistics for spatial data, Wiley New York, 1993.

[4] R. Dahlhaus, Efficient parameter estimation for self-similar processes, Ann. Statist., 17, pp. 1749-1766, 1989.

[5] J. Feder, Fractals, Plenum Press, New York, 1988.

[6] W. Feller. An introduction to probability theory and its applications, Vol II, Wiley, 1970.

[7] P. Flandrin, Wavelet Analysis and Synthesis of Fractional Brownian Motion, IEEE Trans. Information Theory, 38, pp. 910-917, 1992.

[8] R. Fox and M. Taqqu, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, Ann. Statist., 14, pp. 517-532, 1986.

[9] P. Goncalves and P. Flandrin, Scaling exponents estimation from time-scale energy distributions, in IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP-92, San Francisco, pp. V.157-V.160, 1992.

[10] P. Goncalves and P. Abry, Multiple window wavelet transform and local scaling exponent estimation, Research Report, INRIA, 1996.

[11] P. Hall, H.L. Koul and B.A. Turlach, Note on Convergence Rates of Semiparametric Estimators of Dependence Index, Annals of Statistics, 25, pp. 1725-1739, 1997.

[12] W. Hardle, Smoothing techniques, Springer Verlag, 1991.

[13] L. Hudgins, C.A. Friehe and M.E. Mayer, Wavelet transforms and atmospheric turbulence, Phys. Rev Letters, 71, pp. 3279-3282, 1993.

[14] G. Kaiser, A Friendly Guide to Wavelets, Birkhäser, 1994.

[15] I. Karatzas and S. E. Shreve, Brownian Motion and and Stochastic Calculus, Springer-Verlag, New York, 1991.

[16] J.T. Kent and A.T.A. Wood, Estimating the fractal dimension of a locally self-similar Gaussian process by using increments, Journal of the Royal Statistical Society B, 3, pp. 679-700, 1997.

[17] S. Mallat, G. Papanicolaou and Z. Zhang, Adaptive Covariance Estimation of Locally Stationary Processes, Annales of Statistics, 26, pp. 1-47, 1998.

[18] S. Mallat and Z. Zhang, Matching pursuit in a time frequency dictionary, IEEE Trans. Information Theory, 38, pp. 3397-3415, 1993.

[19] V. Kruger, C. Rino, G. Papanicolaou and K. Solna, Analysis of Aerothermal Data, VISTA Inc. Report, 100 View St., Mountain View CA 94041, 1998.

[20] B. B. Mandelbrot and J. Van Ness, Fractional Brownian motion, fractional noises and applications, SIAM Rev., 10, pp. 422-437, 1968.

[21] A. S. Monin and A. M. Yaglom. Statistical fluid mechanics: Mechanics of turbulence. Vol. 1. The MIT Press, 1971.

[22] B. Ninness, Estimation of 1/f noise, IEEE Trans. on Inf. Theory, 44, pp. 32-46, 1998.

[23] H. Omre, K. Solna and H. Tjelmeland, Simulations of random functions on large lattices, in 'Geostatistics', A. Soares editor, pp. 179-199, Kluwer Academic Publishers, 1993.

[24] G. Papanicolaou, K. Solna and D. Washburn, Segmentation independent estimates of turbulence parameters, SPIE AeroSense Meeting, Orlando Florida, April 1998.

[25] R. F. Peltier and J. Levy-Vehel, Multifractional Brownian motion: definition and preliminary results, Research Report, INRIA, 1995.

[26] R. F. Peltier and J. Levy-Vehel, A new method for estimating the parameters of fractional Brownian motion, Research Report, INRIA, 1994.

[27] D. B. Percival and P. Guttorp, Long-memory processes, the Allan variance and wavelets, in 'Wavelets in geophysics, E. Foufoula-Georgiou and P. Kumar editors, pp. 325-344, Academic Press, 1994.

[28] K. Rachedi, Simulation and estimation of fractal processes, Report, Department of Mathematics, Stanford University, 1998.

[29] B. Ripley, Spatial Statistics, Wiley, New York, 1981.

[30] L. Rossi, G. Kaiser, D. Washburn, Recovery of Kolmogorov Statistics in Thermal Mixing in the Troposphere: The Hazards of Real Data, SPIE Meeting Orlando Florida, April 1998.

[31] J. W. Strohbehn, Laser beam propagation in the atmosphere, Springer Verlag, 1978.

[32] P.M. Robinson, Log-periodogram regression of time series with long range dependence, Annals of Statistics, 23, pp. 1048-1072, 1995.

[33] M.S. Taqqu and V. Teverovsky, Estimators for long-range dependence: an empirical study, Fractals, 3, pp. 785-798, 1995.

[34] A. H. Tewfik and M. Kim, Correlation structure of the discrete wavelet coefficients of fractional Brownian motions, IEEE Transaction on Info. Theory, 38, pp. 904-909, 1992.

[35] D. C. Washburn, D. W. Banton, T. T. Brennan, W. P. Brown, R. R. Butts, S. C. Coy, R. H. Dueck, K. W. Koenig, B. S. Masson, P. H. Peterson, R. W. Praus, G. A. Tyler, B. P. Venet and L. D. Weaver, Airborne laser extended atmospheric characterization experiment (ABLE ACE). Technical report, Phillips Laboratory, Kirtland Air Force Base NM 87117-5776, May 1996.

[36] Y. Wang, J.E. Cavanaugh and C. Song. Self-similarity index estimation via wavelets for locally self-similar processes, preprint, Department of Statistics, University of Missouri, Columbia, MO, 1997.

[37] G. W. Wornell. Signal processing with fractals, a wavelet based approach, Prentice Hall, 1996.

[38] G. W. Wornell and A. V. Oppenheim, Estimation of fractal signals from noisy measurements using wavelets, IEEE Trans. on Signal Proc., 40, pp. 611-623, 1992.