Should Structure Functions Be Used to Estimate Power Laws in Turbulence? A Comparative Study

Gerald A. Nichols–Pagel

Department of Mechanical Engineering, Box 356200, University of Washington, Seattle, Washington 98195–6200

Donald B. Percival

Applied Physics Laboratory, Box 355640, University of Washington, Seattle, Washington 98115–5640

Per G. Reinhall

Department of Mechanical Engineering, Box 356200, University of Washington, Seattle, Washington 98195–6200

Abstract

Second order structure functions are widely used to characterize turbulence in the inertial range because they are simple to compute, particularly in comparison to other second order characterizations such as spectral density functions and wavelet variances. Structure function estimators, however, are highly autocorrelated and, as a result, no suitable theory has been established to provide confidence intervals for turbulence parameters when determined via regression fits through structure function estimates in log/log space. Monte Carlo simulations were performed to compare the performance of structure function estimators of turbulence parameters with corresponding multitaper spectral and wavelet variance estimators. The simulations indicate that these latter estimators have smaller variances than estimators based upon the structure function. In contrast to structure function estimators, the statistical properties of the multitaper spectral and wavelet variance estimators allow for the construction of confidence intervals for turbulence parameters. The Monte Carlo simulations also confirm the validity of the statistical theory behind the multitaper spectral and wavelet variance estimators.

Preprint submitted to Elsevier Science *13 October 2005*

1 Introduction

In his classic work Kolmogorov [5] theorized that, at turbulent scales too small to be directly affected by the energetic motions and too large to be affected by viscosity, the longitudinal and transverse velocity structure functions should only depend on the kinetic energy dissipation rate. This theory results in a statistical description of turbulence in which the structure function and related second order descriptors such as the spectral density function (SDF) and wavelet variance are linear on log/log plots. In particular, the SDF is proportional to $|f|^{\alpha}$, where the constant of proportionality is related to the kinetic energy dissipation rate, and the power law exponent is $\alpha = -5/3$. Corrsin [2] and Obukhov [7] showed independently that the fluctuations of a passive scalar in a turbulent flow should exhibit this same power law behavior with a constant of proportionality that is related to the scalar dissipation rate in addition to the kinetic energy dissipation rate.

Experiments in turbulence often rely on estimates derived from a linear region (in log/log space) to identify the inertial subrange, in order to check for consistency with Kolmogorov–Obukhov–Corrsin (KOC) turbulence, and to test hypotheses concerning corrections to turbulence theory. Estimates of the constant of proportionality are important in estimating kinetic energy and scalar dissipation rates from environmental and laboratory flows. Estimation of the exponent and the constant of proportionality is commonly based upon the log of either a structure function estimator, an SDF estimator or a wavelet variance estimator. It is thus important to understand the statistical merits of the estimators used to characterize measured turbulence. As we demonstrate via examples in Section 2, a typical estimate of the structure function appears to be markedly more stable than certain corresponding SDF estimates (and certain wavelet variance estimates). We demonstrate in this paper (Section 3) that the apparently superior stability of structure function estimates does not necessarily translate into estimates of the exponent and constant of proportionality that are superior to those from suitably chosen SDF and wavelet variance estimates. Moreover, while there is an appealing statistical theory for both SDF and wavelet variance estimators that makes it possible to ascertain the amount of uncertainty in the resulting estimates of the exponent and constant of proportionality, the same does not hold for structure function estimators, for which the sampling theory is quite complicated [3].

The remainder of this paper is organized as follows. After a motivating illustration of the relative merits of structure function and SDF estimators, we define – and consider the basic properties of – estimators for the exponent and the constant of proportionality based upon a multitaper SDF estimator, a structure function estimator and wavelet variance estimators. In Section 3 we use computer experiments to verify the statistical theory for the multi-

taper and wavelet-based estimators and to compare these estimators to the structure function-based estimators. Section 4 summarizes our results and recommendations.

2 Background

As a motivating example, Figure 1 compares two approaches for estimating the exponent α of a process with an SDF that is proportional to a power law $|f|^{\alpha}$ over two decades of frequencies. The left-hand column shows two realizations of a stationary process $\{X_t\}$, with an SDF given by $5f^{-5/3}$ for $0.0025 \leq f \leq 0.25$. The middle column shows two multitaper SDF estimates (Section 2.1), along with least squares lines fitted in log/log space over the frequency range exhibiting the $|f|^{-5/3}$ behavior (the lines are displaced upwards on the plot to make them easier to see). The right hand column shows corresponding plots for structure function estimates (Section 2.2), again with lines two decades in length that represent the theoretical slope and extent of the power law behavior in the simulated series. The multitaper spectral estimates $\hat{S}_X^{(mt)}$ appear very noisy compared to the structure function estimates $\hat{D}_X$. In addition, the proportion of explained variation, R^2, is substantially higher for the structure function estimates. One might conclude that the structure function is the superior estimate here; however, the multitaper estimates of the power law exponent $\alpha \doteq -1.667$ and constant of proportionality $B_1 = 5$ (derived from the slopes and intercepts of the fitted lines) are closer to the true values. In fact, the smoothness of the structure function is the result of a high degree of correlation between structure function estimates that are separated by small to moderate distances. This smoothness masks the variability in the estimated power law exponent and constant of proportionality and leads to the large R^2 statistic for the linear regression. The multitaper estimates, in contrast, rapidly decorrelate with increasing frequency separation; the variability in the estimates is clearly observed. The covariance structure of the multitaper estimate is much simpler, and can be approximated well independently of the data to quantify the variability of the estimated exponent and constant of proportionality and to establish confidence intervals for these parameters.

2.1 Multitaper Spectral Estimation

Multitaper spectral estimation is a technique introduced by Thomson [12] that yields a largely unsmoothed or 'raw' estimator with both low bias and tractable covariance and distributional properties. Multitapering builds upon the notion of a single data taper, which has been used routinely since the 1950s to reduce the bias in spectral estimators. In what follows, we first review how

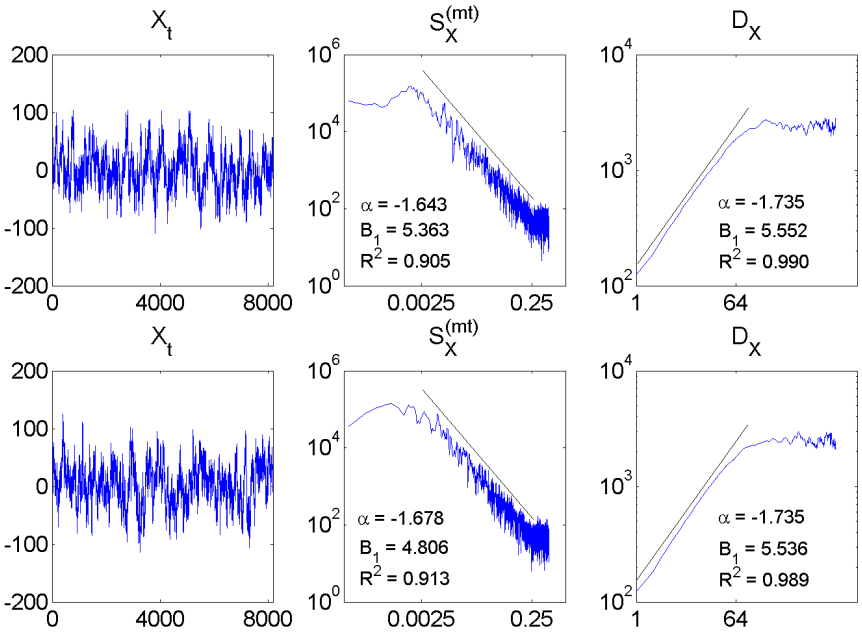

Fig. 1. Comparison of multitaper spectral and structure function estimates

a single data taper is used to form a spectral estimator, after which we define the multitaper spectral estimator $\hat{S}_X^{(mt)}$. We then proceed as in McCoy et al. [6] to formulate an estimator of the exponent of power law processes that is based upon $\hat{S}_X^{(mt)}$ and that is attractive when compared to competing methods.

Suppose we have a time series that can be regarded as a realization of one portion, $X_0, \ldots, X_{N-1}$, of a stationary process with SDF S_X (for simplicity, we assume that $\langle X_t \rangle = 0$, i.e., that the expected value of the process is zero; if this is not a reasonable assumption, the common practice is to replace X_t by $X_t - \overline{X}$ in what follows, where $\overline{X}$ is the sample mean of the time series). A direct spectral estimator of S_X is obtained by computing the magnitude squared of the Fourier transform of the product of the series $\{X_t\}$ and a suitable data taper $\{h_t\}$:

$$\hat{S}_X^{(d)}(f) = \left| \sum_{t=0}^{N-1} h_t X_t e^{-i2\pi ft} \right|^2, \quad |f| \le \frac{1}{2}.$$

The purpose of using the data taper is to obtain an estimator that is approximately unbiased; i.e., $\langle \hat{S}_X^{(d)}(f) \rangle \approx S_X(f)$. Tapering typically 'shrinks' the values of X_t near the beginning and end of the time series toward zero, which effectively causes a loss of information that manifests itself as an increase in the variance of subsequently smoothed estimators.

Multitapering is designed to recover some of the information lost by a single data taper while simultaneously maintaining low bias and decreasing variability to some degree. A multitaper estimator is created by averaging K direct

spectral estimators:

$$\hat{S}_X^{(mt)}(f) = \frac{1}{K}\sum_{k=0}^{K-1} \hat{S}_{X,k}^{(d)}(f),$$

where $\hat{S}_{X,k}^{(d)}$ is the direct spectral estimator obtained using the kth data taper $\{h_{k,t}\}$. The data tapers are chosen so that each generates a direct spectral estimator with good bias properties, and each is approximately uncorrelated with the other $K-1$ spectral estimators. The average of these direct spectral estimators will have good bias properties and a smaller variance than any individual estimator $\hat{S}_{X,k}^{(d)}$. The individual estimators will be nearly uncorrelated if the data tapers are orthogonal [10]:

$$\sum_{t=0}^{N-1} h_{j,t}h_{k,t} = 0 \text{ for } j \neq k.$$

Riedel and Sidorenko [11] suggest using sine tapers:

$$h_{k,t} = \left(\frac{2}{N+1}\right)^{1/2} \sin\left(\frac{(k+1)\pi(t+1)}{N+1}\right).$$

In the inertial range, the SDF of velocity for KOC turbulence has a power law behavior:

$$S_X(f) = C\epsilon^{2/3}\left|f\right|^{-5/3},$$

where C is a universal constant and ϵ is the kinetic energy dissipation rate. On a logarithmic scale, the SDF is linear in frequency with an intercept that is related to the kinetic energy dissipation rate:

$$\log\left(S_X(f)\right) = \log\left(C\epsilon^{2/3}\right) - \frac{5}{3}\log\left(|f|\right).$$

More generally, for an SDF of the form

$$S_X(f) = B_1\left|f\right|^{\alpha}, \tag{1}$$

we can write

$$\log\left(S_X(f)\right) = \log\left(B_1\right) + \alpha\log\left(|f|\right).$$

According to Walden et al. [13], the log multitaper estimator can be written as the sum of two components. The first is nonstochastic and is the true log spectrum plus a known constant, $\psi(K) - \log(K)$, where ψ is the digamma function. The second is stochastic noise η_j whose distribution is dictated by that of a $\log\left(\chi^2_{2K}\right)$ random variable. Hence η_j has a *known* variance $\sigma_\eta^2 = \psi'(K)$, where ψ' is the trigamma function (if K is about 5 or greater, the distribution of η_j is approximately Gaussian). Following McCoy et al. [6], we can estimate α and B_1 by using the linear regression

$$Y_j = c + \alpha x_j + \eta_j, \quad j = 1, \ldots, M, \tag{2}$$

where $Y_j = \log\{\hat{S}^{(mt)}(f_j)\}$, $c = \log(B_1) + \psi(K) - \log(K)$, $x_j = \log(f_j)$ and M is the number of frequency ordinates used in the regression over a range of positive Fourier frequencies f_j. In matrix form we have

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{N},$$

where $\mathbf{Y}$ is the column vector of Y_j variables, $\mathbf{X}$ is a matrix with ones in the first column and the x_j variables in the second column, $\boldsymbol{\Theta} = [c, \alpha]^T$, and $\mathbf{N}$ is the column vector of the η_j variables. The corresponding least squares estimator is

$$\hat{\boldsymbol{\Theta}} = [\mathbf{X}^T\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{Y},$$

with covariance matrix

$$\left\langle [\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}][\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}]^T \right\rangle = [\mathbf{X}^T\mathbf{X}]^{-1}\mathbf{X}^T\boldsymbol{\Sigma}\mathbf{X}[\mathbf{X}^T\mathbf{X}]^{-1}. \tag{3}$$

Here $\boldsymbol{\Sigma}$ is the symmetric Toeplitz covariance matrix for $\mathbf{N}$. Following Walden et al. [13], the $(j, j+\tau)$th element of the covariance matrix for $\mathbf{N}$ can be modeled by

$$s_{\eta,\tau} \equiv \mathrm{cov}\{\eta_j, \eta_{j+\tau}\} \approx \begin{cases} \sigma_\eta^2 \left(1 - \frac{|\tau|}{K+1}\right), & \text{if } |\tau| \le (K+1); \\ 0, & \text{otherwise.} \end{cases}$$

2.2 Structure Function Estimation

Kolmogorov's original hypothesis [5] was formulated in terms of the second order structure function:

$$D_X(\tau) = \langle (X_{t+\tau} - X_t)^2 \rangle.$$

For a process with an SDF given by Equation 1 with $-3 \le \alpha < -1$, the structure function takes the form [14]

$$D_X(\tau) = B_2|\tau|^{-(\alpha+1)}, \text{ where } B_2 = \frac{B_1}{(2\pi)^\alpha \Gamma(-\alpha)\sin(-\pi(\alpha+1)/2)}.$$

After taking the logarithm of both sides, we obtain for positive τ

$$\log(D_X(\tau)) = \log(B_2) - (\alpha+1)\log(\tau). \tag{4}$$

This relationship can be used to estimate α and B_2 by performing a regression involving the usual structure function estimator, namely,

$$\hat{D}_X(\tau) = \frac{1}{N-\tau} \sum_{t=0}^{N-1-\tau} (X_{t+\tau} - X_t)^2,$$

where $N - \tau$ is the number of pairs of sample points separated by the positive distance τ. The regression model in this case is

$$Y_\tau = \log(B_2) - (\alpha + 1)x_\tau + \eta_\tau, \tag{5}$$

where $Y_\tau = \log(\hat{D}_X(\tau))$, $x_\tau = \log(\tau)$, and the expected value of the error term η_τ is assumed to be zero. In contrast to 3, the covariance matrix is difficult to determine because the error terms are highly correlated, with the correlation structure being model dependent.

2.3 Wavelet Variance Estimation

Given a process $\{X_t\}$ that is either stationary or whose increments of a certain order form a stationary process, we can use a jth level wavelet filter $\{\tilde{h}_{j,l}, j = 0, 1, \ldots, L_j - 1\}$ to define a wavelet coefficient process

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j - 1} \tilde{h}_{j,l} X_{t-l} \text{ for } j = 1, 2, 3, \ldots.$$

This process is associated with changes in co-located averages in $\{X_t\}$ on a standardized scale of $\tau_j = 2^{j-1}$. Under mild conditions (see Percival and Walden [8] for details), the process $\{\overline{W}_{j,t}\}$ is stationary with mean zero. The wavelet variance is defined to be the variance of $\overline{W}_{j,t}$:

$$\nu_X^2(\tau_j) = \langle \overline{W}_{j,t}^2 \rangle.$$

This variance offers a decomposition of the variability in $\{X_t\}$ over temporal or spatial scales. If $\{X_t\}$ is stationary with variance σ_X^2, then

$$\sum_{j=1}^{\infty} \nu_X^2(\tau_j) = \sigma_X^2;$$

on the other hand, if $\{X_t\}$ is nonstationary but has stationary increments of a certain order, then

$$\sum_{j=1}^{\infty} \nu_X^2(\tau_j) = \infty.$$

Given $X_0, X_1, \ldots, X_{N-1}$, the jth level maximal overlap discrete wavelet transform (MODWT) wavelet coefficients are found by circular convolution with $\{\tilde{h}_{j,l}\}$:

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j - 1} \tilde{h}_{j,l} X_{t-l \bmod N}, \quad t = 0, 1, \ldots, N-1.$$

The wavelet variance is estimated by computing the sample variance of the MODWT wavelet coefficients:

$$\tilde{\nu}_X^2(\tau_j) \equiv \frac{1}{N} \sum_{t=0}^{N-1} \widetilde{W}_{j,t}^2.$$

The estimate $\tilde{\nu}_X^2(\tau_j)$ is referred to as the biased estimate of the wavelet variance because, whereas $\widetilde{W}_{j,t} = \overline{W}_{j,t}$ and hence $\langle \widetilde{W}_{j,t}^2 \rangle = \nu_X^2(\tau_j)$ when $L_j - 1 \leq t \leq N - 1$, the same does not necessarily hold when $0 \leq t < L_j - 1$. Excluding these latter coefficients gives an unbiased estimator of the wavelet variance:

$$\hat{\nu}_X^2(\tau_j) \equiv \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \widetilde{W}_{j,t}^2 = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \overline{W}_{j,t}^2,$$

where $M_j \equiv N - L_j + 1$.

Percival and Walden [8] show that the wavelet variance can be related to the SDF because the jth level MODWT wavelet filter acts like a band-pass filter with pass-band

$$\frac{1}{2^{j+1}} < |f| \leq \frac{1}{2^j},$$

implying that

$$\nu_X^2(\tau_j) \approx 2 \int_{1/2^{j+1}}^{1/2^j} S_X(f)\, df, \tag{6}$$

where the factor of two comes into play because we use SDFs that are assumed to be two-sided. Substitution of Equation 1 gives:

$$\nu_X^2(\tau_j) \approx 2B_1 \int_{1/2^{j+1}}^{1/2^j} f^\alpha\, df = B_3 \tau_j^{-(\alpha+1)},$$

where

$$B_3 \equiv \frac{2B_1(2^{-(\alpha+1)} - 4^{-(\alpha+1)})}{\alpha + 1}.$$

After taking the logarithm of both sides, we obtain

$$\log\left(\nu_X^2(\tau_j)\right) \approx \log\left(B_3\right) - (\alpha + 1) \log\left(\tau_j\right), \tag{7}$$

which is quite similar to Equation 4 involving the structure function.

Percival and Walden [8] formulate an estimator for α and B_3 by regressing the logarithm of the wavelet variance estimates against the logarithm of the wavelet scale:

$$Y(\tau_j) = \zeta + \beta \log\left(\tau_j\right) + e_j, \tag{8}$$

where $\beta = -(\alpha + 1)$,

$$\zeta = \log\left(B_3\right) = \log\left(\frac{2B_1(2^{-(\alpha+1)} - 4^{-(\alpha+1)})}{\alpha + 1}\right),$$

$$Y(\tau_j) \equiv \log\left(\nu_X^2(\tau_j)\right) - \psi\left(\frac{\eta_j}{2}\right) + \log\left(\frac{\eta_j}{2}\right),$$

e_j has zero mean and variance $\psi'(\frac{\eta_j}{2})$, and η_j is the 'equivalent degrees of freedom' for a scaled chi-squared distribution [8]. We estimate η_j using the following relationship from Percival and Walden [8]:

$$\eta_j = \max\{M_j/2^j, 1\}.$$

We can obtain estimates of β and ζ using an ordinary least squares approach; more appropriately, since the variance of the error term grows with increasing j, we can use a weighted least squares approach with weights $w_j = 1/\psi'(\frac{\eta_j}{2})$. The ordinary least squares estimator takes the same form as the multitaper estimator with $Y_j = Y(\tau_j)$ and $x_j = \log(\tau_j)$. The corresponding weighted least squares estimator is

$$\hat{\mathbf{\Theta}} = [\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{Y}, \tag{9}$$

with covariance matrix

$$\left\langle [\hat{\mathbf{\Theta}} - \mathbf{\Theta}][\hat{\mathbf{\Theta}} - \mathbf{\Theta}]^T \right\rangle = [\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X}]^{-1}, \tag{10}$$

where $\mathbf{\Theta} = [\zeta, \beta]^T$ and $\mathbf{\Sigma}$ is a diagonal matrix with elements $\psi'(\frac{\eta_j}{2})$. Jensen [4] shows that the approximation for the slope in Equation 7 becomes exact as $j \to \infty$ and is quite reasonable even when $j \geq 2$ or 3. The fitness of the approximation for the intercept will be evaluated, by numerical integration of the product of the true SDF and the wavelet squared gain function, in Section 3.

2.4 Scaling range averaging for the estimation of B_1

In turbulent signals, the constant of proportionality in the power law model is often used as a surrogate for the energy of the signal contained in the inertial range. It is through this interpretation that a relationship between the constant of proportionality and the dissipation rate is formed. In the previous sections we have outlined methods for estimating B_1 either directly or indirectly via B_2 and B_3, which are related to the intercept of the SDF, structure function, or wavelet variance, each on a logarithmic scale; however, if the slope of the line for this intercept is not known, the intercept, by itself, does not indicate the signal energy contained within any given band of frequencies or any range of scales.

To form an estimator for B_1 that can be more readily interpreted in terms of the dissipation rate, we propose the following procedure. Note that the expected values of the log/log models 2, 5, and 8 for the multitaper, structure function, and wavelet variance cases all take the form $b+\gamma x$ over an appropriate

scaling region $x_1 \leq x \leq x_2$. The average value of the log/log model over this region is given by

$$A \equiv \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} b + \gamma dx = b + \frac{\gamma}{2}(x_1 + x_2).$$

In the multitaper case, $b = \log(B_1) + \psi(K) - \log(K)$, and $\gamma = \alpha$, so we can estimate B_1 (assuming $\alpha = -5/3$) using

$$\hat{\hat{B}}_1^{(mt)} = \exp\{\hat{A}^{(mt)} + \frac{5}{6}(x_1 + x_2) - \psi(K) + \log(K)\},$$

where $\hat{A}^{(mt)}$ is formed by taking the average values of $Y_j = \log\{\hat{S}^{(mt)}(f_j)\}$ over the region of frequencies for which the power law model is assumed to hold. Based upon similar averages $\hat{A}^{(sf)}$ and $\hat{A}^{(wv)}$ for the structure function and wavelet variance cases, we obtain

$$\hat{\hat{B}}_1^{(sf)} = (2\pi)^{-5/3}\Gamma(5/3)\sin(\pi/3)\exp\{\hat{A}^{(sf)} - \frac{1}{3}(x_1 + x_2)\}$$

and

$$\hat{\hat{B}}_1^{(wv)} = -\frac{1}{3(2^{2/3} - 4^{2/3})}\exp\left\{\hat{A}^{(wv)} - \frac{1}{3}(x_1 + x_2)\right\}.$$

The estimators of B_1 obtained through this procedure will reflect the energy content of the chosen scaling range, regardless of the slope, but will only reflect the intercept of the regression line if the estimated value for α is $-5/3$. This is preferable to the alternative, namely, that the estimate of B_1 reflects the intercept of the regression line, regardless of the slope, and only reflects the energy content of the scaling range when the α estimate is $-5/3$.

3 Results

To compare the relative merits of the estimators described in the previous section, we consider a Gaussian stationary process with an SDF defined over $-0.5 \leq f \leq 0.5$ such that

$$S(f) = \begin{cases} 5 \times 0.0025^{-5/3}, & \text{for } 0 \leq |f| \leq 0.0025; \\ 5|f|^{-5/3}, & \text{for } 0.0025 < |f| \leq 0.25; \text{ and} \\ 5 \times 0.25^{-5/3}, & \text{for } 0.25 < |f| \leq 0.5. \end{cases} \tag{11}$$

This SDF has power law behavior $B_1 |f|^\alpha$ over a two decade range of positive frequencies, with $B_1 = 5$ and $\alpha = -5/3$, and is flat on either side.

3.1 Theoretical Comparison

In this section, the theoretical performance of each of the estimators is considered; first the bias, followed by the variance. This effort serves to validate the computer simulations and provide a basis for understanding observed biases in the wavelet variance and structure function estimators.

It is difficult to define a process in terms of its autocovariance sequence (ACVS) since one must ensure that the ACVS is positive semidefinite. For this reason, we chose to define our process in terms of the SDF. In doing so, we have introduced some bias into the structure function and the wavelet variance based estimates of α and B_1 since the mapping of the power law behavior from the SDF is inexact for bandlimited processes. The considered process was chosen to have exact power law behavior over a two decade range in its SDF. The wavelet variance and structure function do not display exact power law behavior for this band limited process; they are nearly linear over a range of scales that is slightly less than two decades in length. The wavelet variance that corresponds to the chosen SDF (Equation 11) was obtained through numerical integration of

$$\nu_X^2(\tau_j) = \int_{-1/2}^{1/2} |H_j(f)|^2 S_X(f)\, df, \tag{12}$$

where H_j is the transfer function for the filter $\{\tilde{h}_{j,l}\}$. The structure function was obtained from

$$D_X(\tau) = 2(s_{X,0} - s_{X,\tau}), \tag{13}$$

after computing the ACVS $s_{X,\tau}$ from the inverse discrete Fourier transform of the SDF. Regressions were performed (over identical ranges of scale as in the empirical comparisons) on this wavelet variance and structure function to determine the corresponding values for the estimators of the parameters α and B_1. These values differ from the values prescribed in the SDF. This difference is related to the imperfect mapping of the power law behavior of the simulated process rather than the performance of the individual estimators. Tables 1 and 2 summarize the theoretical performance of the estimators of the parameters α and B_1. The LA(16) wavelet was selected because it provides a good approximation to a band-pass filter. Since we know the true SDF for the process we are able to estimate the introduced bias, along with the bias that results from the band-pass filtering approximation in the wavelet variance formulation. It seems most appropriate to compare only the magnitude of the variance, of each of the estimators, to avoid this introduced bias.

The theoretical variances for the estimators of the parameter $\mathbf{\Theta} = [c, \alpha]^T$ in the linear regression model of Equation 2 for the multitaper case are given by the diagonal elements of the covariance matrix shown in Equation 3. Since

Table 1
Theoretical performance of α estimators.

	α	bias^2	theoretical variance	theoretical MSE
$\hat{S}_X^{(mt)}$	-1.666667	0.000000	0.000829	0.000829
$\hat{D}_X$[1,100]	-1.631479	0.001238	no theory	no theory
$\hat{D}_X$[1, 64]	-1.659202	0.000076	no theory	no theory
$\hat{D}_X$[2,100]	-1.625957	0.001657	no theory	no theory
$\hat{D}_X$[2, 64]	-1.656306	0.000107	no theory	no theory
$\hat{\nu}_X^2$ OLS	-1.669780	0.000010	0.002131	0.002140
$\hat{\nu}_X^2$ WLS	-1.664584	0.000004	0.000800	0.000804

$c = \log(B_1) + \psi(K) - \log(K)$, the least squares estimator $\hat{c}$ of c can be used to define an estimator for B_1, namely,

$$\hat{B}_1^{(mt)} \equiv \exp\{\hat{c} - \psi(K) + \log(K)\}.$$

Noting that

$$\mathrm{var}\{\hat{B}_1^{(mt)}\} = K^2 \exp\{-2\psi(K)\}\mathrm{var}\{\exp(\hat{c})\}$$

and assuming that $\hat{c}$ is normally distributed, then $\exp(\hat{c})$ obeys a log normal distribution with parameters $\mu = \log(B_1)+\psi(K)-\log(K)$ and σ^2 given by the upper left-hand element of the covariance matrix in Equation 3. A standard result says that the variance of $\exp(\hat{c})$ is given by $\exp(2\mu + \sigma^2)\{\exp(\sigma^2) - 1\}$, and hence we have

$$\mathrm{var}\{\hat{B}_1^{(mt)}\} = B_1^2 \exp(\sigma^2)(\exp(\sigma^2) - 1).$$

The variance of the least squares estimator $\hat{\alpha}$ of α is given by the lower right-hand element of the covariance matrix in Equation 3. In the wavelet variance case, the theoretical variances for the estimator of the parameters $\mathbf{\Theta} = [\zeta, \beta]^T$ can be obtained from an expression analogous to Equation 3 for the ordinary least squares estimator and from Equation 10 for the weighted least squares estimator. The variance of the wavelet variance estimator of B_1 is

$$\mathrm{var}\{\hat{B}_1^{(wv)}\} = \left(\frac{\alpha+1}{2B_1(2^{-(\alpha+1)} - 4^{-(\alpha+1)})}\right)^2 e^{2\mu+\mathrm{var}\{\hat{\zeta}^{(wv)}\}}(e^{\mathrm{var}\{\hat{\zeta}^{(wv)}\}} - 1),$$

where μ is the actual mean $\log\left(\frac{2B_1(2^{-(\alpha+1)}-4^{-(\alpha+1)})}{\alpha+1}\right)$ and $\alpha = -\frac{5}{3}$ is the actual value of α. The theoretical variances for the wavelet variance estimators of α and β are the same.

Table 2
Theoretical performance of B_1 estimators.

	B_1	bias^2	theoretical variance	theoretical MSE
$\hat{S}_X^{(mt)}$	5.000000	0.000000	0.130876	0.130876
$\hat{D}_X$[1,100]	5.105634	0.011159	no theory	no theory
$\hat{D}_X$[1, 64]	4.753328	0.060847	no theory	no theory
$\hat{D}_X$[2,100]	5.215992	0.046652	no theory	no theory
$\hat{D}_X$[2, 64]	4.801153	0.039540	no theory	no theory
$\hat{\nu}_X^2$ OLS	5.076332	0.005827	0.154553	0.160380
$\hat{\nu}_X^2$ WLS	5.122484	0.015002	0.046261	0.061263

3.2 Empirical Comparison

Realizations of the process described by Equation 11 were generated using the Gaussian spectral synthesis method (GSSM) [9]. GSSM is an approximate frequency domain method that can be used to simulate a zero mean Gaussian stationary process with a specified SDF. A segment of length N is sampled from a harmonic process of length M that is generated using a DFT approach. Percival [9] shows that, by making M large, the ACVS of the simulated series can be made arbitrarily close to the desired ACVS out to lag $N-1$. We experimented with different values of M to determine their effect on the structure function. Since the structure function may be expressed in terms of the ACVS (Equation 13), the structure function of the simulated process will also converge to the desired structure function as M becomes large. Estimates of α and B_1 were obtained using the structure function for 10,000 realizations of the simulated process with $M = 4N$, $M = 8N$, $M = 16N$, and $M = 32N$. No significant change in the estimates of α or B_1 was observed with increasing M, suggesting that $M = 4N$ is large enough to produce accurate results.

Estimates of B_1 and α were obtained with each of the estimators for 10,000 realizations of the simulated process. The multitaper method used $K = 5$ sine tapers, and the regression was performed on frequencies in the power-law portion of the SDF. The structure function was calculated over four different ranges of unitless separations: 1 to 100, 2 to 100, 1 to 64, and 2 to 64. Using Daubechies LA(16) wavelets, ordinary and weighted least squares estimates were obtained on wavelet transform levels $j = 2$ through $j = 7$ using the unbiased estimate of the wavelet variance. The squared bias, variance, and mean squared error (MSE) of each of the estimators are summarized in Table 3 for α, Table 4 for B_1, and Table 5 for estimates of B_1 obtained through scaling range averaging. Figures 2 through 4 show corresponding box and whisker diagrams. For α (Table 3 and Figure 2), the multitaper and weighted least

Table 3
Summary of α estimates.

	$\hat{\alpha}$	bias^2	variance	MSE
$\hat{S}_X^{(mt)}$	-1.666779	0.000000	0.000717	0.000717
$\hat{D}_X$[1,100]	-1.629487	0.001382	0.001160	0.002542
$\hat{D}_X$[1, 64]	-1.657791	0.000079	0.000814	0.000893
$\hat{D}_X$[2,100]	-1.623774	0.001840	0.001404	0.003244
$\hat{D}_X$[2, 64]	-1.654716	0.000143	0.001033	0.001176
$\hat{\nu}_X^2$ OLS	-1.664909	0.000003	0.001848	0.001851
$\hat{\nu}_X^2$ WLS	-1.662119	0.000021	0.000746	0.000767

Table 4
Summary of B_1 estimates.

	$\hat{B}_1$	bias^2	variance	MSE
$\hat{S}_X^{(mt)}$	5.011043	0.000122	0.113948	0.114070
$\hat{D}_X$[1,100]	5.140074	0.019621	0.147397	0.167018
$\hat{D}_X$[1, 64]	4.771829	0.052062	0.065295	0.117358
$\hat{D}_X$[2,100]	5.261158	0.068203	0.219234	0.287438
$\hat{D}_X$[2, 64]	4.826778	0.030006	0.105196	0.135202
$\hat{\nu}_X^2$ OLS	5.122879	0.015099	0.138459	0.153558
$\hat{\nu}_X^2$ WLS	5.137826	0.018996	0.043069	0.062065

squares wavelet estimates have about 10% less variance than the best structure function estimate, which is based on $[1, 64]$. For B_1 estimated via $\hat{B}_1$ (Table 4 and Figure 3), the structure function $[1, 64]$ and $[2, 64]$ estimates have smaller variances than the multitaper estimate, but the weighted least squares wavelet estimate has the smallest variance. For B_1 estimated based upon scaling range averaging (Table 5 and Figure 4), all of the structure functions estimates are better than the multitaper and wavelet variance estimates in terms of variance.

A comparison of Tables 3 and 4 with Tables 1 and 2 indicates very good agreement between theoretical and corresponding empirical values for all estimators. In particular, the empirical estimates of α and B_1 agree well with the theoretical values for each estimator. The theoretical predictions of the variance of the multitaper and wavelet variance estimators of α and B_1 are between 7% and 16% larger than observed, indicating that the sampling theory provides accurate, but somewhat conservative, predictions of the variability of the estimate.

Table 5
Summary of estimates of B_1 obtained using scaling range averaging.

	$\hat{\hat{B}}_1$	bias^2	variance	MSE
$\hat{S}_X^{(mt)}$	5.010497	0.000110	0.064256	0.064366
$\hat{D}_X$[1,100]	4.575704	0.180028	0.024661	0.204689
$\hat{D}_X$[1, 64]	4.612862	0.149876	0.019428	0.169304
$\hat{D}_X$[2,100]	4.601591	0.158729	0.032927	0.191656
$\hat{D}_X$[2, 64]	4.649701	0.122709	0.026435	0.149144
$\hat{\nu}_X^2$	5.091149	0.008308	0.045419	0.053727

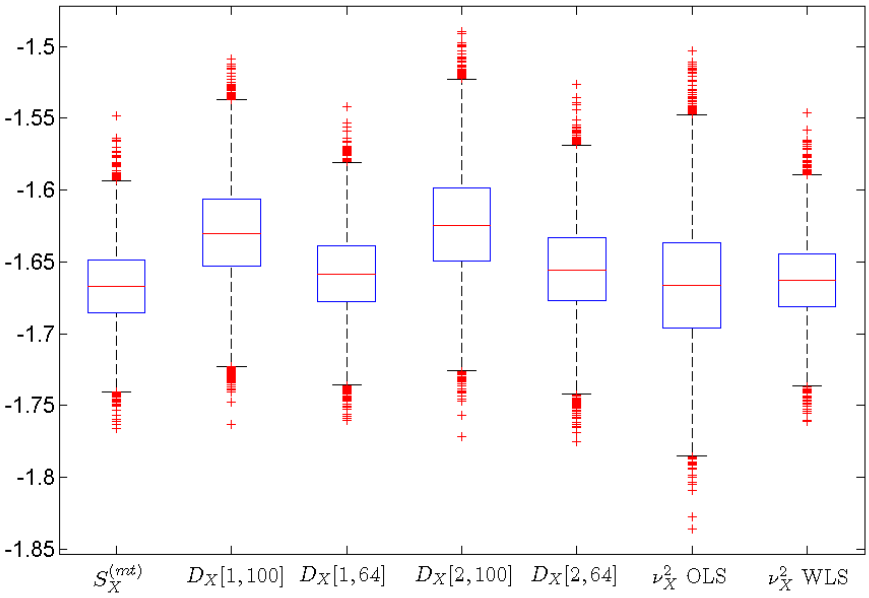

Fig. 2. Box and whisker diagram comparison of α estimates.

4 Discussion

It has been noted in the turbulence literature (see Antonia and Smalley [1], and references therein) that spectra of turbulent signals sometimes display more extensive scaling regions than do structure function estimates of the same signals. In this study a stationary Gaussian process was defined in terms of its SDF with two decades of power law behavior. The corresponding structure function for this process does not display this power law behavior exactly, but instead shows a much narrower range of scales where the slope is nearly linear. Regressions performed over this nearly linear region yielded estimates of the power law exponent and constant of proportionality that are approximately in agreement with the prescribed values. One can also imagine processes that

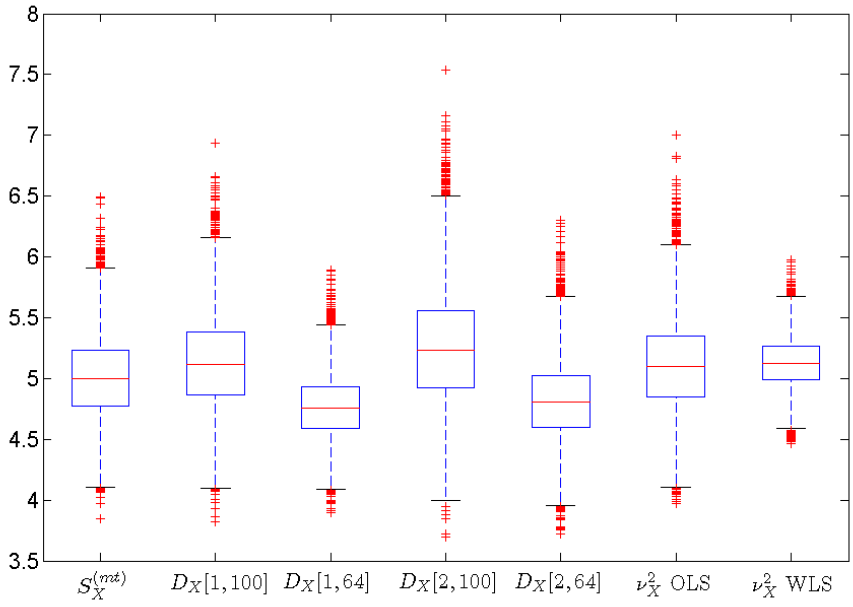

Fig. 3. Box and whisker diagram comparison of B_1 estimates.

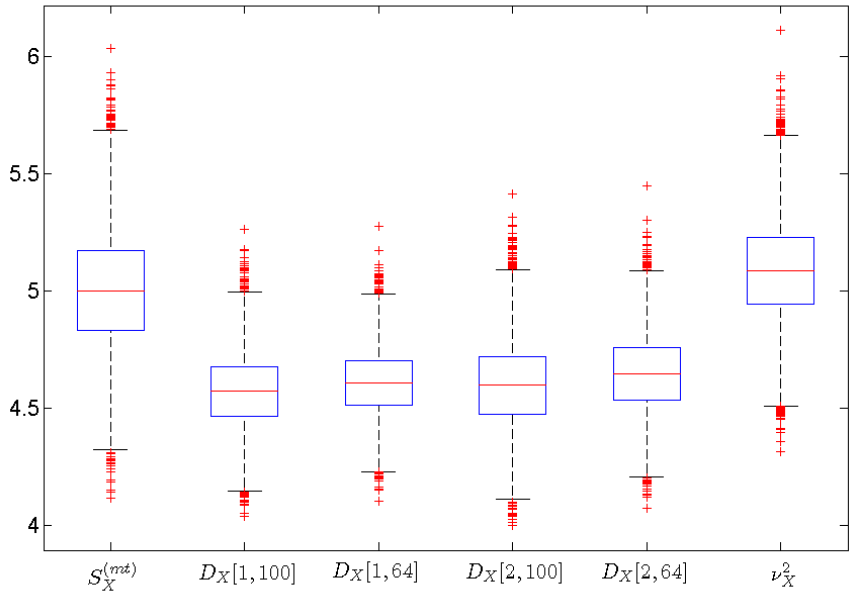

Fig. 4. Box and whisker diagram comparison of estimates of B_1 obtained through scaling range averaging.

display more extensive power law behavior in their structure functions than in their SDFs, but these processes are difficult to define because one must ensure that the ACVS is positive semidefinite. In either case, the process will be more simply described by the characterization displaying the most extensive scaling

region; in practice, there is evidence to suggest that this would be the spectral representation.

Multitaper and wavelet methods did better than the structure function approach in estimating α in terms of variance (see Table 3) while the wavelet method displayed a smaller variance in estimating B_1 from the regression intercept than the structure function (see Table 4). More importantly, the variance and MSE can be computed theoretically for the multitaper and wavelet methods, but not for the structure function method, which is a serious limitation in practical applications. The multitaper estimator, the weighted least squares estimates using the LA(16) wavelet, and structure function [1, 64] estimator have the smallest variance for the estimation of α. The multitaper, structure function [1, 64], structure function [2, 64], and weighted least squares wavelet estimates have the smallest variance for the estimation of B_1 from the regression intercept. The variance of the multitaper estimator is 12% smaller than the structure function [1, 64] estimator for α and 75% larger for B_1. The wavelet variance estimator gives an improvement over the structure function [1, 64] of 8% for α and 34% for B_1. Theoretical confidence intervals can be based upon the multitaper spectral estimates and wavelet variance estimates; however, no theory is available to provide confidence intervals based upon the structure function estimates because of their highly correlated nature. The scaling range averaging approach is recommended, in contrast to the regression intercept based estimates, because it provides an estimate of B_1 that may be interpreted in terms of the signal energy contained in the scaling range, regardless of the slope. The simulations indicate that the best estimates of B_1 are also obtained through scaling range averaging. Scaling range averaging produces estimates with variances that are very similar to the weighted least squares estimator for the wavelet variance; a reduction in variance is observed when scaling range averaging is applied to the multitaper or structure function estimators.

We recommend either the multitaper approach or the wavelet variance approach as the estimator of choice because of their ability to provide accurate estimates with appropriate confidence intervals. The wavelet variance has the best performance as long as the filter length is long enough to act as an adequate approximation to a band-pass filter. It is important to use the weighted least squares approach with the wavelet variance estimator to achieve this performance. The multitaper approach provides quality estimates with very little bias. The multitaper approach is relatively simple to implement and computationally efficient when a fast Fourier transform algorithm is used.

5 Acknowledgments

We thank James J. Riley for numerous discussions. This work was supported by the U.S. Air Force Office of Scientific Research, Contract #FA9451–04–M–0390, AM03.

References

[1] R. A. Antonia and R. J. Smalley Scaling Range Exponents from X-Wire Measurements in the Atmospheric Surface Layer, *Boundary Layer Meteorology* **100** (2001) 439–457.

[2] S. Corrsin, On the spectrum of isotropic temperature fluctuations in isotropic turbulence, *Journal of Applied Physics* **22** (1951).

[3] N. Cressie, Fitting variogram models by weighted least squares, *Journal of the International Association for Mathematical Geology* **17** (1951), 563–586.

[4] M. J. Jensen, Using wavelets to obtain a consistent ordinary least-squares estimator of the long-memory parameter, *Journal of Forecasting* **18** (1999) 17–32.

[5] A. N. Kolmogorov, Local structure of turbulence in an incompressible fluid at very high Reynolds numbers, *Doklady Akad. Nauk SSSR* **30** (1949) 299–303.

[6] E. J. McCoy and A. T. Walden and D. B. Percival, Multitaper spectral estimation of power law processes, *IEEE Transactions on Signal Processing* **46** (1998) 655–668.

[7] A. M. Obukhov, Structure of the temperature field in a turbulent flow, *Izv. Akad. Nauk SSSR, Ser. Geograf. Geofiz.* **13** 1948.

[8] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis* (Cambridge University Press, Cambridge, 2000).

[9] D. B. Percival, Simulating Gaussian random processes with specified spectra, *Computing Science and Statistics* **24** (1992) 534–538.

[10] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications* (Cambridge University Press, Cambridge, 1993).

[11] K. S. Riedel and A. Sidorenko, Minimum biased multitaper spectral estimation, *IEEE Transactions on Signal Processing* **43** (1995) 188–195.

[12] D. J. Thomson, Spectrum estimation and harmonic analysis, *Proceedings of the IEEE* **70** (1982) 1055–1096.

[13] A. T. Walden, D. B. Percival and E. J. McCoy, Spectrum estimation by wavelet thresholding of multitaper estimators, *IEEE Transactions on Signal Processing* **46** (1998) 3153–3165.

[14] A. M. Yaglom, *Correlation Theory of Stationary and Related Random*

Functions I; Basic Results (Springer-Verlag New York, Inc., New York, 1987)