Multiscale denoising of self similar processes

Brani D. Vidakovic
Institute of Statistics and Decision Sciences, Duke University, Durham, NC 27708-0251

Gabriel G. Katul
School of the Environment, Duke University, Durham, NC 27708-0328

John D. Albertson
Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22903

Short title: MULTISCALE DENOISING OF SELF SIMILAR PROCESSES
Abstract. A practical limitation to investigating self-similarity in geophysical phenomena from their measured state variables is that measured signals are typically convolved with instrumentation noise at multiple scales. This study develops and tests a multiscale Bayesian model (BEFE) for separating a $1/f$-like signal from inherent instrumentation noise and contrasts its performance to the Wiener-type (WAS) and Fourier amplitude (FAS) shrinkage methods. The novel feature in BEFE is that the separation is performed in the wavelet domain and involves the use of a Bayesian inference approach guided by existing theoretical power-laws in the filtered signal energy spectrum. We contrast the performance of all three methods for synthetic fractional Brownian motion ($fBm$) signals and turbulent velocity time series collected in the atmospheric boundary layer. A discussion on the advantages and disadvantages of each method is also presented, particularly when the process is not exactly an $fBm$. 
1. Introduction

There are many geophysical phenomena whose statistical properties exhibit spectral characteristics resembling $1/f$ processes. For example, the structure of turbulent water vapor concentration and velocity fluctuations in locally homogeneous and isotropic turbulent flow field (Monin and Yaglom, 1975, Frisch, 1995), mean daily relative humidity (Vattay and Harnos, 1994), precipitation (Gupta and Waymire, 1990; Kumar and Foufoula-Georgiou, 1993), streamflow (Bras and Rodríguez-Iturbe, 1993), river networks (Gupta and Waymire, 1989), clouds (Lovejoy, 1982), and many others (Taqqu, 1987) tend to exhibit spectral properties resembling $1/f$ processes. In practice, measurements of such geophysical phenomena contain observational noise that tend to contaminate or mask the $1/f$ spectral properties.

The objective of this study is to develop and test filtering methods that can efficiently separate white noise, typically associated with measurement errors, from the desired $1/f$-like geophysical process. The proposed methods are based on the premise that the underlying geophysical phenomenon being analyzed is self-similar, but is distorted by the white noise structure.

The filtering methods considered here range from simple Fourier amplitude shrinkage and Weiner-type filtering in the wavelet domain, to a newly proposed Bayesian wavelet approach. The existence of universal power laws in many geophysical phenomena makes the Bayesian approach quite suitable since the prior structure of the statistical model describing the structure of the process can be calibrated using physical
laws and does not require a priori inspection of the data (Wikle et al., 1998; Berliner et al., 1999).

One of the popular models for $1/f$-type processes is the fractional Brownian motion ($fBm(H)$), having a power spectrum decaying as $1/\omega^{2H+1}$, where $\omega$ is the frequency and $H$ is the Hurst exponent. In this study we choose $fBm$ as a theoretical model for $1/f$ processes. Such $fBm$ model is adopted in our evaluation of the filtering methods because its second order properties are well understood and reflect the second order statistics of many geophysical phenomena.

A simple method to filter noise from a $fBm$ signal involves extending the theoretical (and unique) power-law in the Fourier power-spectrum, shrinking all Fourier amplitudes not following this power-law to match the theoretical power-law, retain the phase angle for each frequency, and inverse transforming back to the time domain. In this method it is assumed that the high frequencies in the Fourier domain are predominantly affected by the noise. We contrast such a method with alternative noise separation methods in the wavelet domain that efficiently utilize the theoretical power-law as a prior-calibrator to an underlying Bayesian multiscale model. The selection of the wavelet domain is partly motivated by aptness of wavelets to address self-similarity, noise-whitening capability, unsteadiness, and intermittency of geophysical signals (Farge, 1992; Farge et al., 1992; Vergassola and Frisch, 1991; Hudgins et al., 1993; Kumar and Foufoula-Georgiou, 1994; Katul and Vidakovic, 1998; Katul and Albertson, 1999; Wickerhauser et al., 1994; and Zubair et al., 1990).

The testing of these methods is performed in two stages: The first is to artificially
contaminate a synthetic $fBm(H)$ signal with white noise but varying the signal-to-noise ratio\(^1\)(SNR) and investigate which of the these three denoising methods optimally recovers the original $fBm(H)$ signal. Additionally, the statistical structure of the residual noise is also investigated since the simulated noise used to contaminate the synthetic $fBm(H)$ signal is known. The second stage is to repeat the analysis in the first stage, but using measured turbulent velocity time series collected in the atmospheric boundary layer instead of $fBm$ signal. Details of the experimental setup are presented in Katul \textit{et al.} (1997b) and Katul and Albertson (1999). The latter test provides detailed assessments of these three methods when the geophysical process considered is not exactly self-similar and the distribution of the phase angle is not strictly random. A description of these three denoising methods is considered next.

2. Theory

As discussed in the Introduction, a measured geophysical realization of a self-similar process may be convoluted with white noise at multiple scales. Such noise is typically affiliated with measurement devices (e.g. gas analyzers as in Ray \textit{et al.} (1986) or Katul \textit{et al.} (1996)) or the logging devices. We compare three methods (Fourier Amplitude Shrinkage, Wiener-type Amplitude Shrinkage, and Bayesian Energy Fraction Estimation) to separate the signal from the noise. These methods are chosen because they encompass a wide range of filtering strategies and are outlined below.

\(^1\)SNR is defined as the ratio of standard deviations of signal and noise.
2.1. Fourier Amplitude Shrinkage Method (FAS)

Let

\[ y = f + \eta, \]  

be an additive model for the measurements \( y = (y_1, \ldots, y_N) \), \( N = 2^J \), consisting of the true self-similar signal \( f = (f_1, \ldots, f_N) \) and an error (e.g., instrumentation noise) \( \eta = (\eta_1, \ldots, \eta_N) \), where \( N = 2^J \) and \( J \) is a positive integer.

Let \( F_y(\omega) \), \( F_f(\omega) \), and \( F_\eta(\omega) \) denote the discrete Fourier transform of \( y \), \( f \), and \( \eta \). Because the noise power spectra \( |F_\eta(\omega)|^2 \) is “flat”, it follows

\[ |F_y(\omega)|^2 \geq |F_f(\omega)|^2. \]

Starting with \( |F_f(\omega)|^2 = C_f^2 \omega^{-2H-1} \), \( |F_y(\omega)|^2 \) can be shrunk to estimate \( |F_f(\omega)|^2 \) for all \( \omega \) in which \( |F_y(\omega)|^2 \) exceeds \( C_f^2 \omega^{-2H-1} \), where \( C_f \) is a constant given by \( C_f = |F_y(\omega_0)| \), \( \omega_0 \) is the frequency for which \( |F_y(\omega)|^2 \) departs from \( C_f^2 \omega^{-2H-1} \), and \( \hat{g} \) is defined as an estimator of \( g \). The estimated amplitude is

\[ |\hat{F}_f(\omega)| = \begin{cases} |F_y(\omega)|, & \text{when } \omega \leq \omega_0 \\ C_f \omega^{-H-\frac{1}{2}}, & \text{when } \omega > \omega_0 \end{cases} \]

while the phase angle remains unchanged for all \( \omega \),

\[ (\forall \omega) \quad \arg(\hat{F}_f(\omega)) = \arg(F_y(\omega)). \]

In essence, the filtering model assumes that noise perturbs the magnitude but not the argument when \( F_f(\omega) \) is estimated from \( F_y(\omega) \). With these approximations for the amplitude and phase angle, the denoised signal is recovered as \( \hat{f} = F^{-1}(\hat{F}_f(\omega)) \), where \( F^{-1} \) is the inverse Fourier transform.
2.2. Bayesian Energy Fraction Estimation Method (BEFE)

Bayesian Energy Fraction Estimation Method (BEFE) is derived from results of statistical inference in the wavelet domain using Bayesian modeling strategies. Statistical inference is an inversion process in which the causes are derived from the effects, taking into account the stochastic nature of the model and influence of unexplained, random factors. Bayes’ theorem formalizes this inversion by incorporating information about the model from the observations and producing a probability distribution that is a compromise between our prior knowledge and belief, e.g., existence of power laws, and the measurements. The Bayesian approach provides the unique coherent paradigm for the aforementioned inversion. If our a priori information about the model is expressed in terms of a probability distribution, then Bayes’ theorem gives a recipe for updating the prior distribution to the posterior distribution by accounting for the measurements. Bayes theorem combines the likelihood of the measurements $y$, modeled via probability distribution $f(y|\theta)$, and prior on $\theta$, $\pi(\theta)$, to obtain the posterior $\pi(\theta|y)$ via,

$$\pi(\theta|y) = \frac{f(y|\theta) \cdot \pi(\theta)}{m(y)}, \quad (2)$$

where $\theta$ is object of inference (in our study, the theoretical signal component), and $m(y) = \int f(y|\theta) \cdot \pi(\theta)d\theta$ is the marginal distribution of $y$.

In BEFE, $y$ is first transformed into the wavelet domain by a discrete wavelet transformation. At each decomposition level (i.e., scale) the cumulative energy of transformed measurements is then computed. The essence of the proposed Bayesian estimation method is to incorporate information about the energy decay from the
theoretical power law to split each wavelet coefficient into the signal and noise components. This coefficient splitting is based on the ratio of such estimated energy and the computed total energy at each level. Once the signal component is generated at all scales, the inverse transformation is applied to recover the desired signal. The choice of the wavelet domain is motivated by the decorrelation ("whitening") property of orthogonal wavelet transformations (Flandrin, 1992 and Walter, 1994). This decorrelation allows simple statistical models for which the independence assumption is reasonable. The choice of a Bayesian paradigm is motivated by its aptness to incorporate exact power laws in the energy spectrum of geophysical signals via necessary prior elicitations. Given the likelihood on an arbitrary wavelet coefficient and the prior on its signal part, Bayes’s theorem produces the posterior distribution. The observed wavelet coefficients are split to signal and noise parts by utilizing the derived posterior distribution.

It is assumed that $f$ is a sampled fractional Brownian motion ($fBm(H)$) with zero mean and autocovariance function

$$
\gamma(t, s) = \frac{\sigma^2}{2} (|t|^{2H} + |s|^{2H} - |t - s|^{2H}),
$$

where $s$ and $t$ are time instants and parameters $H$ and $\sigma^2 = \Gamma(1-2H) \frac{\cos \pi H}{\pi H}$ characterize the theoretical spectral properties of the geophysical process, and $\Gamma(x)$ is the Gamma function given by $\int_{0}^{\infty} t^{x-1} e^{-t} dt$. The standard Brownian motion is recovered for $H = 1/2$.

Let $w = \theta + \epsilon$ be the wavelet image of the model (1) where the components in $w$ are organized into a standard wavelet tree-like structure in a multiresolution format
(Mallat, 1989). The level-subvectors are given in the direction of coarsening details,

\[ w = (d^{(J-1)}, d^{(J-2)}, \ldots, d^{(J-K)}, c^{(J-K)}) \quad (4) \]

We assume that the level of finest details, resulting from a single application of the high-pass wavelet filter to the data \( y = c^{(J)} \), is \( d^{(J-1)} \). The vector \( d^{(J-k)} \), associated with the level \( k \), \( 1 \leq k \leq K \leq J \) in the decomposition (4) has \( m_k = 2^{J-k} \) wavelet coefficients.

For a fixed \( k \), \( 0 \leq k \leq K \), we propose the following multivariate normal \((\mathcal{MVN})\), hierarchical model on each of \( d^{(J-k)} \),

\[
[d|\theta, \sigma^2] \sim \mathcal{MVN}_{m_k}(\theta, \sigma^2 I), \quad (5)
\]

\[
[\theta|\tau_k^2] \sim \mathcal{MVN}_{m_k}(0, \tau_k^2 I), \quad (6)
\]

where \( d \) and \( \theta \) denote \( d^{(J-k)} \) and its signal part \( \theta^{(J-k)} \), \( \sigma^2 \) and \( \tau_k^2 \) are the variances of the error and unknown signal part from the level \( k \), and \( I \) is the identity matrix of dimension \( m_k \). In Appendix A we briefly discuss why such independence models such as normal likelihood-normal prior with covariance matrices multiples of \( I \) are suitable for such an approach. Analogous models are not adequate in the time domain because of strong and persistent autocorrelation.

From (5) and (6), Bayes’ theorem yields the posterior on the signal part in level \( k \) as

\[
[\theta|d] \sim \mathcal{MVN}_{m_k}(\frac{\tau_k^2 d}{\sigma^2 + \tau_k^2}, \frac{\sigma^2 \tau_k^2}{\sigma^2 + \tau_k^2} I) \quad (7)
\]

We note that replacing each vector of coefficients \( d \) by the posterior mean \( \frac{\tau_k^2 d}{\sigma^2 + \tau_k^2} \) amounts to traditional Wiener filtering in the wavelet domain.
Even though we derived the posterior distribution for each component of the signal vector, our coefficient splitting will only depend on the cumulative energies in each level. Therefore, using (7), we make an inference on the level-wise energies, $||\theta||^2$. Standard assumption in Bayesian estimation is that the underlying loss function (i.e., the penalty for not estimating the unknown parameter exactly) is the squared error. We depart from this standard assumption and adopt the weighted quadratic loss (see Saxena and Alam, 1982; Robert, 1995)

$$L(\theta, \delta) = w(||\theta||) \cdot (\delta - ||\theta||^2)^2,$$

(8)

where the weights are $w(||\theta||) = \frac{1}{2||\theta||^2 + m_k}$. The Bayes rule for estimating $||\theta||^2$ is (see Robert, 1995; page 70)

$$\delta(||\theta||) = \frac{E^{\theta|d}(w(||\theta||) \cdot ||\theta||^2)}{E^{\theta|d}w(||\theta||)} = \frac{E^{\theta|d}(||\theta||^2/(2||\theta||^2 + m_k))}{E^{\theta|d}(1/(2||\theta||^2 + m_k))},$$

(9)

where $E^{\theta|d}$ is the expectation with respect to the posterior (7).

The rule (9) can be obtained by either numerical integration (involving Bessel functions) or by a simple Monte Carlo simulation technique. In this study we used the latter and its derivation is presented in Appendix B.

Our choice of the loss function in (8) is motivated by the following pragmatic arguments. By rewriting the loss function in (8) as

$$\frac{1}{m_k \cdot (1 + 2||\theta||^2/m_k)} \cdot (\delta - ||\theta||^2)^2,$$

notice that the standard squared-error loss is divided by $m_k \cdot (1 + 2||\theta||^2/m_k)$. The positive factor $1 + 2||\theta||^2/m_k$ considers the average level energy, $||\theta||^2/m_k$, such that an
error made in a level with “energetic” coefficients is less penalized. The factor $m_k$ can be informally thought of as a “dimensionality curse” for which the errors of the same absolute magnitude are penalized less in higher dimensions. Additionally, Saxena and Alam (1982) proved that the rule in (9) is minimax, yet another desirable statistical property of (8). Given $\delta(||d||)$, each coefficient $d$, from the level $k$, is split and the signal part is replaced by its shrunk version $d^* = \sqrt{\delta(||d||)/||d||^2} \cdot d$, where $\delta(||d||)$ is given by (9).

The model in (5, 6) is capable of incorporating second order statistics via suitable selection of the hyperparameter, $\tau_k^2$. By a priori knowing the SNR and level-wise energies of the noised signal the hyperparameter $\tau_k^2$ is elicited to be proportional to $2^{(2H+1)k}$. This is in compliance with the power law that states that the logarithms of average level-wise energies linearly decay when levels vary from coarse to fine. The log-energy decay rate is $-2H - 1$. In the case of $fBm$, elicitation of $\tau_k^2$ is in accordance with Flandrin’s (1992) results, (see the equation (A2), in Appendix A).

2.3. Wiener-Type Amplitude Shrinkage Method (WAS)

As in BEFE, the Wiener-type Amplitude Shrinkage method replaces each wavelet coefficient $d$ from level $k$ by $d^* = \frac{\tau_k^2}{\tau_k^2 + \sigma^2} \cdot d$. This shrinkage is equivalent to Wiener filtering of the signed energies, (i.e., $\operatorname{sign}(d) \cdot d^2$.) In terms of signed energies this shrinkage method has a Bayesian interpretation in which both the model and the prior on its location are Gaussian with variances $\sigma^2$ and $\tau_k^2$, respectively. The WAS method used here departs from traditional Wiener-type shrinkage methods in which the model
is assumed to apply directly on the wavelet coefficients and not on the signed energies, see Simoncelli (1996) and Wornell (1996).

3. Applications to Synthetic and Measured Signals

As stated earlier, the spectral characteristics of many self similar geophysical flow variables resemble those of an $fBm(H)$. For example, the Kolmogorov $K41$ power-laws in atmospheric turbulence for velocity and other scalar quantities, e.g. water vapor or temperature (see Katul et al., 1997a), is given by

$$E(\omega) \propto \omega^{-5/3},$$

which describes well the spectral density function in the inertial subrange (where turbulence is locally isotropic and homogeneous as in Kolmogorov, 1941)). For this purpose, we explore synthetic signals with Hurst exponent $H = 1/3$ corresponding to the log-spectral decay of turbulence for illustration but assess the robustness of our findings to other $H \in [0, 1]$.

In Figure 1 a simulated $fBm$ with $H = 1/3$ is shown, panel (a), along with measured longitudinal velocity ($u$) time series collected in the atmospheric boundary layer, panel (b). As evidenced from panels (c) and (d), the Fourier power spectrum in both signals exhibit an extensive $\omega^{-5/3}$ power law. Hence, it is instructive to illustrate the filtering algorithms using both an $fBm$ signal with $H = 1/3$ and the measured $u$ in which the power-law is approximate and the phase angle is not random.
3.1. Applications to $fBm$ Processes

The $fBm$ signal of size $N = 32768$ was added to a Gaussian noise with zero mean and variance determined by preassigned $SNR$ to mimic various levels of instrumentation noise as shown in Figure 2. In order to investigate the robustness of the signal separation methodology with respect to the noise magnitude, we applied the three filtering methods to all the time series shown in Figure 2.

Figure 3 shows wavelet energies for the contaminated signals with $SNR = 1, 5, 10,$ and $\infty$, and their filtered counterparts $(\hat{f})$ by FAS, WAS, and BEFE. It is clear that all three algorithms produced the same energy spectrum for $(\hat{f})$ despite one decade of $SNR$ variations, with slight deviations at the finest scale of WAS when $SNR = 1$. Such agreements in spectral properties between $\hat{f}$ and $f$ is not surprising given that all three methods aim at recovering the signal component conditioned on known power-laws or pre-assigned $H$. To assess the optimality of the estimated $fBm$ signal $\hat{f}$ we consider the mean-square error, $MSE = \frac{1}{N} \| \hat{f} - f \|^2$, where $N$ is, as before, the length of $f$.

The $MSE$ is calculated for all three filtering methods and for various $SNR$ and shown in Figure 4. Both linear and semi-log representation are used to highlight differences between methods at small and large $SNR$. Among all three methods, WAS best reproduced $f$ for one-decade variation in $SNR$. We note that for the BEFE and WAS methods, the Haar wavelet was used throughout; however, we repeated the same analyses for a wide range of bases functions including Symmlets and Daubechies wavelets with vanishing moments ranging from 2 to 10 and found that WAS and BEFE
still out-performed FAS in terms of minimizing \( \text{MSE} \). Hence, when applied to \( 1/f \)-type signals, these wavelet-based filtering methods are not sensitive to the choice of the analyzing wavelet.

Another measure of method performance is the recovery of the statistical structure of \( \eta \), particularly its normality and “white-noise” spectral properties. The normality is assessed by inspecting the quantile-quantile (or \( Q - Q \)) plots of estimated \( \hat{\eta} = y - \hat{f} \) against \( \eta \) for all three methods. All three methods reproduced the normality of \( \hat{\eta} \) well for a wide range of \( \text{SNR} \) as expected, given the normality of the components (not shown). This test becomes more critical for the turbulence signal which is not Gaussian.

The spectral properties of \( \hat{\eta} \) in all three methods are assessed by comparing their autocorrelation functions (\( \text{acf} \)) with the \( \text{acf} \) of \( \eta \). It is evident from Figure 5 that FAS best reproduced the spectral properties of \( \eta \), while both WAS and BEFE estimated \( \eta \) appear to be correlated up to lag 5. The strong decorrelation obtained by FAS is not surprising for an \( fBm \) process in which the phase angle, by definition, is independent random normal. The existence of finite (and significant) correlation in \( \eta \) estimated by WAS and BEFE is attributed to the poor resolution of time-localized orthonormal wavelet transforms in the frequency domain. It is for this reason, among others, that we repeated the same analysis described above on \( u \), in which the phase angle is not randomly distributed.
3.2. Applications to Velocity Measurements in the Atmospheric Boundary Layer

The measured \( u \) signal of size \( N = 32768 \) and shown in Figure 1 was again added to a Gaussian noise with zero mean and variance determined by preassigned \( SNR \) to mimic various levels of instrumentation noise. To again investigate how robust is the signal separation methodology to the noise magnitude, we applied the three filtering methods to the “contaminated” \( u \) time series.

Figure 6 shows wavelet energies for the contaminated \( u \) signals with \( SNR = 1, 5, 10, \) and \( \infty \), and their filtered counterparts by FAS, WAS, and BEFE. In analogy to the \( fBm \) analysis, it is clear that all three algorithms produced the same energy spectrum for the filtered signals despite one decade of \( SNR \) variations.

To assess the optimality of the recovered \( u \) signal, \( \hat{u} \), the \( MSE \) is also used and the results are shown in Figure 7. Both real and semi-log representation are again used to amplify differences between the three methods at small and large \( SNR \). Among all three methods, BEFE best reproduced the measured \( u \) for one-decade in \( SNR \) variation when compared to FAS and WAS.

As stated earlier, another performance measure is the recovery of the statistical structure of \( \eta \), particularly its normality and spectral properties. In Figure 8, the normality is again assessed by displaying the \( Q - Q \) plots of estimated \( \eta \) by all three methods against the specified \( \eta \) for \( SNR = 1, 5, \) and \( 10 \). All three methods reproduced the normality of \( \eta \) well when \( SNR \) was close to unity as evidenced from Figure 8.
However, with increasing SNR, FAS best reproduced the marginal distribution of the noise, but not its correlation properties discussed next.

To assess the spectral properties of residuals, the acf of all three estimated $\hat{\eta}$ are shown in Figure 9. It is evident from Figure 9 that the residual noise from BEFE and WAS was much more decorrelated than that from FAS. The retention of the phase angle in FAS “leaked” much of the $u$-autocorrelation to $\hat{\eta}$. Interestingly, the estimated $\hat{\eta}$ appear to be correlated up to lag 5 in BEFE and WAS, similar to the $fBm$ analysis shown in Figure 5, further suggesting that such finite correlation is not dependent on the geophysical process analyzed but on the inadequate resolution of time-localized orthonormal wavelet transforms in the frequency domain.

4. Conclusions

We proposed and tested a multiscale Bayesian model (BEFE) for separating a $1/f$ like signal from inherent instrumentation noise and contrasted its performance to the Wiener-type (WAS) and Fourier amplitude (FAS) shrinkage methods. By and large, the difficulty in such filtering is that signal and noise are convolved at multiple scales. In the BEFE method, the separation is performed in the wavelet domain and involves the use of a Bayesian inference approach guided by existing theoretical power-laws in the filtered signal energy spectrum. We demonstrated the robustness of this methodology to large variations in the signal to noise ratio using synthetic $fBm$ signals generated from superposition of a fractional Brownian motion and a preset noise component. The analysis specifically demonstrated the following:
• All three methods recovered the spectral properties of the original time series reasonably well for a wide range of SNR, except WAS for low SNR and for the finest level of details.

• The mean-squared error (MSE) between original and filtered signal were minimized more by BEFE and WAS when compared to FAS. However, the MSE for all three methods are within 25% of each other for small SNR.

• For the fBm process, FAS well reproduced the spectral properties of the noise, particularly the decorrelation property. This, in part, is due to the fact that fBm processes have no “organized” phase angle.

• For the measured turbulent velocity time series (u) for which the phase angle is not random, the spectral characteristics of the noise were poorly reproduced by FAS. However, FAS preserved the Gaussian property of the noise better than BEFE and WAS.

• All in all, differences between BEFE and WAS were minor except for the spectral properties at low SNR and at the finest levels of details in which the filtered signal spectra by BEFE is more consistent with the spectra of the uncontaminated signal.

The broader implication is that each method has strengths and weaknesses depending on the application at hand. For fBm processes, FAS’s performance was comparable in terms of MSE and superior to WAS and BEFE in terms of noise
decorrelation. Hence, for synthetic processes for which the phase angle is random (e.g., $fBm$), we do not find any clear advantages in utilizing Bayesian estimation strategies such as WAS and BEFE.

However, for real geophysical signals in which self-similarity is approximate and the phase angle contains signatures of the process, the BEFE method is recommended. The BEFE out-performed WAS and FAS in terms of minimizing the relative $MSE$ and decorrelating the residuals. This good performance of BEFE relative to FAS is due to the ability of the Bayesian paradigm to incorporate uncertainty about the signal to be estimated in light of existing theoretical power laws. Relative to WAS, it appears that weighted loss function used to define BEFE is more tuned to the geophysical signals when compared to the standard mean square error used in WAS.

Appendix A: Whitening property of wavelet transformations

In this appendix we discuss why simple models such as normal likelihood-normal prior well describe the synthetic $fBm$. The wavelet images of a fractional Brownian motion ($fBm$) with the Hurst exponent $H$ are simple due to an intrinsic self-similarity. Consider $E(d_{jk}d_{jk'})$, where $d_{jk}$s are wavelet coefficients of the $fBm$ signal. The asymptotic correlation structure of detail coefficients is governed by the behavior of

$$
|\Psi(\omega)|^2/|\omega|^{2H+1}
$$

at the origin (Flandrin, 1992), where $\Psi(\omega)$ is the Fourier integral of the decomposing wavelet given by $\int_{\mathbb{R}} \psi(t)e^{-i\omega t} dt$. 
When the decomposing wavelet possesses \( N_m \) vanishing moments, (A1) behaves as \( \omega^{-2N_m-H-1} \). A perfect decorrelation is achieved if and only if \( N_m = H + 1/2 \), which is the case for the pair of Brownian motion \( (fBm(1/2)) \) and Haar’s wavelet. If \( N_m \neq H + 1/2 \), the covariances between coefficients \( d_{jk} \) and \( d_{jk'} \) are nonzero, but their magnitudes decay rapidly (as \( O(|2^{-j}(k - k')|^{2(H-N_m)}) \), when \( |2^{-j}(k - k')| \to \infty \)).

As earlier discussed, the decorrelating property of wavelet transformations allows us to employ the prior in (6) at level \( j \), with the covariance matrix modeled as \( \tau^2 \cdot I_m \), where

\[
\tau^2 = C_\psi(H)2^{-j(2H+1)}. \tag{A2}
\]

The constant \( C_\psi(H) \) depends only on the wavelet basis and the Hurst exponent, \( H \), but not on the level \( j \).

**Appendix B: Calculation of BEFE Bayes rule by Monte Carlo simulation scheme**

As stated earlier, an analytic expression for \( \delta = \delta(\|d\|) \) in (9) does not exist. Here we describe a convenient form of \( \delta \) suitable for a fast Monte Carlo (MC) based simulation.

Since,

\[
\left[ \frac{\|\theta\|^2}{\sigma^2 + \tau_k^2} [d] \right] \sim \lambda^2_m(\lambda), \tag{B1}
\]
where

\[ \lambda = \frac{||\tau_k d||^2}{\sigma^2 r_k^2 (\sigma^2 + r_k^2)} = \frac{\tau_k^2}{\sigma^2} \cdot \frac{||d||^2}{\sigma^2 + \tau_k^2}, \quad (B2) \]

an MC approximation is

\[ \hat{\delta}(||d||) = \frac{\sum_{i=1}^{M} \eta_i}{\sum_{i=1}^{M} \frac{1}{2 \eta_i + m_k}}, \quad (B3) \]

where

\[ \eta_i \equiv \frac{d \cdot \tau_k^2}{\sigma^2} (X^2 + Z), \quad i = 1, \ldots, M \quad (B4) \]

\( X \sim \mathcal{N}(\sqrt{\lambda}, 1) \), and \( Z \sim \chi^2_{m_k - 1} \). In the above, \( \equiv \) means equality in distribution, \( \mathcal{N} \) and \( \chi^2_n \) are symbols for normal distribution and chi-square distribution with \( n \) degrees of freedom, respectively.

This estimator on the level energies will be used in the energy splitting algorithm of wavelet coefficients.

**Acknowledgments.** This research was supported by the National Science Foundation (NSF) Grants DMS-96-26159, EAR-99-03471, and BIR-95-12333 at Duke University. The Wavelab-based MATLAB programs used in the filtering calculations can be obtained from the authors on request.
References


Katul, G., Hsieh, C., Sigmon, J. (1997a). Energy-inertial scale interaction for temperature and


B. D. Vidakovic, Institute of Statistics and Decision Sciences, Duke University, Durham, NC 27708-0251. (e-mail: brani@isds.duke.edu)

G. G. Katul, School of the Environment, Duke University, Durham, NC 27708-0328. (e-mail: gaby@duke.edu)

J. D. Albertson, Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22903. (e-mail: jda4h@virginia.edu)

Received January 25, 2000; revised May 27, 2000; accepted July 23, 2000.
3. Figure Captions

Figure 1. (a) Fractional Brownian motion time series with $H = 1/3$, $fBm(1/3)$, (b) Measured longitudinal velocity fluctuations standardized to have zero mean and unit variance, (c) Fourier power spectrum of $fBm(1/3)$ process from (a), and (d) same as (c) but for the time series in (b). The dashed line in (c) and (d) is $\omega^{-5/3}$ as in K41.

Figure 2. (a) Fractional Brownian motion time series with $H = 1/3$ and with no noise perturbation (i.e. $SNR=\infty$), (b) Same as (a) but contaminated with white normal noise $\epsilon$ producing $SNR = 10$, (c) Same as (b) but for an $SNR = 5$, (d) Same as (c) but for an $SNR = 1$.

Figure 3. (a) Cumulative Haar-wavelet energies for the time series shown in Figure 2 as functions of the scaling level $k$ for $SNR = 1$, 5, 10, and $\infty$, and the cumulative energies of the filtered signals by FAS shown as connected dots. The theoretical power-law corresponding to $H = 1/3$ is shown as dashed, (b) Same as (a) but for the filtered signals by WAS, (c) Same as (a) but for the filtered signals by BEFE.

Figure 4. The variation of the scaled mean-squared error ($MSE \times SNR$) of the $fBm(1/3)$ filtered signals by FAS ($\ast$), WAS ($\circ$), and BEFE ($+$) as a function of $SNR$. Both linear (top) and semi-log (bottom) axis representations are used for contrasting $MSE \times SNR$ amongst the three methods for small and large $SNR$.

Figure 5. The variation of the autocorrelation function ($acf$) for the noise $\hat{\eta}$ estimated by applying FAS ($\ast$), WAS ($\circ$), and BEFE ($+$) to the synthetic time series $y$ shown in Figure 2 (c) (i.e. $SNR=5$). The $acf$ of the original white noise used in Figure 2 (c) is shown in $\circ$.

Figure 6. Same as Figure 3 but for $u$. 
**Figure 7.** Same as Figure 4 but for $u$.

**Figure 8.** Quantile-Quantile plots of the estimated noise ($\hat{\eta} = y - \hat{f}$) for all three filtering methods, where $y$ is the “contaminated” time series of the measured velocity and $\hat{f}$ is the filtered signal by either FAS, WAS, or BEFE method, for $SNR = 1, 5, \text{ and } 10$. The dashed line corresponds to a normal distribution.

**Figure 9.** Same as Figure 5 but for $u$. 
FIGURE 1

(a) fbm(1/3)

(b) $u/\sigma_u$

(c) $E_{fbm}$

(d) $E_u$

frequency (Hz)
FIGURE 2

(a) $\text{fbm}(1/3)$

(b) $\text{fbm}(1/3) + \varepsilon$ (SNR=10)

(c) $\text{fbm}(1/3) + \varepsilon$ (SNR=5)

(d) $\text{fbm}(1/3) + \varepsilon$ (SNR=1)
FIGURE 3

(a) FAS

(b) WAS

(c) BEFE
FIGURE 4

The figure compares the performance of different algorithms (FAS, WAS, BEFE) in terms of mean squared error (MSE) multiplied by signal-to-noise ratio (SNR). The graphs show the relationship between MSE x SNR and SNR on a logarithmic scale. The top graph displays the linear scale, while the bottom graph uses a logarithmic scale for both axes.
FIGURE 5

The figure shows a plot with the lag index on the x-axis and the autocorrelation function (ACF) on the y-axis. The points represent different series labeled as original, FAS, WAS, and BEFE. The y-axis values range from -0.25 to 0.05, and the x-axis values range from 2 to 10.
FIGURE 6

(a) FAS

(b) WAS

(c) BEFE
FIGURE 7

![Graph showing the relationship between mse x snr and snr for different methods: FAS, WAS, and BEFE.](graph.png)