Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2007, Article ID 63219, 7 pages
doi:10.1155/2007/63219

Research Article

Estimation of Spectral Exponent Parameter of $1/f$ Process in Additive White Background Noise

Süleyman Baykut,[1] Tayfun Akgül,[1,2] and Semih Ergintav[2]

[1] *Department of Electronics and Communications Engineering, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey*
[2] *TÜBİTAK Marmara Research Center, Earth and Marine Sciences Institute, 41470 Gebze, Kocaeli, Turkey*

Received 29 September 2006; Revised 5 February 2007; Accepted 29 April 2007

Recommended by Abdelhak M. Zoubir

An extension to the wavelet-based method for the estimation of the spectral exponent, γ, in a $1/f^\gamma$ process and in the presence of additive white noise is proposed. The approach is based on eliminating the effect of white noise by a simple difference operation constructed on the wavelet spectrum. The γ parameter is estimated as the slope of a linear function. It is shown by simulations that the proposed method gives reliable results. Global positioning system (GPS) time-series noise is analyzed and the results provide experimental verification of the proposed method.

1. INTRODUCTION

$1/f^\gamma$ processes, also referred to as self-similar processes, are observed in many diverse fields and have gained importance in various signal processing applications from geophysical records to biomedical signals, from economical indicators to internet network traffic [1–6]. $1/f^\gamma$ processes are generally characterized by a power-law relationship in the frequency domain, that is, the empirical (or measured) power spectra of such processes are considered to be of the form [1]

$$S_x(\omega) \sim \frac{\sigma_x^2}{|\omega|^\gamma} \tag{1}$$

over some decades of frequency ω, where σ_x^2 is a finite nonzero constant and γ is the so-called spectral exponent (or sometimes it is called the self-similarity parameter). In general, $1/f^\gamma$ processes can be modeled by fractional Gaussian noise (fGn) and fractional Brownian motion (fBm). fBms are zero-mean, normally distributed, nonstationary random processes with $1 < \gamma < 3$, whereas fGns are zero-mean, normally distributed, stationary incremental processes of fBms with $-1 < \gamma < 1$ [1, 7]. $1/f^\gamma$ processes are also named as colored noise. White noise having a flat spectrum is the special case of colored noise, where the spectral exponent $\gamma = 0$. For $\gamma = 1$, it is called flicker noise and for $\gamma = 2$, it is known as classical Brownian motion (random walk process).

The importance of such processes is due to the fact that they can be modeled by a single parameter γ which can be used for diagnosis, prediction, and control purposes in many applications. Therefore, an accurate estimation of γ is needed. However, estimation of this parameter is not often straightforward, especially when the data is considered to be corrupted by additive white noise. For this case, the measured power spectrum $S_x(\omega)$ is

$$S_x(\omega) \sim \frac{\sigma_x^2}{|\omega|^\gamma} + \sigma_g^2, \tag{2}$$

where σ_g^2 is the variance of the white noise. Now, estimation of the underlying characteristics of such processes becomes challenging simply because there is a single equation with more than one unknown parameter.

Although it may seem straightforward to separate a $1/f^\gamma$-type process from additive white noise, one has to know the domination of this process among frequency regions exactly. There are several methods for γ estimation from noisy measurements [3–5, 8–11]. Among them, the conventional ones attempt to estimate the parameters of the processes in the spectral domain [3, 4]. In this approach, σ_x^2, σ_g^2, and γ are extracted from the estimated noise power spectral density using a least-square fit algorithm to the spectrum in (2). The spectral density of the $1/f^{\gamma}$ process rapidly decays towards the higher frequency regions so that the white noise spectrum tends to dominate the rest of the spectrum which makes a reliable estimation of γ difficult.

There are also some other methods such as approximate and exact maximum-likelihood estimation methods in the time—[3–5] and in the wavelet—[8, 9, 11] domains. In [3–5], a maximization of a likelihood function is suggested in the time domain. This approach, however, is complex (i.e., matrix inversion is required) and time consuming (i.e., an iterative technique). In [8], a parameter estimation algorithm for $1/f^{\gamma}$ process in white background noise based on iterative maximum likelihood in wavelet domain is proposed. This method, when compared with other methods, is considered to have relatively low computational complexity. In [11], a modified version is utilized using the discrete wavelet transform with Haar basis.

In this paper, a simple and practical extension to the wavelet-based γ estimation method [8] is proposed, where γ is estimated directly (without iteration) by using a differentiation operation in the wavelet domain.

The paper is organized as follows. First, the wavelet-based γ estimation technique is briefly summarized, and then a simplified version is proposed. The effectiveness of this extension is examined using synthetic data with known parameters. The method provides promising results and GPS time series noise is analyzed to provide experimental verification.

2. WAVELET-BASED γ ESTIMATION IN THE PRESENCE OF WHITE NOISE

Wavelet-based γ estimation method relies on calculating the wavelet coefficients and investigating if the variances of the wavelet coefficients follow the power-law relationship [1]. The method utilizes the orthonormal wavelet transform of a process $x(t)$ to estimate the spectral exponent

$$x_n^m = \int_{-\infty}^{\infty} x(t)\psi_n^m(t)dt. \tag{3}$$

Here, x_n^m are the wavelet transform coefficients of the signal $x(t)$, n and m are the translation (location) and dilation (scale) indices, respectively. $\psi_n^m(t)$ are the normalized dyadic dilations and integer translations of the mother wavelet $\psi(t)$ which is $\psi_n^m(t) = 2^{m/2}\psi(2^m t - n)$. The wavelet transform acts as the whitening process for a $1/f^{\gamma}$ process where the corresponding wavelet coefficients appear to be zero-mean, uncorrelated, or weakly correlated (within and along scales) random variables. The variances of the uncorrelated or weakly correlated wavelet coefficients along scales satisfy a power-law relationship [1]

$$\text{var}\{x_n^m\} = \sigma^2 2^{-\gamma m}, \tag{4}$$

where σ^2 is a positive real constant which is proportional to σ_x^2. In the labeling scheme used, higher scales are related to the higher (and wide) frequency regions, whereas lower scales are related to the lower (and narrow) frequency ranges, respectively. Taking base-2 logarithm of both sides of (4) yields a straight line whose slope is the estimated γ parameter

$$\log_2(\text{var}\{x_n^m\}) = c - \gamma m. \tag{5}$$

Here, c is a constant equal to $\log_2(\sigma^2)$.

If the corresponding process is corrupted by additive white noise, γ cannot be estimated by a linear fit in linear-log scale. Consider the noisy process $r(t)$ as the superposition of colored ($x(t)$) and white ($g(t)$) noise

$$r(t) = x(t) + g(t). \tag{6}$$

The wavelet coefficients of $r(t)$ are obtained from (3):

$$r_n^m = \int_{-\infty}^{\infty} [x(t) + g(t)]\psi_n^m(t)dt. \tag{7}$$

In this study, the discrete dyadic wavelet transform is used to obtain the wavelet coefficients, where the number of scales M is related with the data length N by $M = \log_2(N)$. Statistical independence implies that the wavelet coefficients of $r(t)$ are

$$r_n^m = x_n^m + g_n^m, \tag{8}$$

and the variances of these coefficients are related according to [8]

$$\text{var}\{r_n^m\} = \text{var}\{x_n^m\} + \text{var}\{g_n^m\}. \tag{9}$$

The power spread of white noise is uniform throughout the entire spectrum, hence the variance of the wavelet coefficients of the white noise component in each scale is equal to the variance of the white noise process (σ_g^2) [1, 8, 11]. Therefore, considering (4), (9) becomes

$$(\sigma_r^m)^2 = \sigma^2 2^{-\gamma m} + \sigma_g^2. \tag{10}$$

In [8], a log-likelihood function is expressed as a function of the unknown parameters σ^2, σ_g^2, and γ. There, three cases are analyzed: (i) all of the parameters are unknown; (ii) σ^2 and γ are unknown, σ_g^2 is known; (iii) σ^2 and γ are unknown, $\sigma_g^2 = 0$. In the first two cases, an iterative expectation-maximization (EM) algorithm is utilized to estimate the parameters. In the third case where σ_g^2 is known to be 0 (the noise free case), it is shown that the iterative EM algorithm is not needed, therefore the parameters can be estimated directly by using (4).

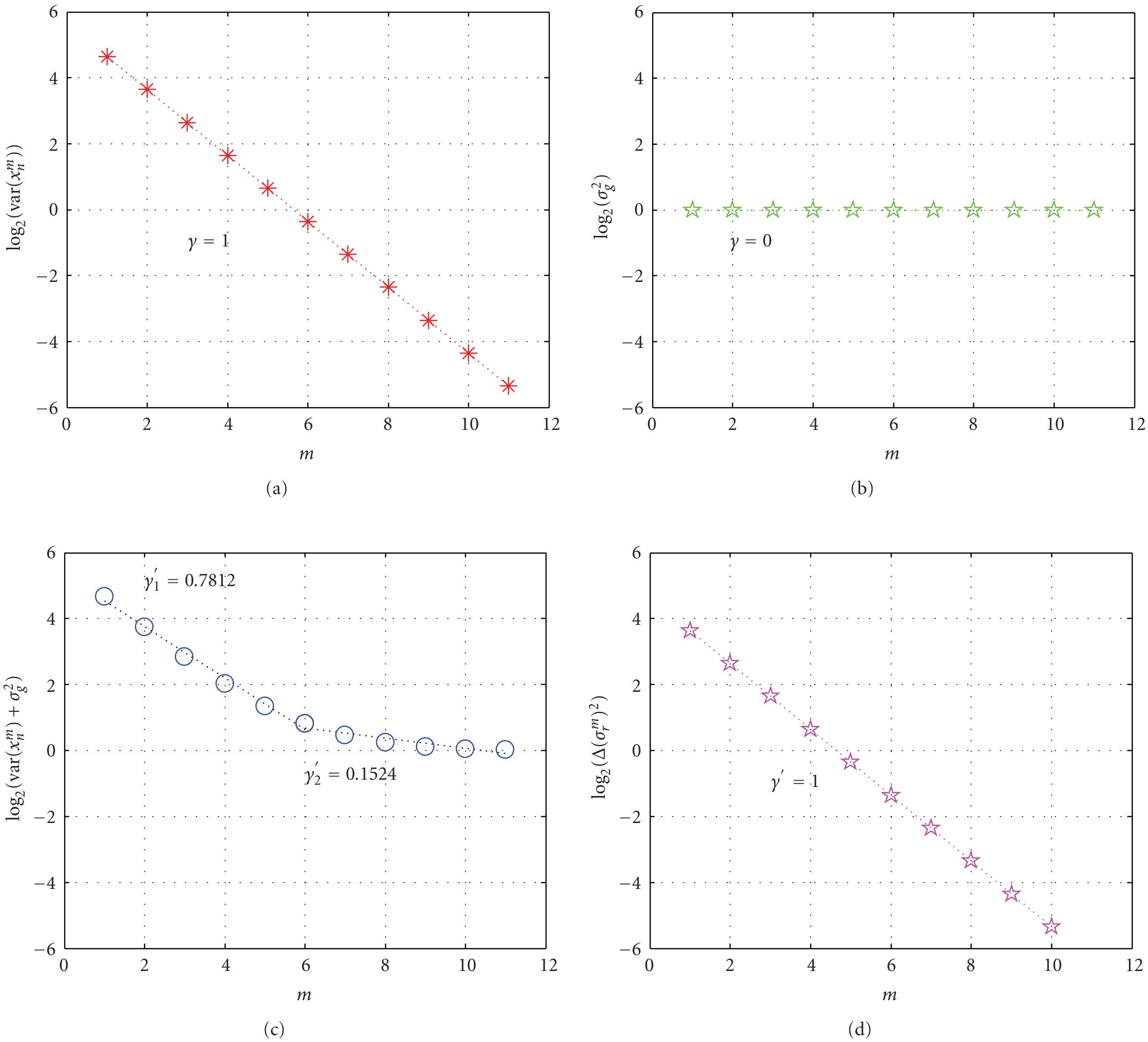


Figure 1: Base-2 logarithm of the variances of the wavelet coefficients (a) of the process $x(t)$ (with $\gamma = 1$), (b) of white noise ($\gamma = 0$), (c) of the compound signal, (d) the difference sequence of the logarithm of the variances of the wavelet coefficients in (b). (Note that the values in the figures are theoretically chosen.)

2.1. Proposed method

In this study, a simple extension is proposed to estimate the parameters directly, without using any iterative minimization technique even when white noise exists.

When $\gamma > 0$, the process has lower power at higher scales which are more affected by the white noise component than at the lower scales. Therefore, if the logarithm of both sides of (10) is taken, a curve like behavior is observed, instead of a straight line, as is illustrated in Figure 1 where simulation results are given for flicker noise ($\gamma = 1$) and additive white noise ($\gamma = 0$). Here, the data length is considered to be 2048. In Figures 1(a), 1(b), and 1(c), the base-2 logarithm of the variances of the wavelet coefficients versus the scales are plotted for flicker noise, white noise, and the compound signal, respectively. The plot of the compound signal forms a knee-like shape with broken line around scale 6 which means that the higher scales are affected more than the lower scales. If one attempts to fit a line to one of the curves in the figure, the slope does not accurately provide the true γ value. Initially, two different slopes of 0.7812 and 0.1524 are observed due to the white noise corruption. It is relevant to mention here that in Figure 1 the values on the plots are theoretically chosen. In practice, due to the limited data length, the variances of the wavelet coefficients cannot be determined exactly, therefore, the plots may not be perfect lines or curves for practical data.

In order to estimate γ, σ^2, and σ_g^2 directly and reliably, a difference sequence of (10) can be constructed along scales as

$$\triangle(\sigma_r^m)^2 = (\sigma_r^m)^2 - (\sigma_r^{m+1})^2 = \sigma^2 2^{-\gamma m} - \sigma^2 2^{-\gamma(m+1)} \quad (11)$$

which eliminates the constant noise term σ_g^2 yielding

$$\triangle(\sigma_r^m)^2 = \sigma^2(1 - 2^{-\gamma})2^{-\gamma m}. \quad (12)$$

By taking the base-2 logarithm of both sides of (12), γ and σ^2 can be estimated by fitting a linear equation (a straight line) in the least-square sense. Note that the slope here is identical to the slope of the linear equation in (5). Finally, σ_g^2 can be estimated by substituting γ and σ^2 in (10). When the above operations are applied to the compound process of Figure 1(c), the γ estimation plot becomes as in Figure 1(d). Here, the slope is 1 which is equal to the γ parameter of the underlying process (in this case, flicker noise).

Note that if one assumes the fit error as Gaussian distributed, the line fit in the least-square sense corresponds to a maximum likelihood estimation which means that the proposed line regression method becomes a line fitting problem to data corrupted by Gaussian noise.

The performance of this approach is examined by simulations in the next section below.

3. SIMULATION RESULTS

In this section, the performance of the proposed technique is examined on synthetic data. The data set is constructed as synthesized $1/f^\gamma$ processes superposed with white Gaussian noise having different SNR[1] values from -20 dB to 20 dB, with increments of 4 dB, and where γ varies from 0.5 to 2.0 with increments of 0.25. In addition, the data length is set to $N = 2^i$, where i is varied from 10 to 14 with increments of 1. Using the wavelet-based synthesis method, the data sets of $K = 100$ trials are generated for each combination of N, SNR, and γ given above. Although the wavelet-based γ estimation is shown to be empirically insensitive to the choice of the wavelet bases [8], in the simulations, among many available wavelet basis, Haar basis is used as suggested in [11].

The proposed technique is applied to each data set. Then, the mean values, the variances, and the root-mean-square (RMS) errors of the estimates are calculated. Here, RMS is defined as $\text{RMS} = \sqrt{(1/K)\sum_{k=1}^{K}(\gamma - \hat{\gamma})^2}$, where γ is the theoretical, $\hat{\gamma}$ is the estimated parameter, and K is the total number of trials.

In Figure 2(a), the RMS errors of the estimated $\hat{\gamma}$ versus SNR values are plotted for fixed data length of 4096. We see smaller RMS errors for smaller γ (≤ 1) even for poor SNR($\leq$ 0 dB.) When γ is relatively higher (≥ 1.50), the $1/f^\gamma$ process is dominated by white noise in the high frequency regions. Notice that to have better estimates, we need to observe the logarithm of the variances of the wavelet coefficients at lower scales which requires longer data. In Figure 2(a), the RMS errors of the estimates are asymptotically bounded below as the SNR increases.

The data length dependence is observed by the results given in Figure 2(b), where the RMS errors of the estimated $\hat{\gamma}$ versus the data length N are shown for a fixed SNR value of 0 dB. Here, estimation errors decrease with the increasing data length. Note that when γ increases, the dependency of the estimation method on the data length decreases. These results are similar to the ones given in [8, 9].

In Figure 3, the mean of $\hat{\gamma}$ versus SNR for fixed data length $N = 4096$ is provided. For poor SNR, the method underestimates $\hat{\gamma}$ for larger values of γ, whereas it overestimates $\hat{\gamma}$ for smaller γ. The standard deviations of the $\hat{\gamma}$ estimates decrease with the increasing SNR. Note that since there are less coefficients in the small scales, the first 3 scales are not used in the γ estimation method for statistical reasons. This limitation causes a bias of $\hat{\gamma}$ estimates for higher SNRs as is evident in Figure 3. However, when the SNR is high, the proposed technique gives similar results to the noise-free case and the wavelet-based method in [8].

4. REAL DATA ANALYSIS—GPS NOISE

There are various signal processing applications where the signals contain colored and white noise together. In some applications, estimation of noise characteristics is critical. For example, noise in electronic devices is observed to be the sum of $1/f^\gamma$ processes and white noise which are induced independently by microscopic defects and different physical mechanisms [10]. Separation of these two processes is essential for the quality and reliability of the devices determined by means of noise measurements. Another important example can be given from geophysics. It is shown that the GPS coordinate time series error can be conveniently characterized by the superposition of $1/f^\gamma$ noise (time-correlated) and white noise (time invariant) processes [2–5]. The estimations of GPS noise characteristics (i.e., the spectral exponent of the $1/f^\gamma$ process, the variance of white noise, and the mixture ratio of these processes) are crucial since they are used to obtain the model parameters which characterize the surface displacement velocity of the earth (as linear slope), seasonal motions (as periodic components), relaxation after an earthquake (as logarithmic decay), and so forth. The accuracy and the precision of the estimated model parameters depend on the accurate estimation of relevant noise characteristics [2–5].

[1] SNR is defined as the ratio of the $1/f^\gamma$ noise variance to the white noise variance ($\text{SNR} = 10\log_{10}(\sigma_x^2/\sigma_g^2)$).

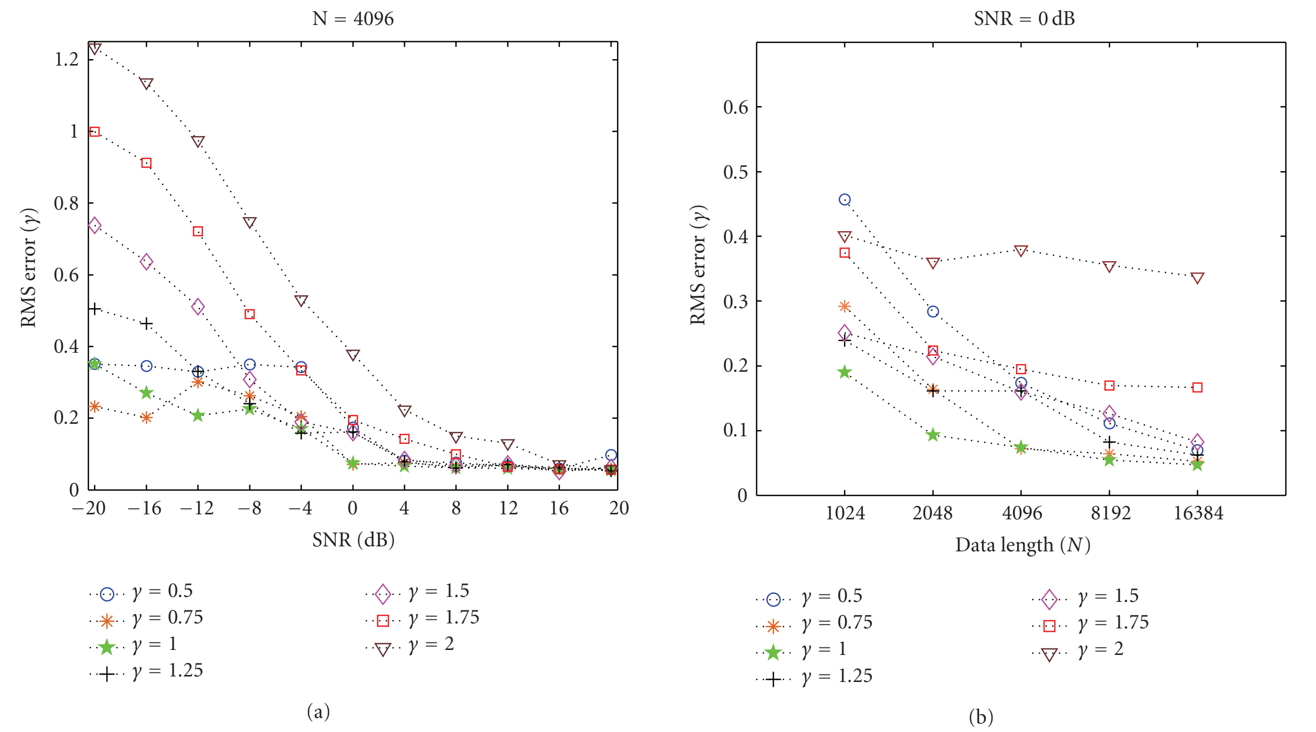


FIGURE 2: (a) The RMS errors of the estimated $\hat{\gamma}$ versus SNR for a fixed data length $N = 4096$; (b) The RMS errors of the estimated $\hat{\gamma}$ versus various data length N for SNR = 0 dB.

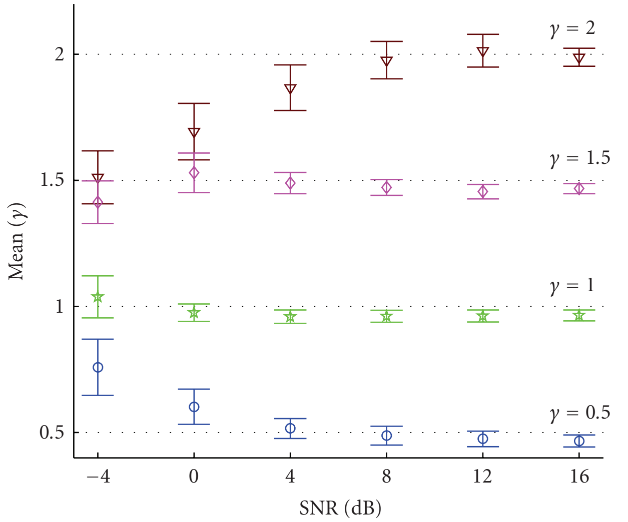


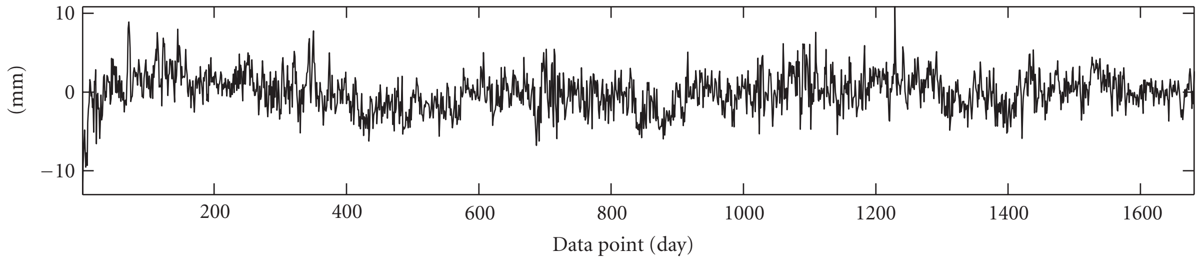
FIGURE 3: The mean values of the estimated $\hat{\gamma}$ as a function of SNR for $N = 4096$. The vertical lines indicate the standard deviations.

For real data, the GPS coordinate time series noise is analyzed. We present the analysis of the north components of GPS data obtained from *TUBI* site run by TÜBİTAK Earth and Marine Sciences Institute. After the preprocessing procedures (outlier cleaning and small gap filling as in [3–5]), the GPS noise is obtained as the difference (residual) between the model and the observed data. Then, the proposed technique is applied to the residual signal.

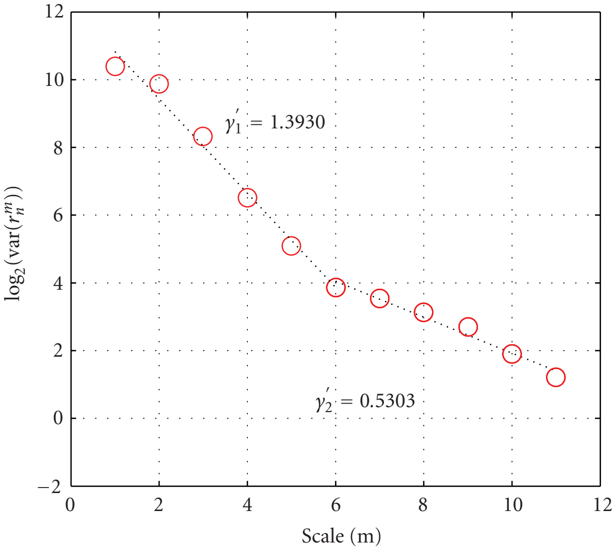
In Figure 4(a), the *TUBI* GPS noise data is plotted. In Figure 4(b), the base-2 logarithms of the variances of the wavelet coefficients along scales are given. The existence of white noise appears as a broken-line around the 6th scale. After applying the difference operator to the variances of the wavelet coefficients, a linear progression is observed as shown in Figure 4(c). The estimated $\hat{\gamma}$ value is close to 1 (to be exact, $\hat{\gamma} = 1.0194$).

5. CONCLUSIONS

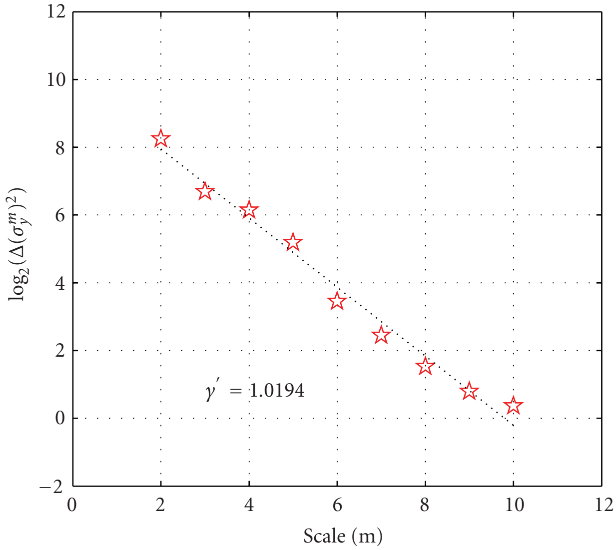
An extension to the wavelet-based spectral exponent estimation method has been proposed. The method can be used to obtain the spectral exponent of a $1/f^{\gamma}$ process with additive white noise whose parameters are unknown. The method is based on a difference operation realized in the wavelet domain from the variances of the wavelet coefficients for each scale which eliminates the unknown noise parameters yielding a direct $\hat{\gamma}$ estimation as the slope of a linear function. The method gives reliable results on synthetic data even for relatively low SNR. Note that for higher SNR, the method becomes similar to the wavelet-based method in the noise-free case. For the data with relatively high spectral exponent ($\gamma \geq 1.50$), the domination of white noise on $1/f^{\gamma}$ process is effective. For this case, estimation of γ is difficult.



(a)



(b)



(c)

FIGURE 4: (a) GPS noise obtained from TUBI GPS station. (b) The logarithm of the variances of the wavelet coefficients of the data in (a). Initially two different slopes of 1.3930 and 0.5303 are observed due to the white noise corruption. (c) The logarithmic difference sequence obtained from the values in (b). The spectral exponent is estimated as $\hat{\gamma} = 1.0194$.

Analysis of real GPS noise shows that such data can be modeled as the superposition of flicker noise ($\gamma = 1$) and white noise ($\gamma = 0$), as suggested by some GPS experts.

ACKNOWLEDGMENTS

This work is supported by TÜBİTAK MRC TURDEP Project no: 5057001, EU 6. Frame Foresight Project (Contract no: 511139), and in part by TÜBİTAK CAYDAG Project no: 106Y090.

REFERENCES

[1] G. W. Wornell, "Wavelet-based representations for the $1/f$ family of fractal processes," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1428–1450, 1993.

[2] D. C. Agnew, "The time-domain behavior of power-law noises," *Geophysical Research Letters*, vol. 19, no. 4, pp. 333–336, 1992.

[3] J. Langbein and H. Johnson, "Correlated errors in geodetic time series: implications for time-dependent deformation," *Journal of Geophysical Research*, vol. 102, no. B1, pp. 591–604, 1997.

[4] A. Mao, C. G. A. Harrison, and T. H. Dixon, "Noise in GPS coordinate time series," *Journal of Geophysical Research*, vol. 104, no. B2, pp. 2797–2816, 1999.

[5] S. D. P. Williams, Y. Bock, P. Fang, et al., "Error analysis of continuous GPS position time series," *Journal of Geophysical Research*, vol. 109, no. B3, pp. 1–19, 2004.

[6] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, 1994.

[7] B. B. Mandelbrot and J. W. van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, 1968.

[8] G. W. Wornell and A. V. Oppenheim, "Estimation of fractal signals from noisy measurements using wavelets," *IEEE Transactions on Signal Processing*, vol. 4, no. 3, pp. 611–623, 1992.

[9] B. Ninness, "Estimation of $1/f$ noise," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 32–46, 1998.

[10] L. Du, Y. Zhuang, and Y. Wu, "$1/f^\gamma$ noise separated from white noise with wavelet denoising," *Microelectronics Reliability*, vol. 42, no. 2, pp. 183–188, 2002.

[11] L. M. Kaplan and C.-C. J. Kuo, "Fractal estimation from noisy data via discrete fractional Gaussian noise (DFGN) and the Haar basis," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3554–3562, 1993.

Süleyman Baykut is currently a Ph.D. candidate and Research Assistant in the Department of Electronics and Communications Engineering at Istanbul Technical University, Istanbul, Turkey. He received his B.S. degree (2002) in electrical-electronics engineering from Istanbul University, Istanbul, Turkey, and his M.S. degree (2004) in telecommunication engineering from Istanbul Technical University, Istanbul, Turkey. His research interests include fractal signal processing, $1/f$ (power-law) processes, noise analysis, and underwater acoustics. Part of his Ph.D. study is supported by The Scientific and Technical Research Council of Turkey (TÜBİTAK-BIDEB).

Tayfun Akgül has been a Professor at the Department of Electronics and Communications Engineering in Istanbul Technical University (ITU), Istanbul, Turkey, since 2002, and Chief Senior Researcher (part-time) in the Earth and Marine Sciences Institute at TÜBİTAK Marmara Research Center, Kocaeli, Turkey. His ongoing research is in the area of signal/image processing, array processing, acoustics, speech, and geophysical signal processing. Between 1999 and 2002, he was the Chief Senior Researcher in the Information Technologies Research Institute at TÜBİTAK Marmara Research Center. He was an Assistant Professor and later an Associate Professor in the Department of Electrical Engineering at Çukurova University in Turkey. Also, between April 1997 and November 1998, he was a Visiting Assistant Professor and later a Research Associate Professor in the Electrical and Computer Engineering Department at Drexel University, USA. Between 1996 and 1997, he was a NATO postdoctoral fellow at the University of Pittsburgh. He received his Ph.D. degree in electrical engineering from the University of Pittsburgh in April 1994. He is a Senior Member of the IEEE, and currently serving as a Member-at-Large in the IEEE Publication Services and Products Board.

Semih Ergintav is a Chief Senior Research Geophysicist at TÜBİTAK Marmara Research Center (MRC), Turkey and a Deputy Director of the Earth and Marine Sciences Research Institute, MRC, TÜBİTAK. He holds the Ph.D. degree in seismic data processing and modeling. His ongoing research is in analyzing and modeling of the GPS time series, conventional and unconventional signal processing of geophysical data and earthquake process.