"Athens Conference on Applied Probability and Time Series Analysis. Volume II: Time Series Analysis in Memory of E. J. Hannan".
Appeared in "Lecture Notes in Statistics", Vol. 115 (1996) 420-432, Copyright Springer Verlag.

Semi-parametric graphical estimation techniques for long-memory data.

Murad S. Taqqu
Vadim Teverovsky [1] [2] [3]

ABSTRACT This paper reviews several periodogram-based methods for estimating the long-memory parameter H in time series and suggests a way to robustify them. The high frequencies tend to bias the estimates. Using only low frequencies eliminates the bias but increases the variance. We hence suggest plotting the estimates of H as a function of a parameter which balances bias versus variance and, if the plot flattens in a central region, to use the flat part for estimating H. We apply this technique to the periodogram regression method, the Whittle approximation to maximum likelihood and to the local Whittle method. We investigate its effectiveness on several simulated fractional ARIMA series and also apply it to estimate the long-memory parameter H in computer network traffic.

1 Introduction

Time series with long memory have been considered in many fields including hydrology, biology and computer networks. Unfortunately, estimating the long memory (long-range dependence) parameter H in a given data series is a difficult task. Some of the more accurate estimation methods, e.g. the Whittle estimator, are parametric in that the exact form of the spectral density must be known *a priori*. If the assumed spectral density function is not the correct one, then the estimated H may be biased. In this paper we review several periodogram-based methods for estimating H and propose a graphical extension in order to reduce the bias.

The methods we will try to modify are the periodogram regression method originally introduced by Geweke and Porter-Hudak [4], the Whittle approximate likelihood method (see Fox and Taqqu [3], Beran [1]) and the "local Whittle" method developed by Robinson [18]. An empirical study of

---

[1] Research partially supported by NSF grants NCR-9404931 and DMS-9404093 at Boston University.
[2] *AMS Subject classification:* 60G18, 62-09.
[3] *Keywords:* Long-range dependence, long memory, periodogram, Whittle.

these methods can be found in Taqqu, Teverovsky and Willinger [23] and in Taqqu and Teverovsky [22].

We will plot the estimates of H as a function of a parameter which balances bias versus variance. For the periodogram regression and for the local Whittle methods we vary the fraction of frequencies which is used in the estimation of H. For the Whittle method we aggregate the given data series and then apply the method assuming a Fractional Gaussian Noise (FGN) model. The goal is to achieve a "flattening" of the plot of the estimates of H for a certain range of values of the varying parameter. We will then use the value of $\widehat{H}$ given by the flat part of the plot as an estimate of the long-memory parameter.

We test our methods on simulated Gaussian FARIMA$(1, d, 1)$ series and also apply them to two very long data series of Ethernet traffic, which are believed to exhibit long memory, i.e. have $H \in (1/2, 1)$. A FARIMA$(1, d, 1)$ series ($H = d + 1/2$) is defined by:

$$(1 - \phi B)(1 - B)^d X_i = (1 - \theta B) Z_i, \tag{1.1}$$

where Z_i are a Gaussian white noise sequence, ϕ, θ and d are the parameters of the process and B is the backward shift operator, $BX_i = X_{i-1}$. For more information about FGN and FARIMA see Samorodnitsky and Taqqu [20] and Beran [1].

We consider the regression on the periodogram method in Section 2, the "aggregated Whittle" in Section 3 and the "local Whittle" in Section 4. Our conclusions will be presented in Section 5.

2 Regression on the Periodogram

This method is based on the periodogram of the time series, defined by

$$I(\nu) = \frac{1}{2\pi N} \Big| \sum_{j=1}^{N} X_j e^{ij\nu} \Big|^2, \tag{1.2}$$

where ν is the frequency, N is the number of terms in the series, and X_j are the data. A series with long-range dependence has a spectral density proportional to $|\nu|^{1-2H}$ close to the origin. Since $I(\nu)$ is an estimator of the spectral density, a regression of the logarithm of the periodogram vs. the logarithm of the frequency ν should give a coefficient of $1 - 2H$.

This method was first developed by Geweke and Porter-Hudak [4] in a slightly different version which is referred to as the GPH estimator. Their estimator regresses on $\log |2 \sin(\nu/2)|$ instead of $\log |\nu|$. Clearly, at very small frequencies the two are equivalent. A FARIMA$(0, d, 0)$ model has the spectral density $|2 \sin(\nu/2)|^{1-2H}$, whereas a Fractional Gaussian Noise (FGN) model has a spectral density closer, although not equal, to $|\nu|^{1-2H}$.

The function $C|\nu|^{1-2H}$ lies approximately half-way between these two theoretical spectral densities. For plots see [22].

Whatever the independent variable, not all of the periodogram frequencies are used in the regression, since the proportionality to $|\nu|^{1-2H}$ or $|2\sin(\nu/2)|^{1-2H}$ holds only for ν close to the origin. Finding the right cut-off is very important. Geweke and Porter-Hudak [4] suggested using a cut-off of $M = \sqrt{N}$ where N is the series length, and M is the number of frequencies used for the regression.

In our work we regress the logarithm of the periodogram on $\log|\nu|$, as a compromise between the FARIMA$(0,d,0)$ and FGN models (see Fig. 2 [22]). In [23] we used this method on simulated series of both FGN and FARIMA$(0,d,0)$, focusing on Fourier frequencies $\nu_j = 2\pi j/N$, $j = 1,\ldots,M$, with $N = 10,000$ and $M \simeq 500$. For those types of series the method works quite well. On the other hand, when series such as FARIMA$(1,d,1)$ are used, the periodogram can be biased because of the short-range dependence structure.

There have been many attempts to justify the periodogram regression method. Yajima [26] showed that the discrete Fourier transforms of long memory series become independently normal as the sample size N approaches infinity, but his work applies only to a fixed finite number of frequencies that does not change with the sample size. Künsch [13] showed that the normalized periodogram is asymptotically i.i.d. χ^2 if there are a fixed finite number of varying frequencies that tend to zero with N but such that the smallest frequency multiplied by $N^{1/2}$ approaches infinity. He also proposed a variation on the GPH estimator by ignoring some portion L of the smallest frequencies. The variation was motivated by a desire to distinguish between monotonic trends and long-range dependence. Robinson [19] has established the consistency and asymptotic normality of a similar modification of the GPH estimator. He uses $\log|\nu|$ as the regressor, and does not fix the number of frequencies used in the regression. Robinson shows that, if $M = o(N^{4/5})$ and

$$\frac{L}{M} + \frac{M^{1/2}\log M}{L} + \frac{(\log N)^2}{M} \to 0,$$

then the periodogram for frequencies between $2\pi L/N$ and $2\pi M/N$ is asymptotically i.i.d. χ^2, and the H estimated from the regression is consistent and asymptotically normal. Robinson has also suggested using an estimator based on the averaged (or integrated) periodogram [17], and derived criteria for choosing M [16], but they depend, unfortunately, on H. The estimator, moreover, is biased downward, especially for H close to 1.

Hurvich and Beltrao [9] have shown that the periodogram is, in fact, asymptotically neither i.i.d. nor χ^2 distributed at the very low frequencies of order $1/N$. This does not contradict the results of Künsch and Robinson, since they were dealing with frequencies that tend to zero much more slowly. It implies a bias for the GPH estimator which is noticeable at the very low frequencies only. In the same paper, Hurvich and Beltrao propose tapering the periodogram as a way to reduce the bias. They provide in [10] some criteria for choosing M and L, based on frequency domain cross-validation and a minimization procedure over M and L. Simulation results are presented to show that their method is slightly better than the GPH estimator, and better, as well, than the variation proposed by Künsch and Robinson. Beran, in his recent book [1], presents several of the above results.

In a recent paper, Hurvich, Deo and Brodsky [11] have established some asymptotic properties of the original (GPH) periodogram estimator. They show that if the spectral density is $f(\nu) = |2\sin(\nu/2)|^{1-2H} f^*(\nu)$, then the bias behaves asymptotically like

$$-\frac{2\pi^2}{9}\frac{f^{*''}(0)}{f^*(0)}\frac{M^2}{N^2}.$$

Here f^* is an even, positive, continuous function on $[-\pi, \pi]$, with certain other regularity properties that ensure that it posseses only short-range dependence.

Hurvich, Deo and Brodsky also show that the variance of the GPH estimator is equal to $\pi^2/(24M)$ and thus obtain a theoretical expression for the mean squared error (MSE) of the estimator. They obtain a formula for the M which minimizes the MSE of the GPH estimator. This M^{OPT} is proportional to $N^{4/5}$, but the proportionality constant again depends on f^*. Even if the asymptotic relations were to hold exactly, M can not be chosen *a priori*.

Instead of picking *a priori* a value of M (whether blindly or using the M^{OPT} above) and then finding an estimate of H, we suggest letting M vary through a fairly wide range. The various estimates of H are then plotted as a time series. Starting at large values of M (small values of N/M), we would expect to find a range where the estimates of H are incorrect because of the short-range effects. Then, as M decreases (N/M increases), the short-range effects should disappear and the value of H obtained should represent the true long-memory dependence. Thus there should be a period of relative stability, where the estimates of H are approximately constant. Then, if we move to smaller M's, we will get into a region where the estimates of H are very scattered and unreliable because there are not enough frequencies left to have an accurate regression. Thus, we should expect to see a flat region somewhere in the middle of the plot of the estimates of H and we can estimate an overall H from that region. In practice, of course, the region will not be completely flat.

We have applied this method to 10 simulated series of FARIMA$(1, d, 1)$ of length $N = 10,000$, with parameters $\phi = 0.3, \theta = 0.7, d = 0.3$ ($H = 0.8$). We used the above method with $N/M = 4(4)200$. For each series, a line was drawn connecting the estimates at the various values of M. The plot is given

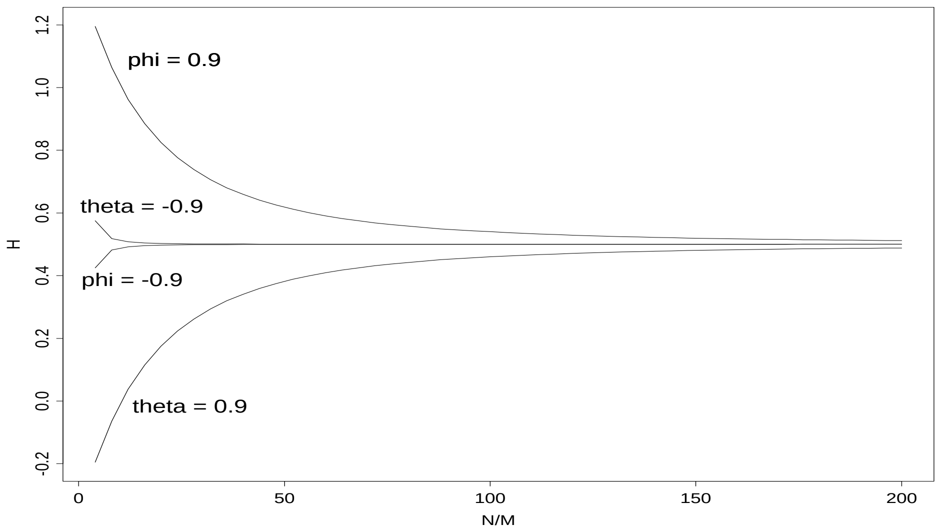

FIGURE 1. Periodogram method applied to the theoretical spectral density of AR(1) processes with $\phi = \pm 0.9$ and MA(1) processes with $\theta = \pm - 0.9$.

in Figure 2. In addition the plot contains some "theoretical" results related to this method: we used the theoretical spectral density of the FARIMA process in place of the periodogram in the regression, and generated the "estimates" of H at the same values of M as above. This is illustrated by the thickest curve in the plot, which starts at approximately 0.5 and levels off at 0.8, which is the nominal value of H. The other two curves in the plot denote the approximate 95% confidence intervals ($\pm 1.96\sigma$) around the "theoretical" result, where $\sigma = \pi/\sqrt{24M}$ is the asymptotic standard deviation (see [19], [11]). It can be seen that the simulated results stay within these bounds almost all of the time, and that most of the lines flatten out appreciably at higher N/M. On the other hand, the bounds are large enough that the estimates of H thus obtained can be fairly scattered.

We also include the "theoretical" curves for several other series, namely MA(1) with $\theta = \pm 0.9$ and AR(1) with $\phi = \pm 0.9$ (see Figure 1). These are extreme choices, that include both "good" and "bad" case scenarios. In accordance with the results of Hurvich, Deo and Brodsky the magnitude of the bias is the same in the cases $\theta = 0.9$ and for $\phi = 0.9$. The magnitude of the bias is also the same, but much smaller, in the cases $\theta = -0.9$ and $\phi = -0.9$. All the "theoretical" curves, however, are different, illustrating the fact that care has to be exercized when using a periodogram estimator.

In real life one does not have the luxury of either examining ten replicates of a time series or using their theoretical spectral density. One has only one actual time series to analyze. We applied the graphical extension to the periodogram regression method to two very long series of Ethernet traffic, considered by Leland, Taqqu, Willinger and Wilson [14], [25]. The first Ethernet data set gives the number of bytes per 10 milliseconds passing through a monitoring system during a "normal traffic hour" in August 1989. The second gives the number of packets sent per 10 milliseconds during a "normal traffic hour" in August 1989. Thus, each of the series is $360,000$ long. The authors obtain estimates of approximately $H = 0.8$ for the *byte* data, and $H = 0.9$ for the *packet* data, by using several different methods (see also Teverovsky and Taqqu [24]). The results of using the graphical periodogram regression method on these series are shown in Figure 5 ($N/M = 10(10)1000$). The curves do flatten out eventually, and H can be estimated to be approximately 0.70 to 0.75 for the *byte* data and approximately 0.90 to 0.95 for the *packet* data.

3 Aggregated Whittle Estimator

The Whittle estimator involves the function

$$Q(\eta) := \int_{-\pi}^{\pi} \frac{I(\nu)}{f(\nu;\eta)} d\nu, \tag{1.3}$$

where $I(\nu)$ is the periodogram (see (1.2)) and $f(\nu;\eta)$ is the spectral density at frequency ν, where η denotes the vector of unknown parameters. The Whittle estimator is the value of η which minimizes the function Q. When dealing with fractional Gaussian noise or fractional ARIMA$(0, d, 0)$, η is simply the parameter H or d. If the series is assumed to be FARIMA(p, d, q) then η also includes the unknown coefficients in the autoregressive and moving average parts.

Ted Hannan contributed in an important way to the development of the (classical) Whittle's method. For a stationary Gaussian time series with spectral density $f(\nu, \beta)$, $-\pi < \lambda < \pi$, Whittle's method, which provides an estimate of β, requires replacing the inverse covariance matrix that appears in the Gaussian likelihood by a Toeplitz (covariance) matrix with spectral density $1/f$ and then maximizing the quadratic form. Hannan, in the seminal paper [6], applied Whittle's method to finite variance ARMA time series. He proved that the estimator is consistent and asymptotically normal. An ARMA time series, however, has short-range dependence because the correlations decrease exponentially fast. Fox and Taqqu [3] extended this result to Gaussian time series with long-range dependence such as fractional Gaussian noise or fractional ARIMA by appealing to a central limit theorem for weighted quadratic forms whose weights are chosen in such a way as to compensate for the long-range dependence. Fox and Taqqu's result, which was later generalized to the full maximum likelihood by Dahlhaus [2], is the basis of one of the most commonly used techniques for estimating the intensity of long-range dependence in Gaussian time series (see Beran [1]). Giraitis and Surgailis [5] extended Fox and Taqqu's result to finite variance innovations without Gaussian assumptions and Heyde and Gay [8] to random fields. Mikosch, Gadrich, Klüppelberg and Adler [15] went back to the original Hannan's ARMA result and investigated what happens if the Whittle estimator is used on ARMA time series with infinite variance; Kokoszka and Taqqu [12] then provided the extension to FARIMA sequences with long-range dependence and infinite variance.

The Whittle method is parametric. If the assumed spectral density function is not the correct one, then the estimated H obtained by the Whittle method may be incorrect. To alleviate this problem, one may try to aggregate the series, namely consider

$$X_k^{(m)} := \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i$$

where X_i is the original time series and m is the aggregation level. As the aggregation level becomes large, $X^{(m)}$ approaches FGN if X_i is Gaussian with long-range dependence (see Taqqu [21]). Thus if we aggregate the series at a large enough level, we can use the Whittle estimator with the FGN spectral density model to estimate H. Since the appropriate m is not known *a priori*, we again estimate H at different levels of aggregation, plot the results, and try to find a region where the graph is approximately flat. This idea was first used in Leland, Taqqu, Willinger and Wilson [14], in the context of Ethernet series.

We applied this method to the same simulated series of FARIMA$(1, d, 1)$ considered above with $d = 0.3$ ($H = 0.8$). We chose $N/m = 4(4)100$. For each series, a line was drawn connecting the estimates at the various values of m. The plot is given in Figure 3. As m grows the graph flattens to $\widehat{H} = 0.8$. Although individual estimates might range from 0.7 to 0.9, most of the estimates seem fairly close to 0.8.

We have also produced the aggregated Whittle plots for the two Ethernet series introduced in Section 2. We used $N/m = 10(10)1000$ (Figure 6). After initial periods of rise or fall, the plots for both of the series flatten out. The estimate of H seems to be about 0.9 for the packet series and about 0.75 to 0.8 for the byte series. (Refer also the Ethernet papers [14], [25]).

4 Local Whittle Estimator

This method shares features with both of the previous methods. Like the regression on the periodogram, local Whittle only assumes the behavior of the spectral density close to the origin and just like the Whittle method it involves minimizing a modification of the function Q in (1.3). It was recently introduced by Robinson [18] who derived its asymptotic properties. Taqqu and Teverovsky [22] studied its robustness. Assume

$$f(\nu) \sim G(H)|\nu|^{1-2H} \quad \text{as} \quad \nu \to 0.$$

For such a spectral density, the discrete analogue of the function Q in (1.3) is

$$Q(G,H) := \frac{1}{M}\sum_{j=1}^{M}\left(\frac{I(\nu_j)}{G\nu_j^{1-2H}} + \log G\nu_j^{1-2H}\right) \tag{1.4}$$

where $\nu_j = 2\pi j/N$. Note that: (a) the frequencies are integrated (summed) only up to $2\pi M/N$; (b) the spectral density f is replaced by its asymptotic form. M is assumed to satisfy:

$$\frac{1}{M} + \frac{M}{N} \to 0 \quad \text{as} \quad N \to \infty.$$

Replacing the constant G by its estimate $\widehat{G} = M^{-1}\sum_{j=1}^{M} I(\nu_j)/\nu_j^{1-2H}$, one minimizes $R(H) := Q(\widehat{G}, H) - 1$, that is

$$R(H) = \log\left(\frac{1}{M}\sum_{j=1}^{M}\frac{I(\nu_j)}{\nu_j^{1-2H}}\right) - (2H-1)\frac{1}{M}\sum_{j=1}^{M}\log\nu_j.$$

Robinson [18] shows that under some additional technical conditions, $\widehat{H} = \arg\min R(H)$ converges in probability to the actual value H_0, as $N \to \infty$. Moreover, under slightly stronger assumptions on f and M,

$$M^{1/2}(\widehat{H} - H_0) \to_d \text{Normal } (0, 1/4), \quad \text{as} \quad N \to \infty. \tag{1.5}$$

This last relation yields asymptotic values for the standard deviations, confidence intervals, etc. The asymptotic variance is smaller than for the regression on the periodogram method.

Note that the larger the value of M, the smaller the variance of $\widehat{H}$. On the other hand, if the series is not "ideal", e.g. if it is FARIMA$(1, d, 1)$, then M has to be small in order to avoid bias. In such a case we only want to use the frequencies which are close to zero, since at larger frequencies the short-range behavior of the series will affect the form of the spectral density. In a previous paper [22] we have examined several FARIMA series at several discrete values of M. In particular, we found that M had to be no more than approximately $N/32$ when $N = 10,000$ in order to achieve reasonable accuracy for the series we used.

Here, we introduce the continuous version of the above method, where we will calculate H for many values of M and plot the resulting estimates. If the range of values of M is chosen correctly then we should again see a fairly constant region when the estimates of H converge to the correct long-memory parameter.

We apply this method, with $N/M = 4(4)200$, to the same simulated FARIMA series used for the two previous estimators. As in the regression on the periodogram method, we plot the "theoretical" values of the estimates and the approximate confidence intervals in Figure 4. The "theoretical" values are obtained by using the spectral density of the known process in place of the periodogram in the estimator. The confidence intervals correspond to the standard deviations (1.5) around these "theoretical" values. The simulated results stay almost always within these bounds. Henry and Robinson [7] have recently derived results for the bias of this estimator. These are very similar to the results obtained for the regression on the periodogram.

We have used this estimator to estimate H for the two Ethernet series discussed above using $N/M = 10(10)1000$. The results are shown in Figure 7 and are very similar to the results for the regression on the periodogram. Based on the local Whittle method, $\widehat{H}$ is 0.75 for the *byte* series and 0.9 for the *packet* series. The estimates for all three methods are hence fairly consistent.

We should also mention that there is a very large downward spike for the packet data on the left edge of the plot for all three methods. This is due to a large spike in the periodogram of the data series occuring at a frequency of approximately $2\pi/64$. Thus, when this frequency is included in the periodogram regression or the function Q for the Whittle estimators it causes a spike in the estimates of H. Once M is small enough for this frequency not to be included, the estimates of H tend to be more consistent.

5 Conclusion

The original Whittle method for estimating the long memory parameter H is parametric: it is highly accurate if the underlying model is correct. Other methods, such as the regression on the periodogram and the local Whittle use a cut-off parameter to ensure robustness. It is difficult, unfortunately, to decide *a priori* on the best choice of values for these parameters.

We have introduced several modifications to these methods, in order to increase robustness with respect to short memory effects and applied them to two real-life series of Ethernet traffic of length $N = 360,000$. The graphical technique used for the aggregated Whittle estimator in [14], [25], was extended to the regression on the periodogram and the local Whittle methods. In the aggregated Whittle method, the data is aggregated at various levels and the estimates of H are computed for the aggregated series. In the other two methods, the values of H are computed for various cut-off values. The results in all cases are displayed in a plot which should first display variability and then become approximately flat. Since at that point, the short-range dependencies are no longer affecting the estimates to

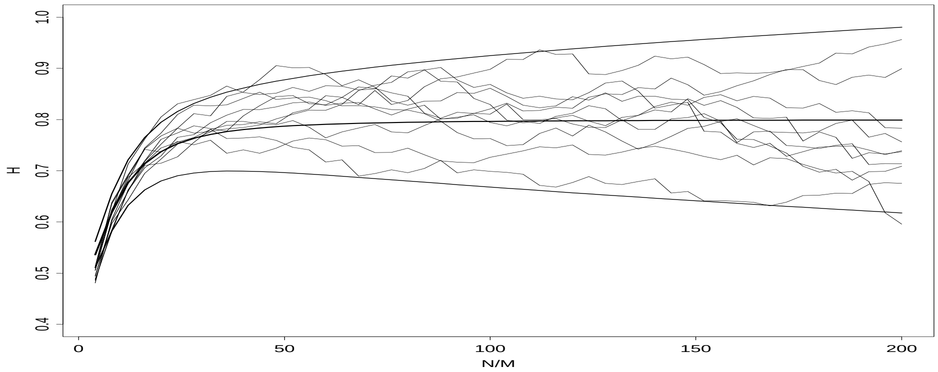

FIGURE 2. Periodogram regression estimates for 10 series of FARIMA(.3, .3, .7).

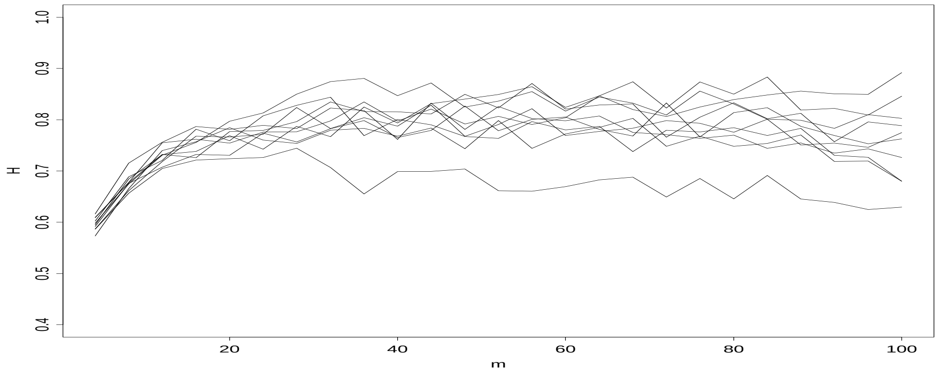

FIGURE 3. Aggregated Whittle estimates for 10 series of FARIMA(.3, .3, .7).

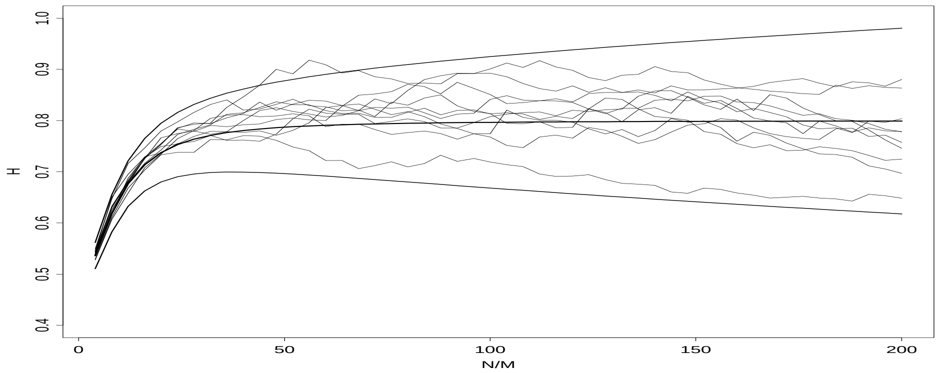

FIGURE 4. Local Whittle estimates for 10 series of FARIMA(.3, .3, .7).

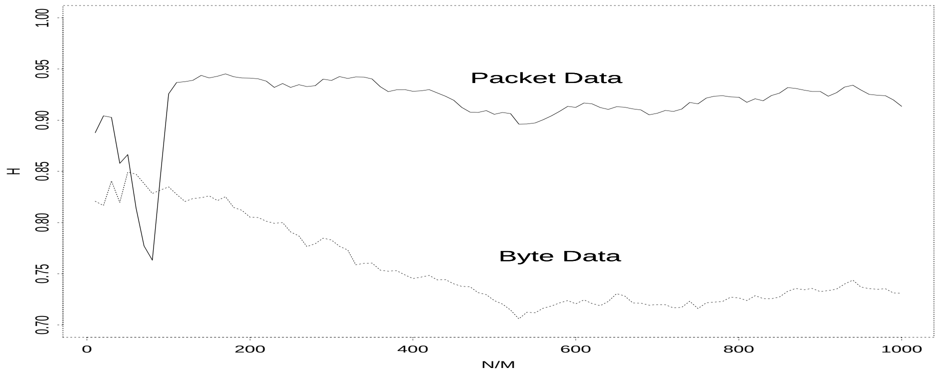

FIGURE 5. Periodogram regression estimates for Ethernet series.

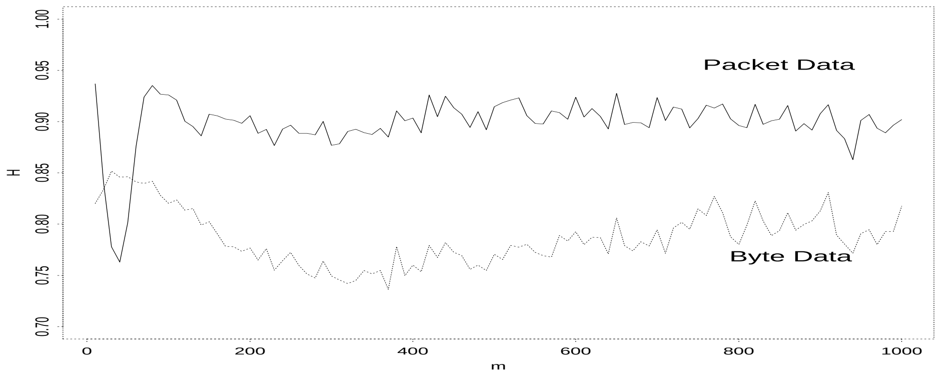

FIGURE 6. Aggregated Whittle estimates for Ethernet series.

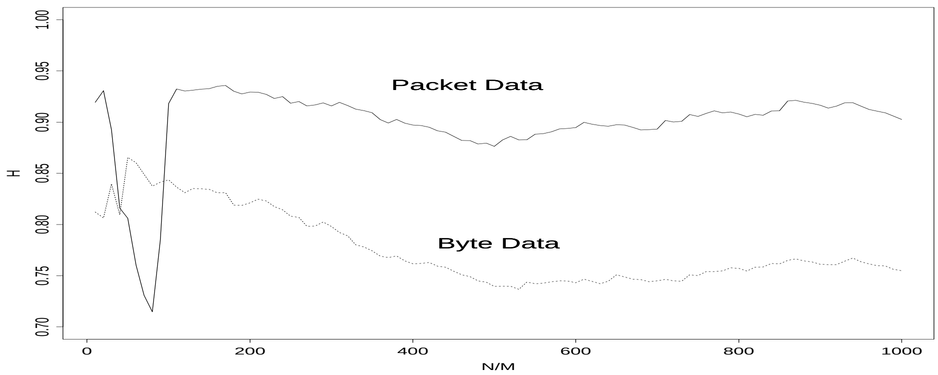

FIGURE 7. Local Whittle estimates for Ethernet series.

any noticeable degree, the resulting value of H should be fairly unbiased.

It is difficult to evaluate these methods in a quantifiable way since the estimates of H are, of necessity, somewhat subjective. This is why we presented several plots of the results using simulated series of length $N = 10,000$. We also presented corresponding "theoretical" graphs for several other types of series. The results seem to confirm that the estimators do behave as expected and that it is preferable to use these plots of $\widehat{H}$ rather than the $\widehat{H}$ corresponding to a fixed *a priori* value of N/M.

6 References

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

[2] R. Dahlhaus. Efficient parameter estimation for self similar processes. *The Annals of Statistics*, 17(4):1749–1766, 1989.

[3] R. Fox and M. S. Taqqu. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, 14:517–532, 1986.

[4] J. Geweke and S. Porter-Hudak. The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4:221–238, 1983.

[5] L. Giraitis and D. Surgailis. A central limit theorem for quadratic forms in strongly dependent linear variables and application to asymptotical normality of Whittle's estimate. *Probability Theory and Related Fields*, 86:87–104, 1990.

[6] E. J. Hannan. The asymptotic theory of linear time series models. *Journal of Applied Probability*, 10:130–145, 1973.

[7] M. Henry and P. M. Robinson. Bandwidth choice in Gaussian semiparametric estimation of long range dependence. In *Proceedings of the Athens Conference on Applied Probability and Time Series Analysis*, Springer-Verlag, New York, 1996. Time series volume in honour of E. J. Hannan. Appears in this volume.

[8] C. C. Heyde and R. Gay. Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. *Stochastic Processes and their Applications*, 45:169–182, 1993.

[9] C. M. Hurvich and K. I. Beltrao. Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. *Journal of Time Series Analysis*, 14:455–472, 1993.

[10] C. M. Hurvich and K. I. Beltrao. Automatic semiparametric estimation of the memory parameter of a long memory time series. *Journal of Time Series Analysis*, 15:285–302, 1994.

[11] C.M. Hurvich, R. Deo, and J. Brodsky. The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long memory time series. Preprint, 1995.

[12] P. S. Kokoszka and M. S. Taqqu. Parameter estimation for infinite variance fractional ARIMA. To appear in *The Annals of Statistics*, 1995.

[13] H. Künsch. Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability*, 23:1025–1030, 1986.

[14] 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*, 2:1–15, 1994.

[15] T. Mikosch, T. Gadrich, C. Klüppelberg, and R. J. Adler. Parameter estimation for ARMA models with infinite variance innovations. *The Annals of Statistics*, 23:305–326, 1995.

[16] P. M. Robinson. Rates of convergence and optimal bandwidth in spectral analysis of processes with long range dependence. *Probability Theory and Related Fields*, 99:443–473, 1994.

[17] P. M. Robinson. Semiparametric analysis of long-memory time series. *The Annals of Statistics*, 22:515–539, 1994.

[18] P. M. Robinson. Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, 1995. To appear.

[19] P. M. Robinson. Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, 23:1048–1072, 1995.

[20] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance.* Chapman and Hall, New York, London, 1994.

[21] M. S. Taqqu. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31:287–302, 1975.

[22] M. S. Taqqu and V. Teverovsky. Robustness of Whittle-type estimates for time series with long-range dependence. Preprint, 1995.

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

[24] V. Teverovsky and M. S. Taqqu. Testing for long-range dependence in the presence of shifting means or a slowly declining trend using a variance-type estimator. Preprint, 1995.

[25] W. Willinger, M. S. Taqqu, W. E. Leland, and V. Wilson. Self-similarity in high-speed packet traffic: analysis and modeling of Ethernet traffic measurements. *Statistical Science*, 10:67–85, 1995.

[26] Y. Yajima. A central limit theorem of Fourier transforms of strongly dependent stationary processes. *Journal of Time Series Analysis*, 10:375–383, 1989.

Murad S. Taqqu and Vadim Teverovsky
Boston University
Department of Mathematics
111 Cummington Street
Boston, MA 02215-2411, USA
Email: murad@math.bu.edu, vt@math.bu.edu