

Spectral Bandwidth Selection for Long Memory

Grace Lee Ching Yap¹ & Wen Cheong Chin²

¹ School of Applied Mathematics, Faculty of Engineering, (Malaysia Campus), 43500 Semenyih, Selangor, Malaysia

² Faculty of Management, SIG Quantitative Economics and Finance, Multimedia University, 63100 Cyberjaya, Selangor, Malaysia

Correspondence: Grace Lee Ching Yap, School of Applied Mathematics, Faculty of Engineering, (Malaysia Campus), 43500 Semenyih, Selangor, Malaysia. E-mail: grace.yap@nottingham.edu.my

Received: March 14, 2016

Accepted: April 20, 2016

Online Published: June 7, 2016

doi:10.5539/mas.v10n8p63

URL: <http://dx.doi.org/10.5539/mas.v10n8p63>

The research is financed by Fundamental Research Grant Scheme (FRGS) FRGS/1/2014/SG04/UNIM/03/1.

Abstract

Long-memory parameter estimation using log-periodogram regression relies largely on the frequency bandwidth and the order of estimation. Literature shows that a data-dependent plug-in method for the bandwidth significantly increases the MSE's. In a long memory time series with mild short range effect, a simple approach to determine the bandwidth size is suggested based on the spectral analysis. Monte Carlo simulation results and empirical applications show that the proposed bandwidth selection performs satisfactorily.

Keywords: long memory, log-periodogram, spectral density, optimal bandwidth

1. Introduction

There has been research on estimation of long memory parameter d in time series using periodogram-based semi-parametric estimate, namely the local Whittle (Kunsch, 1987; Robinson, 1995a), average periodogram (Robinson, 1994) and log-periodogram (Geweke & Porter-Hudak, 1983; Robinson, 1995b). Semi-parametric estimation procedures are desired in the time series analysis of financial measurements sampled at high frequencies (Barros, Gil-Alana & Payne 2014; Bollerslev et al., 2013; Garvey & Gallagher, 2013) as they allow the estimation of the long-run characteristics (low frequency behaviour) of the time series without the knowledge of the short-run (high frequency) structure. Amongst these methods, log-periodogram (LP) regression proposed by Geweke and Porter-Hudak (1983) (GPH) has become a popular tool for statistical inference in empirical research due to its simple implementation, pivotal asymptotic normality and robustness as a result of the local condition (Arteche & Orbe, 2009). Nonetheless, it has been criticized due to its finite-sample bias (Agiakloglou, Newbold & Wohar 1993). To reduce the bias, Andrews and Guggenberger (2003) (AG) proposed a bias-reduced log-periodogram estimator (BRLP) \hat{d}_r , $r \in \mathbb{Z}^+$. Literature shows that the rate of convergence to zero of the mean-squared error (MSE) of \hat{d}_{GPH} is of order $n^{-\frac{4}{5}}$ and it is of order $n^{-\frac{4+4r}{5+4r}}$ with \hat{d}_r . Also, the rate of convergence of the latter exceeds that of the local Whittle estimator (Robinson, 1995a) and the average periodogram (Robinson, 1994), provided the spectral density is sufficiently smooth. Nonetheless, as a semi-parametric parameter estimate, the bandwidth m plays an important role on the performance of \hat{d}_r . A large bandwidth reduces the variance at the cost of increased bias, and the estimates of the memory parameter vary significantly with the choice of m . To balance the squared bias and variance, optimal bandwidth selection is usually proposed to minimize the approximation of the MSE or the root mean-squared error (RMSE).

There are basically three formal procedures to optimal bandwidth. Hurvich and Deo (1999) proposed a plug-in method that minimizes an asymptotic approximation of the MSE, Giraitis, Robinson and Samarov (2000) introduced an adaptive LP that chooses m to adapt to an unknown local to zero spectral smoothness, and Arteche and Orbe (2009) suggested a bootstrap-based bandwidth choice that minimizes a local bootstrap MSE. The adaptive LP does not give unique choices of the bandwidth but only the bandwidths with an optimal growth rate. The bootstrap-based bandwidth is claimed to be robust, if the signals from the data are not misinterpreted. Nonetheless, it is noted that this technique depends largely on the choice of resampling width and the range of bandwidths considered for optimization. On the other hand, the plug-in method for selecting the number of

frequencies to minimize asymptotic MSE of \hat{d} or \hat{d}_r is easy to implement but it is usually not adequate as it depends on the unknowns to be estimated (Andrews & Guggenberger, 2003; Arteche, 2004; Delgado & Robinson, 1996; Henry & Robinson, 1996; Henry, 2001). Besides, as the convergence rate is measuring the asymptotic properties that will not be reflected in finite samples, an improvement in the search for an optimum bandwidth is deemed crucial. It is believed that such effort is encouraging as much of the research work rely on the long-memory parameter estimation that is done by arbitrary choice of bandwidth (Charfeddine, 2014; Choi, Yu, & Zivot, 2010; Garvey & Gallagher, 2013). In a long memory time series with mild short range effect, the search for the optimum bandwidth can be done easily with spectral analysis. The following section reviews the log-periodogram regression estimator. Section 3 considers an alternative approach for the bandwidth selection using the spectral analysis. Section 4 provides a simulation study in finite samples, section 5 illustrates the proposed method with empirical examples and section 6 offers the concluding remarks.

2. Log-periodogram Regression Estimator

The spectral density of a semi-parametric model for a stationary Gaussian long-memory time series $\{Y_t: t = 1, \dots, n\}$ is given by

$$f(\lambda) = |\lambda|^{-2d} f^*(\lambda), \text{ as } \lambda \rightarrow 0 \quad (1)$$

Where $d \in (-0.5, 0.5)$ determines the low frequencies property. Specifically when $0 < d < 0.5$, the series exhibits long memory. $f^*(\cdot)$ is an even, positive continuous function on $[-\pi, \pi]$ with $0 < f^*(0) < \infty$. It determines the high frequencies properties of the series, relating to the short-term correlation structure. A model that takes a fractional difference of order d , a p -order autoregressive (AR) and q -order moving average (MA), abbreviated as ARFIMA (p, d, q) introduced by Granger and Joyeux (1980) and Hosking (1981) is a special case of long-range process satisfying Eq. (1). Based on this, Robinson (1995b) wrote the GPH estimator \hat{d} in the form of regression model below.

$$\log I(\lambda_j) = (\log f^*(0) - C) + dX(\lambda_j) + \log \left(\frac{f^*(\lambda_j)}{f^*(0)} \right) + \varepsilon_j \text{ for } j = 1, \dots, m \quad (2)$$

Where $I(\lambda_j) = \frac{1}{2n\pi} \left| \sum_{t=1}^n Y_t \exp(i\lambda_j t) \right|^2$, $\lambda_j = \frac{2\pi j}{n}$, $X(\lambda_j) = -2\log|\lambda_j|$, $\varepsilon_j = \log \left(\frac{I(\lambda_j)}{f^*(\lambda_j)} \right) + C$,

$C = 0.577216 \dots$ is the Euler constant and m is a positive integer smaller than or equal to $\left\lfloor \frac{n}{2} \right\rfloor$, $[x]$ being the largest integer part of x .

To avoid the bias in \hat{d} (Hurvich et al., 1998) due to the error term $\{\varepsilon_j\}$ that is not i.i.d. (Hurvich & Beltrao, 1993; Kunsch, 1986; Robinson, 1995b), AG proposed a bias-reduced log-periodogram regression (BRLP) that gives \hat{d}_r , which is basically adding regressor λ_j^{2r} , $r \geq 1$ to the pseudo-regression model Eq.(2). The asymptotic bias of \hat{d}_r is of order $\left(\frac{m}{n}\right)^{2+2r}$, compared to that of the GPH estimator \hat{d}_{GPH} which has the order of $\left(\frac{m}{n}\right)^2$. As such, a larger bandwidth size is preferred for \hat{d}_r because with m being the effective sample size to estimate d , the standard deviation of the estimator declines at the approximate rate $\frac{1}{\sqrt{m}}$, which in effect, yields low RMSE.

Various effort have been made to determine the suitable bandwidth size m . For the plug-in method, Hurvich et al. (1998) suggested an optimal bandwidth choice for \hat{d}_{GPH} , of which following the lines of Delgado & Robinson (1996), it can be estimated using Taylor expansion of $f^*(\lambda_j)$ about $\lambda = 0$. As f^* is unknown, this method is not fully automatic and the estimation can be poor for some $f^*(\lambda)$ functions. For \hat{d}_r estimator, AG proposed the plug-in MSE optimal choice of bandwidth. This method has the similar problem in GPH estimator, that is, $f^*(\lambda)$ is unknown, and a non-parametric estimation is involved, which leads to an increase in the MSE of \hat{d}_r . Essentially, there is still no clue as to how large a bandwidth size should be taken. As the performance of the estimator varies according to the bandwidth size, a more efficient approach for the data-dependent choice of bandwidth is desired in finite samples.

3. An Alternative Bandwidth Selection Method for Long-memory Time Series

The spectral density is the variance per unit frequency (Chatfield, 2004), of which the auto-covariance function

$\gamma_X(k)$ can be expressed as a cosine transform of the spectral density, shown in Eq.(3).

$$\gamma_X(k) = \int_0^\pi f(\lambda) \cos(\lambda k) d\lambda \tag{3}$$

In other words, $f(\lambda)d\lambda$ is the contribution of the variance of a time series with frequencies in the range $(\lambda, \lambda + d\lambda)$. As persistence or positive autocorrelation is the characteristic of a long-memory, the variance in the time series due to the long range effect is explained at the low frequencies, and it diminishes on the high frequencies. On the other hand, a process with mild short range effect has quite a flat spectral density, particularly on the interval $\lambda \in (\frac{1}{\pi}, \pi)$ (see Figure 1).

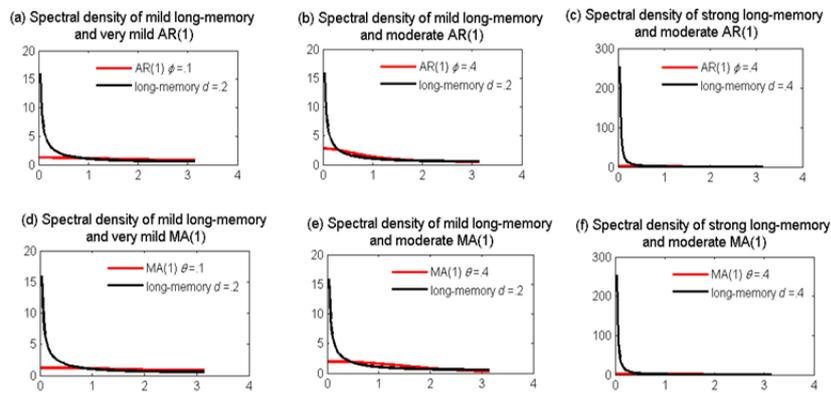


Figure 1. Spectral density of short- and long-range effects

The characteristic of a stationary long-memory process with mild short-range effect can be examined via the autocorrelation function (ACF) and its spectral density. Briefly, Figure 2 shows the characteristics of ARFIMA(0, d, 0), ARFIMA(1, d, 0) and ARFIMA(0, d, 1) processes with the plots of the average modified periodogram, ACF and partial autocorrelation function (PACF). Basically, a long-memory process with mild short-range effect has dominant long-memory characteristics, of which the autocorrelations are not too big, and the ACF decay hyperbolically with its spectral density diverges at zero. However, a significant short-range effect increases the strength of the autocorrelations at low lags, and the ACF may tend to decay more rapidly (see Figure 2(b)). For a detailed discussion of the characteristics of long-memory process, see Baillie (1996).

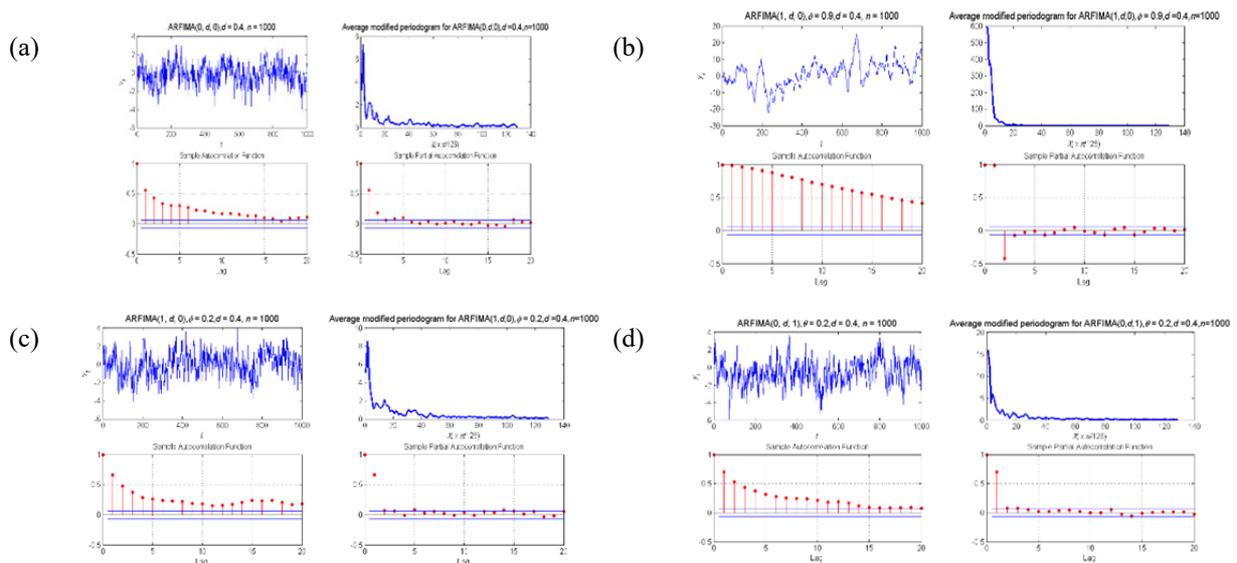


Figure 2. Characteristics of various ARFIMA processes

To obtain an accurate long-memory parameter estimation, we should include as much the information pertaining the long range behaviour as possible. GPH suggested a bandwidth size $m = \lfloor \sqrt{n} \rfloor$, but Qu (2011) showed that a larger bandwidth $m = \lfloor n^{0.7} \rfloor$ produces better results as the spurious long-memory affects the periodogram only up to $j = O(\sqrt{n})$. This is supported by the empirical analyses (Charfeddine, 2014; Choi et al., 2010; Garvey & Gallagher, 2013), of which a bandwidth size larger than $\lfloor \sqrt{n} \rfloor$ is adopted. Concentrating on the scenario of long-memory process of which the short-range effect is mild, this paper proposes to include as much the frequencies as possible up to the point of where the spectral density is low enough to be defined as noise. This approach is practicable as the spectral density f^* of a mild short-range effect is rather flat on the interval $j \in \left(\lfloor \sqrt{n} \rfloor, \lfloor \frac{n}{2} \rfloor \right)$. From Eq.(1), we can deduce that f is not affected by f^* on the interval $j \in \left(\lfloor \sqrt{n} \rfloor, \lfloor \frac{n}{2} \rfloor \right)$.

Hence, the inclusion of as many frequencies with significant power spectrum as possible in the BRLP model is warranted. The spectral densities on the high frequencies are very low, and the signals are indistinguishable from the noise. These signals are regarded as insignificant, and they are to be removed like noise elimination. The challenge to this is to define a level of spectral density that is low enough to be treated as noise. As the effect of spurious long-memory is dominant on $j < \sqrt{n}$, one may take $f(\lambda_{po}) = f\left(\frac{2\pi}{\sqrt{n}}\right)$ as the reference for the variance per unit frequency that explains the long-memory trait. The noise is then defined as those spectral density that has less than 5% of the amplitude at λ_{po} , computed as $f_c = 0.05 * [f(\lambda_{po}) - f_{\min}(\lambda)] + f_{\min}(\lambda)$, where $f_{\min}(\lambda) = \min_{\lambda \in [0, \pi]} f(\lambda)$ represents the spectral density of white noise in the time series. It can be shown that the bandwidth selected with this criterion satisfies $m = O(\sqrt{n})$, thus meeting the statistical assumptions for a regression method. Following this procedure, we avoid the bias due to insufficient data, and it is believed that more accurate long-memory parameter estimation can be attained.

Since the spectral density of a time series is unknown, an unbiased and consistent estimate of the spectral density is needed. Although periodogram is a natural estimate of spectral density, it not a consistent estimator because its variance does not decrease as $n \rightarrow \infty$ (Jenkins & Watts, 1968). To improve the spectral estimation, this paper uses the average modified periodogram $S(\lambda)$ by the Welch method (Welch, 1967) with a frequency resolution of $\frac{2\pi}{nfft}$, where $nfft = \max\{256, \text{the next power of 2 greater than the length of the segments}\}$. The spectral

density is then estimated by $S(\lambda_i), i = 0, 1, \dots, n_s$, where $n_s = \lfloor \frac{nfft}{2} \rfloor$. By overlapping the segments with one

half their fixed length (L), Welch (1967) reported that the variance of this estimator is a function of $\frac{L}{n}$. That is,

the variance of $S(\lambda)$ improves as n increases. This implies that when the sample size is small, $f(\lambda_{po})$ should be estimated from a broader neighbourhood in $S(\lambda_{po})$. To simplify the work, we propose that for $n < 1000$,

$f(\lambda_{po})$ is estimated by the largest signal in the interval $\left[\lambda_{po} - \frac{5\pi}{n_s}, \lambda_{po} + \frac{5\pi}{n_s} \right]$. As the sample size increases

($n \geq 1000$), the variance in the estimator decreases, and hence, the search interval for the estimate of $f(\lambda_{po})$

can be drawn closer to the point λ_{po} , say $\left[\lambda_{po} - \frac{3\pi}{n_s}, \lambda_{po} + \frac{3\pi}{n_s} \right]$ on the $S(\lambda)$ plot. Let's denote this estimate as

$\hat{f}(\lambda_{po})$. To include as many frequencies with significant spectrum, we suggest that moving from the last point π ,

the cutoff frequency λ_c for the bandwidth is set at the first frequency that has the amplitude equals to \hat{f}_c . This cutoff frequency suggests a preliminary search for an optimum bandwidth for the long memory parameter estimation. To refine the bandwidth, the graphical method via the plot of \hat{d}_r against m by Taqqu and

Teverovsky (1996) is referred, whereby the optimum bandwidth is searched from the flat region in the plot of \hat{d}_r against m around the Fourier frequency $\left[\frac{\lambda_c * n}{2 * n_s}\right]$ concluded from the spectral analysis. The flat region is defined as the region of frequencies of which \hat{d}_r are almost similar, say in a neighbourhood of three estimates with a standard deviation of less than 10^{-2} . The average frequency of such region gives the optimum bandwidth, and this point will determine the number of frequencies to be included in the log-periodogram regression for the long-memory parameter estimation.

The above procedure to search for the optimum bandwidth in the finite sample can be summarized in the steps below:

1. Examine the plots of modified periodogram, ACF and PACF to check if the process is with mild short range effect.
2. Compute λ_{po} and identify $\hat{f}(\lambda_{po})$ from the modified periodogram.
3. Compute the minimum amplitude $\hat{f}_c = 0.05 * [\hat{f}(\lambda_{po}) - \hat{f}_{\min}(\lambda)] + \hat{f}_{\min}(\lambda)$.
4. Identify the cutoff frequency λ_c , where $\hat{f}(\lambda_c) = \hat{f}_c$.
5. Search for the flat region in the plot of \hat{d}_r against m around the Fourier frequency $\left[\frac{\lambda_c * n}{2 * n_s}\right]$.
6. The optimum bandwidth is the average frequency of the region identified in step (5).

Whilst the above procedure suggests the number of frequencies needed in the BRLP model, the accuracy of \hat{d}_r depends on r . It is noted in AG that except for extremely large sample sizes, a small value of r is preferred because the variance of \hat{d}_r for fixed m increases quickly as r increases. The need for a large value of r is evident if $f^*(0)$ is close to unboundedness due to large nonzero derivatives of f^* on all positive even orders at zero. Hence, for an ARFIMA(1, d , 0) process with AR parameter $\phi \rightarrow 1$, it is preferable to choose $r = 2$ to $r = 1$ for the estimator \hat{d}_r . On the other hand, an ARFIMA(0, d , 1) process with MA parameter $\theta > 0$ does not have the problem of unboundedness at $f^*(0)$, indicating that \hat{d}_1 is preferred. It is believed that with the optimal bandwidth and the proper selection of r , the BRLP model can be improved and \hat{d}_r can be an efficient estimator for the long-memory parameter.

4. Monte Carlo Experiment

In this section, we compare the finite sample behaviour of the BRLP estimators \hat{d}_1 and \hat{d}_2 using the plug-in MSE optimal choice of bandwidth and the bandwidth size following the proposed method outlined in Section 0. We consider stationary Gaussian ARFIMA (1, d , 1) processes with AR parameter ϕ and MA parameter θ . The time series generated takes the form in Eq.(4). Without loss of generality, the series is normalized to zero-mean.

$$(1 - \phi L)(1 - L)^d Y_t = (1 + \theta L)\varepsilon_t, t = 1, \dots, n, \quad (4)$$

where L is a backshift operator and ε_t is an iid standard normal random variable.

Focusing on the scenario of long-memory with mild short-range effect, we consider the processes with the combination of parameters $d = 0, .2, .4$, $\phi = 0, .1, .2, .3, .4$ and $\theta = 0, .1, .2, .3, .4$. We do not examine the cases with negative parameters as these cases are of low empirical relevance. To examine if the performance of the proposed bandwidth selection is consistent, we run the simulation in sample sizes $n = 512, 1000$ and 2000 . We generate 1000 realizations for each run of combination, and the procedure is repeated for 100 replicates. To evaluate the performance of the bandwidth selection method in the BRLP estimators, the average of root mean squared error (RMSE) of these 100 replicates, each with 1000 realizations, is obtained for each of the processes considered in this study. These results are compared to the average of the minimum RMSE that tracks the BRLP estimates within the bandwidth size from $m = 10, 11, \dots, \left[\frac{n}{2}\right]$. We report in Figures 3 and 4 the results for ARFIMA(1, d , 0), $\phi = 0, .1, .2, .3, .4$ and ARFIMA(0, d , 1), $\theta = 0, .1, .2, .3, .4$ respectively with $d = .2, .4$ and $n = 512, 2000$. It can be seen that in both ARFIMA(1, d , 0) and ARFIMA(0, d , 1) processes, the proposed bandwidth produces the average RMSE's that are closer to that of the minimum RMSE that examine \hat{d}_r over

the bandwidth size from $m = 10, 11, \dots, \lfloor \frac{n}{2} \rfloor$. Figure 5 reports that the proposed method consistently suggests a bandwidth size that is closer to the one that gives the minimum RMSE in the Monte Carlo simulations, of which \hat{d}_2 is preferred for the ARFIMA $(1, d, 0)$ processes, especially when $\phi \geq .2$. The plug-in MSE-optimum bandwidth (AG method) is rather insensitive towards ϕ 's (and θ 's), and the bandwidths suggested are relatively small. Indeed, the average RMSE's due to AG method in Figure 3 are for \hat{d}_1 , as this method reports \hat{d}_1 over \hat{d}_2 owing to the small bandwidth. In other words, the selection of estimator does not follow the choice of the one that minimizes the Monte Carlo RMSE. Interestingly, even in the case without long memory effect, the proposed method suggests the bandwidths that are closer to the one that gives the minimum RMSE in the Monte Carlo simulations, hence, giving the smaller average RMSE compared to the plug-in AG method (see Table 1). These results demonstrates that with the mild short-range component, the long-memory parameter should be estimated using a larger bandwidth size, which takes almost all signals except the noise.

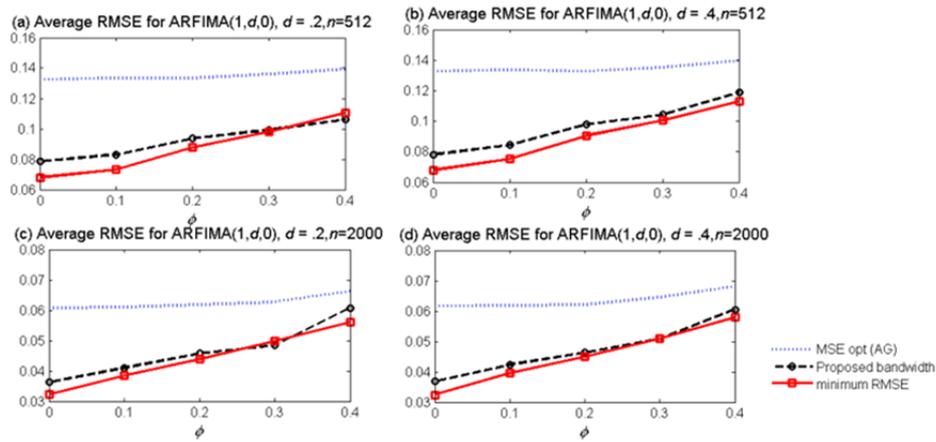


Figure 3. Average RMSE of the BRLP estimators with the plug-in MSE optimal bandwidth (AG) and the proposed bandwidth selection method compared to the minimum RMSE for ARFIMA(1, d, 0) processes

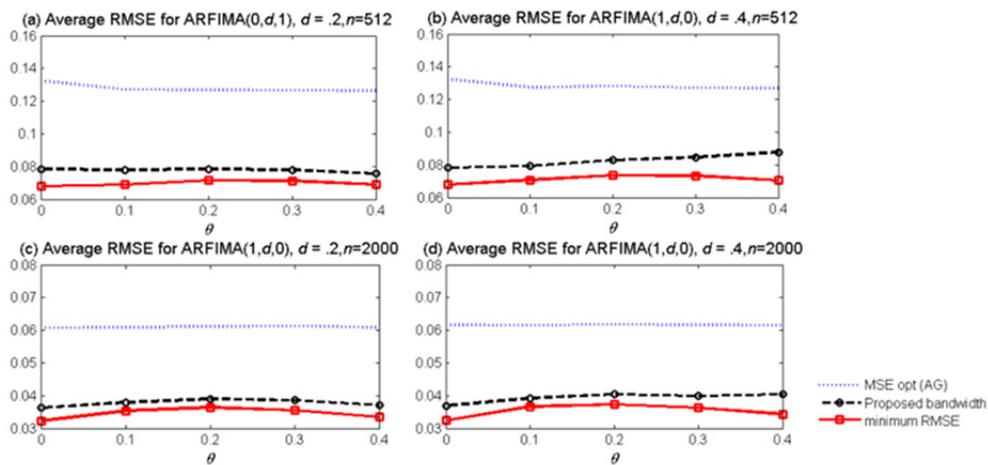


Figure 4. Average RMSE of \hat{d}_1 with the plug-in MSE optimal bandwidth (AG) and the proposed bandwidth selection method compared to the minimum RMSE for ARFIMA(0, d, 1) processes

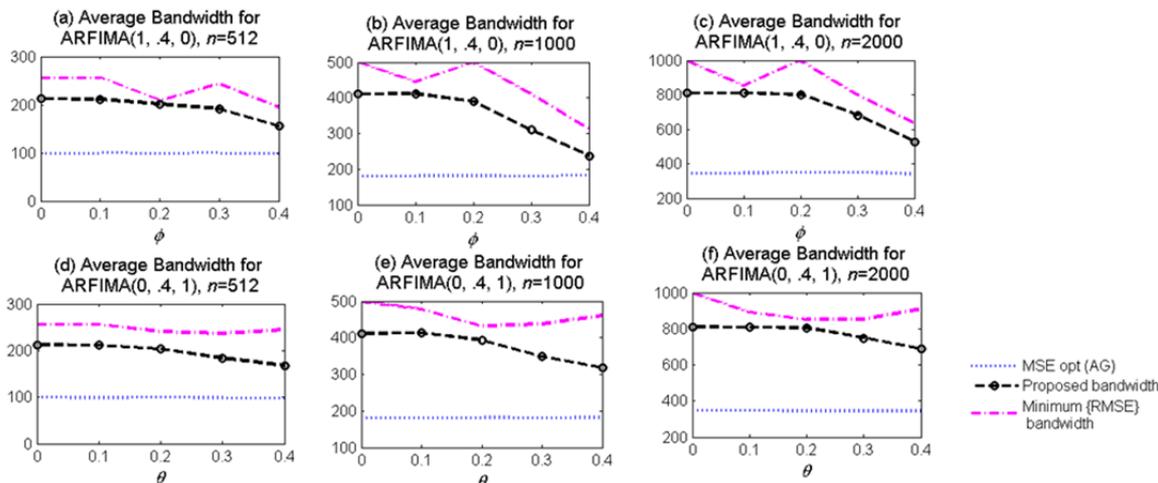


Figure 5. Average bandwidth size of various selection methods

Table 1. Average RMSE and the bandwidth size used for ARFIMA(1,0,0), $n = 2000$

ϕ	Minimum RMSE	MSE-optimum (AG)	Proposed bandwidth
0	.0323 (988)	.061 (348)	.0361 (822)
.1	.037 (851)	.061 (348)	.0406 (810)
.2	.0434 (958)	.061 (348)	.0456 (799)
.3	.0488 (798)	.0633 (347)	.0477 (788)
.4	.0549 (633)	.0661 (344)	.0575 (786)

Note. Indicates the bandwidth size.

5. Empirical Examples

This section shows the application of the proposed bandwidth selection strategy to two real time series, namely the Nile river minimum water levels during years 622 through 1284 and the de-seasonalized volatility of the 5-minute returns of Kuala Lumpur Composite Index (FBMKLCI) from 30th April 2013 to 31st December 2013. The first time series consists of 663 observations and it has been widely discussed in the long memory literature. The second series is an example of high-frequency data on returns of financial assets which spurs a large amount of research relating to modeling and predicting the realized volatility. Both series are displayed in Figures 6 and 8. Based on the plots of the average modified periodogram, ACF and PACF in these figures, we can conclude that both the time series have long memory with a very weak short memory component. In each series, the long memory parameter is estimated using \hat{d}_1 , and the performance of the proposed bandwidth is checked with that of the plug-in MSE-optimal bandwidth. The local bootstrap-based bandwidth is used as a means of substantiation.

For the Nile River data, the procedure in Section 0 suggests an optimal bandwidth $\hat{m} = 301$ giving $\hat{d}_1 =$

.3785. The bandwidth estimate of the MSE-optimal choice is $\hat{m} = 111$ and the corresponding long memory parameter estimate is $\hat{d}_1 = .4578$. Taking the idea of approximating the MSE by a bootstrap MSE, a bandwidth estimate using local bootstrap by Arteche & Orbe (2009) is referred. To focus around the bandwidth estimates of the proposed method and the MSE-optimal choice, the range of bandwidths considered for optimization is set at $100 \leq m \leq 320$, and the resampling width is $k_n = 5$. The optimum bandwidth in terms of MSE is $\hat{m} = 306$, leading to $\hat{d}_1 = .3795$ with a bootstrap MSE of .0027. The local bootstrap-based bandwidth is very close to the proposed bandwidth, suggesting that a larger bandwidth size gives a better long memory parameter estimate in terms of MSE. To further examine the performance of these procedures, each pair of estimates (\hat{m}, \hat{d}_1) is used to compute the spectral density in Eq.(1) following the approach of Delgado & Robinson (1996). The spectral density estimates are compared to the plot of periodogram shown in Figure 7. As the spectral density estimate due to the proposed bandwidth is closer to the periodogram, this procedure of bandwidth search for the BRLP estimator seems to provide the more plausible approximation.

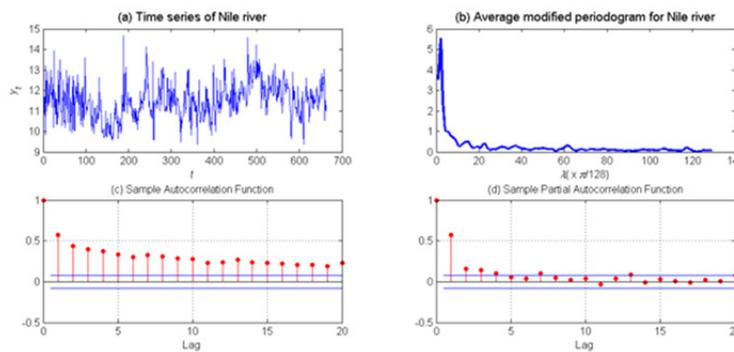


Figure 6. Characteristics of Nile River minimum water levels during years 622 through 1284

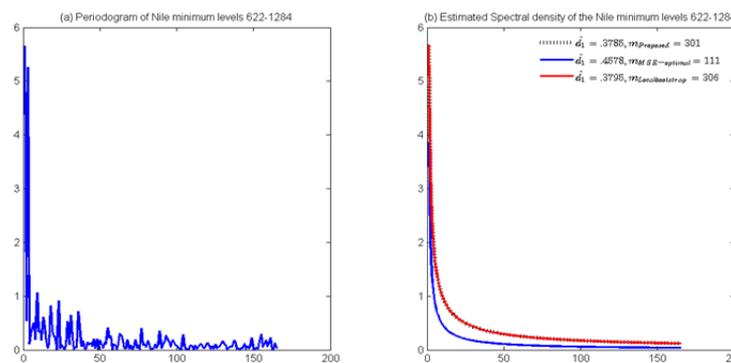


Figure 7. Periodogram and spectral density estimates for Nile River minimum water levels

The 5-minute price data of Kuala Lumpur Composite Index (FBMKLCI) from 30th April 2013 to 31st December 2013 consists of $n = 12,096$ observations with 72 price indexes per day. The 5-minute returns are computed as the difference of the log prices, that is $r_t = \ln(p_t) - \ln(p_{t-1})$, $t = 2, \dots, n$, and the series of log-squared returns is taken as a proxy for the volatility of FBMKLCI. Following the spirit of Deo et al. (2006), the seasonality of the log-squared returns is expected to be periodic with a period of 72. The series of volatility is thus de-seasonalized, and the graphical analysis of the series is depicted in Figure 8. The results of the long memory parameter estimation corresponding to the bandwidth estimates due to the proposed method and the MSE-optimal choice are shown in Table 2.

As the difference between the bandwidths suggested by the proposed method and the MSE-optimal choice is large, we may not be able to obtain a satisfactory local bootstrap-based bandwidth from the interval that spans between these numbers. The extended residuals in the local bootstrap procedure show a marked structure, and hence, a narrower search interval is deemed necessary. To examine the performance of the bandwidth choices,

we execute the local bootstrap twice, one targeting the search interval $m \in [1300, 1800]$ and another one $m \in [4200, 4700]$, both using a resampling width $k_n = 5$. The extended residuals of both cases showed stable behaviour (see Figure 9). Based on the bootstrap MSE in Table 2, it can be concluded that the proposed bandwidth suggests a more reasonable approximation, whilst the MSE-optimal choice bandwidth seems too low. The performance of these procedures are further gauged by their spectral density estimates, compared to the periodogram of the de-seasonalised volatility of FBMKLCI (see Figure 10). Similar to the Nile River data, the spectral density estimate with the proposed bandwidth is closer to the periodogram. These examples confirm that in a long memory time series with weak short memory component, the MSE-optimal choice of bandwidth may be too low. Alternatively, an adequate bandwidth for the BRLP estimator can be obtained by including all the Fourier frequencies except those of which the spectrums are low enough to be treated as noise. The proposed method of bandwidth selection clearly gives a better long memory parameter estimate in terms of MSE.

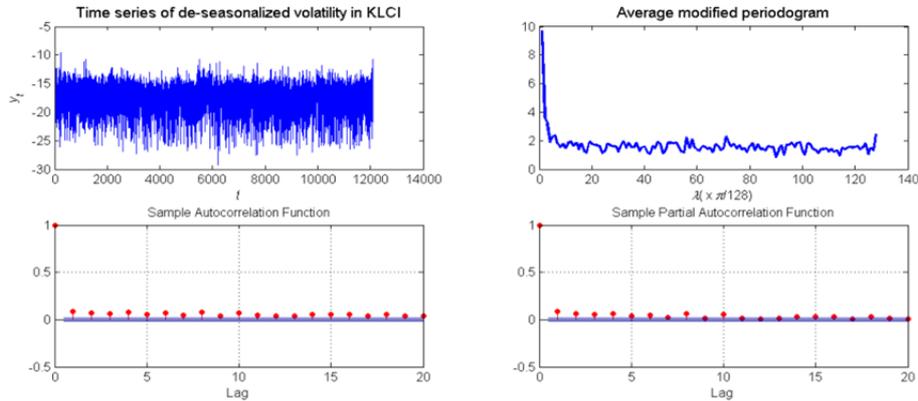


Figure 8. Characteristics of the de-seasonalized volatility of the 5-minute returns of Kuala Lumpur Composite Index (FBMKLCI) from 30th April 2013 to 31st December 2013

Table 2. Results for the bandwidth selection methods, the corresponding long memory parameter estimates and the validation via local bootstrap for the de-seasonalized volatility of FBMKLCI

	\hat{m}	\hat{d}_1	Bootstrap MSE
Proposed method	4668	.136	-
MSE-optimal choice	1356	.2687	-
Local bootstrap	4523 ^(a)	.177 ^(a)	1.97×10^{-4}
	1672 ^(b)	.249 ^(b)	4.26×10^{-4}

(a) \hat{m} is searched from the interval $[4200, 4700]$ with $k_n = 5$.

(b) \hat{m} is searched from the interval $[1300, 1800]$ with $k_n = 5$.

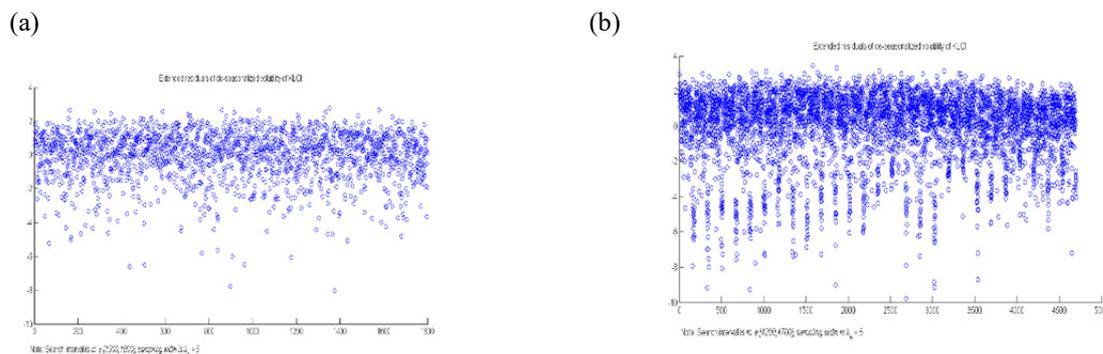


Figure 9. Extended residuals of the local bootstrap for the bandwidth search around the proposed bandwidth and the MSE-optimal bandwidth choice for the de-seasonalized volatility of FBMKLCI

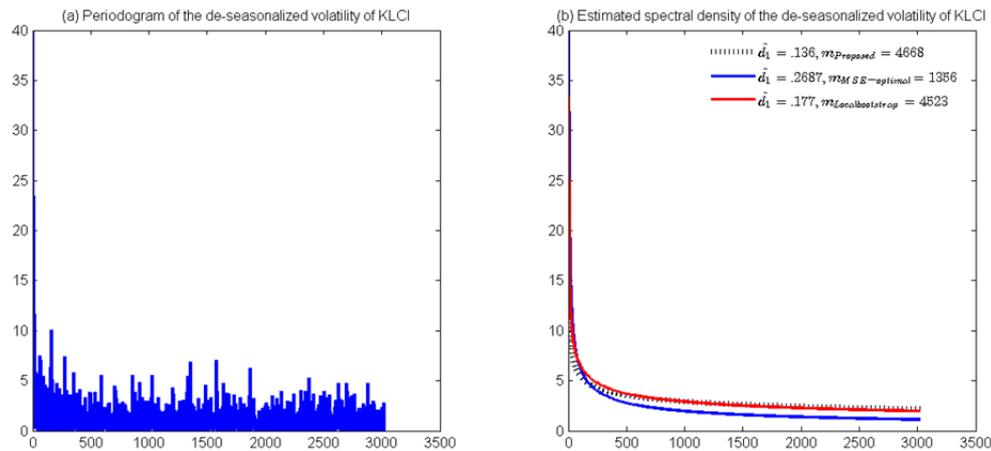


Figure 10. Periodogram and spectral density estimates for the de-seasonalized volatility of the 5-minute returns of FBMKLCI

6. Conclusion

BRLP estimator \hat{d}_r reduces the bias in \hat{d}_{GPH} , but its performance depends largely on the bandwidth size m and the order r . Working on the long-memory process with weak short-range component, this paper proposes to take as much the Fourier frequencies in the BRLP regression model, excluding the high frequencies that correspond to low spectrums that are associated with noise. The Monte Carlo simulation results confirmed the theoretical analyses, which suggest that for a mild short-range ARFIMA(1, d , 0) with large sample size, \hat{d}_2 using the proposed bandwidth gives a good estimate for the long-memory parameter, whereas for a mild short-range ARFIMA(0, d , 1), \hat{d}_1 with the proposed bandwidth is preferred. The results are quite consistent throughout the degrees of long-memory and the sample sizes. The advantage of the proposed method for the bandwidth selection in the long memory series with very weak short memory component is demonstrated in the empirical examples, of which the proposed bandwidth performs better in terms of MSE, and it gives a closer estimate to the spectral density. It is believed that the accuracy in the long memory parameter estimation can greatly improve the analysis of economic time series as pertains to modeling and forecasting.

Acknowledgments

This research is funded by Fundamental Research Grant Scheme (FRGS) FRGS/1/2014/SG04/UNIM/03/1.

References

- Agiakloglou, C., Newbold, P., & Wohar, M. (1993). Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis*, 14, 235-246. <http://dx.doi.org/10.1111/j.1467-9892.1993.tb00141.x>
- Andrews, D. W. K., & Guggenberger, P. (2003). A bias-reduced log-periodogram regression estimator for the long-memory parameter. *Econometrica*, 71(2), 675-712. <http://dx.doi.org/10.1111/1468-0262.00420>
- Arteche, J. (2004). Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. *Journal of Econometrics*, 119, 131-154. [http://dx.doi.org/10.1016/S0304-4076\(03\)00158-1](http://dx.doi.org/10.1016/S0304-4076(03)00158-1)
- Arteche, J., & Orbe, J. (2009). Bootstrap-based bandwidth choice for log-periodogram regression. *Journal of Time Series Analysis*, 30(6), 591-617. <http://dx.doi.org/10.1111/j.1467-9892.2009.00629.x>
- Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73, 5-59.
- Barros, C. P., Gil-Alana, L. A., & Payne, J. E. (2014). Long range dependence and breaks in energy prices. *Energy Sources, Part B: Economics, Planning, and Policy*, 9(2), 196-206. <http://dx.doi.org/10.1080/15567249.2012.753959>
- Bollerslev, T., Osterrieder, D., Sizova, N., & Tauchen, G. (2013). Risk and return: Long-run relations, fractional cointegration, and return predictability. *Journal of Financial Econometrics*, 108, 409-424. <http://dx.doi.org/10.1016/j.jfineco.2013.01.002>
- Charfeddine, L. (2014). True or spurious long memory in volatility: further evidence on the energy futures

- markets. *Energy Policy*, 71, 76-93. <http://dx.doi.org/10.1016/j.enpol.2014.04.027>
- Chatfield, C. (2004). *The Analysis of Time Series, An Introduction* (6th ed.). New York: Chapman & Hall.
- Choi, K., Yu, W.-C., & Zivot, E. (2010). Long memory versus structural breaks in modeling and forecasting realized volatility. *Journal of International Money and Finance*, 29(5), 857-875. <http://dx.doi.org/10.1016/j.jimonfin.2009.12.001>
- Delgado, M. A., & Robinson, P. M. (1996). Optimal spectral bandwidth for long memory. *Statistica Sinica*, 6, 97-112.
- Deo, R., Hurvich, C.M., & Lu, Y. (2006). Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment. *Journal of Econometrics*, 131, 29-58. <http://dx.doi.org/10.1016/j.jeconom.2005.01.003>
- Garvey, J., & Gallagher, L. A. (2013). The economics of data: using simple model-free volatility in a high-frequency world. *North American Journal of Economics and Finance*, 26, 370-379. <http://dx.doi.org/10.1016/j.najef.2013.02.011>
- Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long-memory time series models. *Journal of Time Series Analysis*, 4, 221-238. <http://dx.doi.org/10.1111/j.1467-9892.1983.tb00371.x>
- Giraitis, L., Robinson, P. M., & Samarov, A. (2000). Adaptive semiparametric estimation of the memory parameter. *Journal of Multivariate Analysis*, 72, 183-207. <http://dx.doi.org/10.1006/jmva.1999.1865>
- Granger, C. W. J., & Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15-39. <http://dx.doi.org/10.1111/j.1467-9892.1980.tb00297.x>
- Henry, M. (2001). Robust automatic bandwidth for long memory. *Journal of Time Series Analysis*, 22(3), 293-316.
- Henry, M., & Robinson, P. M. (1996). Bandwidth choice in Gaussian semiparametric estimation of long range dependence. In P. M. Robinson & M. Rosenblatt (Eds.), *Athens Conference on Applied Probability and Time Series, Lecture Notes in Statistics, 115, Vol II* (pp. 220-232). New York: Springer.
- Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68, 165-176.
- Hurvich, C. M., & Beltrao, K. I. (1993). Asymptotics for the low-frequency ordinates of the periodogram of a long memory time series. *Journal of Time Series Analysis*, 14, 455-472. <http://dx.doi.org/10.1111/j.1467-9892.1993.tb00157.x>
- Hurvich, C. M., & Deo, R. S. (1999). Plug-in selection of the number of frequencies in regression estimates of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, 20, 331-341. <http://dx.doi.org/10.1111/1467-9892.00140>
- Hurvich, C. M., Deo, R. S., & Brodsky, J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, 19, 19-46. <http://dx.doi.org/10.1111/1467-9892.00075>
- Jenkins, G. M., & Watts, D. G. (1968). *Spectral Analysis and Its Applications*. Holden-Day.
- Kunsch, H. (1987). Statistical aspects of self-similar processes. In Y. Prohorov & V. V Sazanov (Eds.), *Proceedings of the First World Congress of the Bernoulli Society* (pp. 67-74). Utrecht: VNU Science Press.
- Kunsch, H. R. (1986). Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability*, 23, 1025-1030. <http://dx.doi.org/10.2307/3214476>
- Qu, Z. (2011). A test against spurious long memory. *Journal of Business and Economics Statistics*, 29(3), 423-438. <http://dx.doi.org/10.1198/jbes.2010.09153>
- Robinson, P. M. (1994). Semiparametric analysis of long-memory time series. *The Annals of Statistics*, 22(1), 515-539.
- Robinson, P. M. (1995a). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, 23, 1630-1661.
- Robinson, P. M. (1995b). Log-periodogram regression of time series with long-range dependence. *Annals of Statistics*, (23), 1048-1072.
- Taqqu, M. S., & Teverovsky, V. (1996). Graphical methods for bandwidth selection in semiparametric estimation

of long memory. In P. M. Robinson & M. Rosenblatt (Eds.), *Athens Conference on Applied Probability and Time Series Analysis, Vol II: Time Series Analysis* (pp. 420-432). New York: Springer.

Welch, P. D. (1967). The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans, Audio and Electroacoust, Au-15*, 70-73.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).