Journal of Scientific Research and Development 2 (10): 86-92, 2015
Available online at www.jsrad.org
ISSN 1115-7569

Bandwidth selection method for bias-reduced log-period gram estimator in decaying spectral density

G.L.C. Yap*, K.W. Liew

Department of Applied Mathematics, the University of Nottingham Malaysia Campus, Semenyih, Malaysia

Abstract: Long-memory parameter estimation using bias-reduced log-period gram regression (BRLP) is proven efficient as it eliminates the first and higher order of biases of the (Geweke & Porter-Hudak, 1983) (GPH) estimator. Nonetheless, its performance relies largely on the frequency bandwidth and the order of estimation. Literature suggests a data-dependent plug-in method for selecting the frequency bandwidth that minimizes the asymptotic mean-squared error (MSE). The optimal rate of convergence to zero of the MSE is faster than that of the GPH and the other semi-parametric estimators when the normalized spectrum at zero is sufficiently smooth. However, this choice of bandwidth significantly increases the MSE's over the finite sample minimum MSEs due to the non-parametric estimation problem in the unknown term within the plug-in method. To obtain the optimum bandwidth in the decaying spectral density, this paper suggests an alternative approach that relies on spectral analysis, with the idea of low pass filter applied in signal processing. Monte Carlo simulation results for stationary ARFIMA (1, *d*, 0) processes show that the proposed method for the bandwidth selection perform well relative to the MSE optimal choice of bandwidth, and the estimation performance is improved with the sample size.

Key words: Long-memory parameter estimation; Bias reduction; Frequency domain; Spectral density

1. Introduction

Semi-parametric estimation procedures are common in the time series analysis of financial measurements sampled at high frequencies (Barros et al., 2014; Bollerslev et al., 2013; Garvey & Gallagher, 2013). Following these methods, the long-range characteristics (low frequency behavior) of the time series can be estimated without the knowledge of the short-range (high frequency) structure. One of the popular tools for long-memory estimation in empirical research is log-period gram regression (LP) proposed by (Geweke & Porter-Hudak, 1983; Robinson, 1995) (GPH) due to its simple implementation, pivotal asymptotic normality and robustness as a result of the local condition (Arteche & Orbe, 2009).

The GPH estimator $\hat{d}$ is the least squares estimate of the long memory parameter d, in the regression model that takes the first m harmonics of the logged period gram against a simple function of Fourier frequency. It has been criticized due to its finite-sample bias (Agiakloglou et al., 1993). To overcome this, Andrews & Guggenberger (2003) proposed a bias-reduced log-period gram estimator (BRLP) $\hat{d}_r$ which is the same as $\hat{d}$ except that it includes additional regressors in the form of the Fourier frequencies to the power $2k$ for $k = 1, \dots, r, r \in Z^+$, in the pseudo-regression model. The performance of the estimator is usually evaluated based on the mean squared error (MSE) or root mean-squared error (RMSE), of which estimator with the minimal MSE (or RMSE) is preferred. 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, an optimal bandwidth, that is, an m value that minimizes the MSE or RMSE is sought.

There are basically three approaches to determine optimal bandwidth, namely the plug-in method that minimizes an asymptotic approximation of the MSE (Hurvich & Deo, 1999), the adaptive LP (Giraitis, Robinson, & Samarov, 2000) that uses an adaptive LP that adapt to an unknown local to zero spectral smoothness, and a bootstrap-based bandwidth choice (Arteche & Orbe, 2009) that minimizes a bootstrap MSE. The adaptive LP method does not produce optimal bandwidth but it only offers bandwidths with optimal growth rate which can be changed arbitrarily, whilst the bootstrap-based method is rather tedious as the optimal bandwidth is obtained by searching for the minimum bootstrap MSE amongst the points within a predefined interval of bandwidths. The plug-in method is easy to implement but it is usually not adequate as it depends on unknowns to be estimated (Andrews & Guggenberger, 2003; Arteche, 2004; Delgado & Robinson, 1996; Henry & Robinson, 1996; Henry, 2001). This paper suggests an alternative approach to identify the optimum bandwidth for the

* Corresponding Author.

long-memory parameter estimation using the BRLP estimator. This approach examines the long- and short-range effect on the spectral density of which optimal bandwidth is determined such that the short-range effect and the noise are filtered out. The following section reviews the BRLP model. Section 3 considers the proposed method for the bandwidth selection. Section 4 provides a simulation study in the finite sample, and section 5 offers the concluding remarks.

2. Literature review

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

$$f(\lambda) = |\lambda|^{-2d} f^*(\lambda) \quad (1)$$

Where d is the long-memory parameter and $f^*(\cdot)$ 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 and q-order moving average, abbreviated as ARFIMA (p, d, q) introduced by Granger & Joyeux (1980) and Hosking (1981) is a special case of long-range process satisfying Eq. (1). The long-memory parameter can be estimated using the first m ($m < n$) period grams in the log-period gram regression proposed by Geweke & Porter-Hudak (1983) (GPH). Robinson (1995) writes this model in the form of Eq.(2).

$$\log I(\lambda_j) = (\log f^*(0) - C) + dX(\lambda_j) + \log\left[\frac{f^*(\lambda_j)}{f^*(0)}\right] + \varepsilon_j \quad (2)$$

Where

$I(\lambda_j) = \frac{1}{2n\pi}\left|\sum_{t=1}^{n} Y_t \exp(i\lambda_j t)\right|^2$ for $j = 1, \dots, m$, $\lambda_j = \frac{2\pi j}{n}$, $X(\lambda_j) = -2\log|\lambda_j|$, $\varepsilon_j = \log\left[\frac{I(\lambda_j)}{f(\lambda_j)}\right] + C$, and $C = 0.577216\dots$ is the Euler constant.

The GPH estimator $\hat{d}$ is the least squares estimator of d in the regression model Eq. (2). The dominant bias comes from the term $\log\left[\frac{f^*(\lambda_j)}{f^*(0)}\right]$. To eliminate this bias term, regressors $\lambda_j^{2r}, r \geq 1$, are added to the pseudo-regression model Eq.(2), giving rise to the bias-reduced log-period gram regression (BRLP) (Andrews & Guggenberger, 2003). Assuming $f^*(\lambda)$ is smooth of order s at zero for some $s \geq 1$ BRLP is written in Eq.(3).

$$\log I(\lambda_j) = (\log f^*(0) - C) + dX(\lambda_j) + \sum_{k=1}^{\min\left([\frac{s}{2}], r\right)} \frac{b_{2k}}{(2k)!}\lambda_j^{2k} + R_j + \varepsilon_j, \quad j = 1, \dots, m \quad (3)$$

Where

R_j = error term other than ε_j

$$b_k = \frac{d^k}{d\lambda^k}\log f^*(\lambda)\Big|_{\lambda=0}$$

$[\cdot]$ is the greatest integer function

Similar to the GPH estimator, $\hat{d}_r$ is the least squares estimator of the coefficient on $X(\lambda_j)$ in the BRLP model. The problem of long-memory parameter estimation lies in the uncertain number of frequencies needed to estimate the behavior of $f^*(\lambda)$ for $\lambda \to 0$. The optimal choice of the bandwidth proposed by Andrews & Guggenberger (2003) (hence called AG method) is the m value that minimizes the asymptotic MSE of $\hat{d}_r$. However, this method involves an unknown b_{2k} that can only be estimated non-parametrically, and its rate of convergence is quite slow especially when $r > 0$. As a result, it may lead to a large MSE in $\hat{d}_r$. On the other hand, there are other challenges in the bandwidth selection, especially when there is a sufficient density of individuals with close-to-unit-root behavior which produces an aggregate long memory (Robinson, 2003). Alternative to AG method, graphical method by Taqqu & Teverovsky (1996) seems reasonable. This method argues that at large bandwidth m, the estimates of d are incorrect due to the short-range effects. As m decreases, the short-range effects disappear and $\hat{d}_r$ should represent the true long-memory dependence. Nonetheless, this causes instability in the estimates of d due to insufficient frequencies for the log-period gram regression. Hence, the optimum bandwidth is identified as a point from the flat region in the plot of $\hat{d}_r$ against m. Unfortunately, as argued by Henry (2001), the flat region in the plot is not always obvious. As such, to overcome these difficulties, a more efficient approach for the data-dependent choice of bandwidth is desired.

3. Bandwidth selection for BRLP using the spectral analysis and low pass filter

This paper suggests collecting the Fourier frequencies near the origin that have significant spectrum. The spectrum of a time series is the distribution of variance of the series as a function of frequency (Chatfield, 2004). A peak in the spectrum represents relatively high variance in the respective frequency band, whilst frequencies with small spectrum contain no significant signal and a flat spectrum indicates that the variance is evenly distributed over the frequencies. A time series with long-range characteristic has positive autocorrelation with low-frequency spectrum, which means the spectral density has a decaying pattern with the variance tends to be higher at the low frequencies. Hence, an appropriate amount of low frequencies need to be determined for $\hat{d}_r$ such that the frequency band contains sufficient information for the long-range characteristic, yet it is not contaminated with the signals from the short-memory traits.

To improve the spectral estimation, this paper uses the average modified period gram $S(\lambda)$ (Welch,

1967) to estimate the spectral density. For real-valued time series, $S(\lambda)$ is produced with a frequency resolution of $\frac{2\pi}{nfft}$, where $nfft$ = max{256, the next power of 2 greater than the length of the segments}. In other words, spectral density is estimated by $S(\lambda_i)$, i = 0,1, … , r_s, where $r_s = \lfloor\frac{nfft}{2}\rfloor$. $S(\lambda)$ is used to identify the frequency band close to origin that contains large spectrums so that sufficient information can be assembled for $\hat{d}_r$ in the BRLP model.

As long-memory parameter determines the low frequencies properties of a series, this paper proposes the use of a low pass filter (Van de Vegte, 2002). This requires a cutoff frequency λ_c to be identified, and all frequencies of the input signal in the interval $|\lambda| \le \lambda_c$ are passed with equal gain and all the frequencies outside this interval (which are related to the short-term correlation structure) are completely filtered out (Shenoi, 2006). As such, the accuracy of $\hat{d}_r$ relies heavily on the choice of the cutoff frequency. The procedure to identify the cutoff frequency is detailed below.

Working on $S(\lambda)$, the procedure begins with an arbitrary potential optimum frequency λ_{po}. Based on the remark by Qu (2011), frequencies in the neighborhood of $\frac{2\pi}{\sqrt{n}}$ suggest the possibility of distinguishing between the true and the spurious long-memory from the spectral domain. This leads to setting a potential stop point $po = \lfloor\frac{2r_s}{\sqrt{n}}\rfloor$, which corresponds to bandwidth $m_{po} = \lfloor\sqrt{n}\rfloor$ suggested by Geweke & Porter-Hudak (1983). Working from r_s and regress up to $po - 5$, the cutoff frequency λ_c is determined as the first frequency that exceeds the minimum criteria, $cri = \min\{S(\lambda_a)\}$ + k*range{$S(\lambda_a)$}, $a = po - 5, \ldots, r_s$, where k is a proportion factor. As the spectral density of a long-memory process is a decreasing function, this is somewhat equivalent to discarding the signal that has less than k% of the information of the long memory process. This rule seems reasonable to give an appropriate cutoff frequency, as the error in the log-period gram regression increases for frequencies far from origin. However, following Arteche & Orbe (2009), it is noted that the enlargement of the error is significant when the characteristic polynomial of the autoregressive (AR) process has a root that is close to unity. This indicates that some intervention to this rule is needed. As these cases are associated with slow decaying auto-correlation and monotonic decaying spectral density, the characteristics in ACF and $S(\lambda)$ plots need to be identified. An interaction of long memory parameter (d) and the AR parameter (ϕ) gives the characteristics as described in Table **1**.

Table 1: Characteristics in ACF and $S(\lambda)$ due to the parameters d and ϕ

	large	small
large	ACF: very slow decay ( ): extremely fast decay	ACF: very slow decay ( ): fast decay
small	ACF: slow decay ( ): fast decay	ACF: no obvious decaying pattern ( ): rather flat

The auto-correlation in a rapid decaying ACF converges to zero after a few low-order lags but a very slow decaying ACF is likely to have significant auto-correlations even after 10 lags. By checking on the ACF up to 20 lags, we define an ACF as very slow decay if none of the auto-correlations is less than 1.1*standard error of the ACF. The process with possibly large ϕ can be identified as the one that depicts rather a monotonic decay, that is, most of the differences between the consecutive auto-correlations $r_i - r_{i-1}$, i = 2, … , 20 are negative, and a polynomial fit to the ACF should return an order (ord) that is not higher than 3. On the other hand, based on some simulation results in Section 4, it is noticed that a fast decaying $S(\lambda)$ converges before the frequency $\lambda = \frac{\pi}{4}$. We proposed to measure the characteristic of fast decaying $S(\lambda)$ using (i) the range of $S(\lambda_j)$, $j = po - 5, \ldots, r_s$, (ii) the number of relative signal that is less than .001, $rs_j = \frac{S(\lambda_j)}{\max\{S(\lambda)\}} <$ .001, $j = po - 5, \ldots, r_s$, and (iii) the mean of the relative signals in the interval $j = \lfloor\frac{r_s}{10}\rfloor, \ldots, \lfloor\frac{r_s}{4}\rfloor$. The following section examines the relationship between these statistics and the parameters d and ϕ, and subsequently, an alternative bandwidth selection in relation to these parameters is suggested.

4. Monte Carlo experiment

4.1. Alternative procedure for optimum bandwidth

This section begins with a data generation according to a stationary Gaussian ARFIMA $(1, d, 0)$ process with the AR parameter ϕ. The time series generated takes the form in Eq.(4). Without loss of generality, the series is normalized to zero-mean.

$$(1 - \phi L)(1 - L)^d Y_t = \varepsilon_t, \quad t = 1, \ldots, n \qquad (4)$$

Where ε_t = iid standard normal random variable

Before the optimum bandwidth can be determined, we examine the performance of the markers in the plots of ACF and $S(\lambda)$ due to various values of ϕ's and ds. We examine the cases with d = 0, .1, ... , .4 and the AR coefficient taking the values ϕ = 0, .1, ... , .9 with sample size n = 512. Each combination is repeated with 1000 replicates. In each replicate, the markers in these plots, namely the proportion of the negative differenced auto-correlations (pn), the mean of relative signals (rs), the number of relative signals that are less than .001 (N_{rs}), and the range are computed. Taking the averages of all replicates, the plots are concluded with the statistics (i) Avr_pn, (ii) Avr_rs, (iii) Avr_N_{rs}, and (iv) Avr_range, shown in Fig. 1). Although a short memory AR process with a large ϕ gives the similar effect like a long memory process with small d, they can be identified using the results in Fig.Fig. 1), which is then summarized as a combination of markers shown in Table 2.

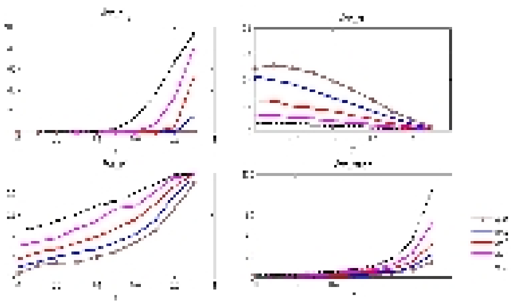

Fig. 1: Statistics related to ACF and $S(\lambda)$ for ARFIMA (1,d,0), n =512

Table 2: Markers to identify moderate to large ϕ

		range	N_{rs}	rs	pn	ord	ACF_0
large	large	≥40	≥40	≤.04	≥.8	≤3	No
	small	[10,40]	≤5	≤.04	≥.8	≤3	Yes
moderate	large	[5,30]	[10,40]	≤.04	≥.8	≤3	No
	small	[1,10]	≤5	[.05,.4]	<.8	>3	Yes

*ACF_0 denotes a case where there is no autocorrelation in the first 20 lags such that ≤1.1 , =1,...,20 where is the standard error of the autocorrelations

With the aim to identify the optimum bandwidth for the ARIMA(1, d, 0) process with possibly large ϕ, we examine the relationship between the bandwidth size m and the parameters ϕ and d. To be consistent, we use the same model in Eq. (4), with the long memory parameter $d \in$ {0, .2, .4} and extend the sample sizes to n = {512, 1000, 2000}. The long memory parameter d is estimated based on BRLP, and the RMSE's of $\hat{d}_1$ and $\hat{d}_2$ are calculated as functions of m, for $m = 20, 21, \ldots, \left[\frac{n}{2}\right]$. In each simulation replicate, the bandwidth size that gives the minimum RMSE is obtained, and the average of these values is taken as the optimum bandwidth for the corresponding parameter combination. The results of the Monte Carlo simulation are reported in Fig. 2). To give a better comparison, we report the proportion of the bandwidth size $\frac{2m}{n}$. It can be seen that the proportion can be regarded as a function of ϕ, and it is quite consistent throughout the ds and the sample sizes. As such, we only need to guess the possible values of ϕs (using results of Fig.(1) and Table 2) in order to identify the search interval for the potential optimum bandwidth. Based on Fig.(2), we observe that when ϕ is large, the proportion of the bandwidth size is close to .1 and a moderate ϕ corresponds to a proportion of the bandwidth size about .3. As such, we force the cutoff frequency to be searched from the set {[.1 r_s], ... , r_s} for a process with a potentially large AR parameter, and {[.3 r_s], ... , r_s} for the case with moderately large AR parameter. This is in line with the goal to avoid the short-range effect which is not negligible on the low frequencies when ϕ is large. As spectrum is the variance per unit frequency, a peak in $S(\lambda)$ is regarded as the signal that stands out from the 'noise floor'. Hence, in such cases, working on the search interval, we restrict the signals for the long-range effect up to the point where the modified period gram shows an abrupt fall. To be conservative, the cutoff frequency λ_c is set at the hump immediately after this abrupt fall. Taking all the possible ϕs values, we suggest the rule to set the cutoff frequency as shown in Fig.(3). Note that in the event when the plots do not satisfy any of the pre-defined characteristics, we set a higher minimum criteria (cri) for the case suspected of strong AR(1) (with pn ≥ .8, PACF(1)≥.8 and ord ≤ 3). This is in line with the aim to force the bandwidth to be close to the origin when $\phi \to 1$.

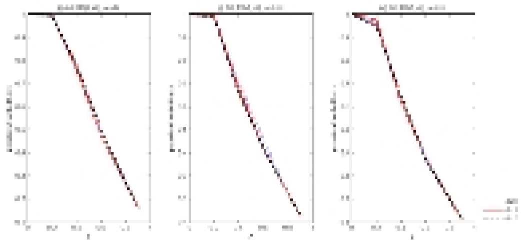

Fig. 2: Optimum proportion of bandwidth size for various parameter combinations

This paper suggests to identify the flat region in the plot of $\hat{d}_r$ against m around $\frac{\lambda_c * n}{2 * \pi}$, which is supposedly the number of frequencies that explain the variation due to the long-range dependence. The flat region is defined as the region of frequencies of which $\hat{d}_r$ are almost similar, say in a neighborhood 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-period gram regression for the long-memory parameter estimation.

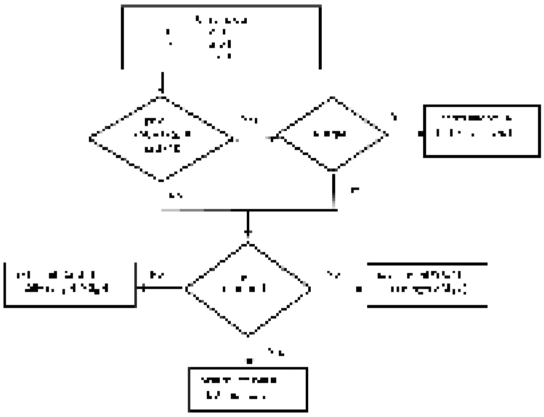

Fig. 3: Flow chart to identify the criteria for the cutoff frequency

In short, the procedure to identify the optimum bandwidth involves the steps below:
(i) Determine the possible values for ϕ based on the plots of ACF and $S(\lambda)$, refer to Table 2.
(ii) Identify the search interval based on the value of ϕ suggested in step (i). Fig.(3) is taken as a reference
(ii) Determine the cutoff frequency
(iv) Identify the optimum bandwidth (flat region around $\frac{\lambda_c * n}{2 * \pi}$)

4.2. Performance evaluation

The proposed bandwidth selection procedure is examined via the finite sample performance of the BRLP estimators $\hat{d}_1$ and $\hat{d}_2$ following the recommendation by Andrews & Guggenberger (2003) to use a relatively small value of r, such as one or two, for better finite sample performance. The results of these estimators using the proposed bandwidth selection is compared to that of the MSE optimum bandwidth selection (AG method). Their performances are gauged by the RMSE in the BRLP estimators.

To estimate the long-memory parameter in the time series with decaying spectral density, the performance of $\hat{d}_r$ using the AG method and the proposed method is examined using the data generation explained in sub-section 4.1, by which each combination is run 100 times with 1000 simulation replicates in sample sizes of 512 and 1000, and 2000. To have a close comparison, the biases, standard deviations, RMSE's are calculated, and the coverage probabilities of the nominal 95% confidence intervals (CI's) are obtained using Eq. (6.3) in the paper of (Andrews & Guggenberger, 2003), except that $z_{.95}$ is replaced by $z_{.975}$ for a 95% CI. Besides, the results are checked against the minimum RMSE in the Monte Carlo simulation replicates that examine $\hat{d}_1$ and $\hat{d}_2$ as functions of m, for m = 20, 21, ... , $\lfloor \frac{n}{2} \rfloor$.

In general, the proposed method works well with $\hat{d}_r$ by consistently returning small RMSE. Fig.(4a) shows the RMSE's of $\hat{d}_2$ of these procedures from 100 sets of experiments for the simulated ARFIMA (1, d, 0) process with ϕ = .9, d = .4 in sample size n = 512. The results of the proposed method are close to the minimum RMSE in the Monte Carlo replicates. However, since a substantial amount of frequencies are filtered out, it is straight forward to know that the parameter estimation in the case of ϕ = .9 would not be ideal, as shown by the large RMSE in Fig.(4a) and low coverage probabilities of the nominal level .95 in Fig.(4b), compared to the results of the case with ϕ = 0, d = 0 in Fig.(4c) and Fig.(4d). This is consistent with the remark by Andrews & Guggenberger (2003) that $\hat{d}_r$ does not perform well for ϕ = .9. Nonetheless, the proposed method performs better than the MSE optimal bandwidth (AG method) in both cases, and the confidence intervals produced are in general conservative.

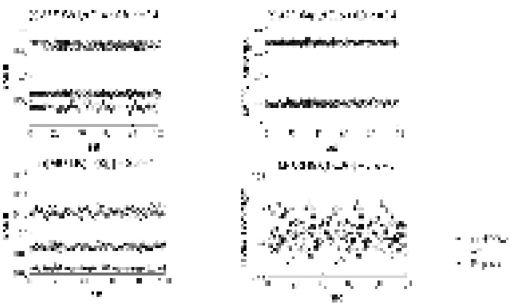

Fig. 4: RMSE and the coverage probability of $\hat{d}_1$ and $\hat{d}_2$ for ARFIMA (1, d, 0) processes respective to the bandwidth selection methods

Table 3 reports the statistics of the averages in RMSE, bias, standard deviation and the 95% CI coverage probabilities for $\hat{d}_1$ and $\hat{d}_2$ using the bandwidth suggested by AG and the proposed method for ARIMA(1, d, 0) processes with ϕ = 0, .2, .4, .6, .8, .9, d = .4 and n = 512. The results are compared with the minimum RMSE in the Monte Carlo simulation replicates. It is noted that the proposed method outperforms the MSE optimal bandwidth by consistently returning the RMSE that is closer to the minimum RMSE. Besides, it is noted that as a whole, the proposed method gives smaller bias and better coverage probability, except for the case when ϕ = .6. This is rather probable as a strong long memory process with moderate ϕ is similar to a weak long memory process with a close to unit-root AR. This leads to difficulty to correctly identify value for ϕ, and hence the search interval. On the other hand, the MSE optimal bandwidth (AG method) is not sensitive to the values in ϕ. This can be a major drawback in the long-memory parameter estimation using log-period gram regression as it is clear from Eq.(2) that function f^{*} contributes to the accuracy of the estimation. Interestingly, it is observed that when ϕ is large (.8 < ϕ < 1), the minimum RMSE values are in favor of $\hat{d}_2$ for both methods of bandwidth selection. This is rather straight forward as more terms are needed in the Taylor expansion when ϕ is large.

The experiment continues to examine the performance of the proposed bandwidth selection method in the BRLP estimator as the sample size increases. Fig.(5) shows the results of the proposed method compared to the AG method relative to the minimum RMSE for various parameter combinations of ϕ, d and sample sizes. It is observed that the proposed method consistently produces estimates with smaller RMSE (except for the case when $\phi = .6$), and the improvement in the cases with $\phi \to 1$ is significant as the sample size increases.

Table 3: Statistics of $\hat{d}_1$ and $\hat{d}_2$ for ARFIMA (1, d, 0) with d = .4 and n = 512

			MSE optimal bandwidth (AG method)					Proposed method				
		Min{RMSE}	Avg m	Avg RMSE	Avg Bias	Avg Std Dev	Avg 95% CI	Avg m	Avg RMSE	Avg Bias	Avg Std Dev	Avg 95% CI
0	1	.0657	99	.1262	.0144	.1253	.9308	147	.0979*	.0136	.0969	.9399
	2	.085	110	.1483	.0184	.1471	.9421	148	.1261	.0105	.1256	.9479
.2	1	.0864	99	.1259	.0195	.1243	.9348	150	.1114*	.0323	.1066	.8749
	2	.0878	110	.1479	.0193	.1465	.943	150	.1229	.0069	.1226	.9581
.4	1	.1131	99	.1333	.0374	.1279	.9163	117	.1282*	.0642	.1109	.9014
	2	.1114	109	.1508	.0247	.1487	.9352	118	.1422	-.0003	.1422	.9516
.6	1	.1507	96	.1645	.095	.134	.805	73	.1682	.0799	.148	.877
	2	.1482	109	.1577*	.0626	.1448	.917	74	.1914	.0114	.1911	.94
.8	1	.2303	88	.3044	.2678	.1446	.4148	56	.257	.201	.16	.7354
	2	.227	106	.2601	.2075	.1566	.65	58	.2323*	.1022	.2085	.9176
.9	1	.34	88	.5178	.4978	.1422	.0585	58	.4591	.4295	.1621	.2296
	2	.3366	104	.4624	.4335	.1609	.1829	59	.3642*	.2962	.2118	.6733

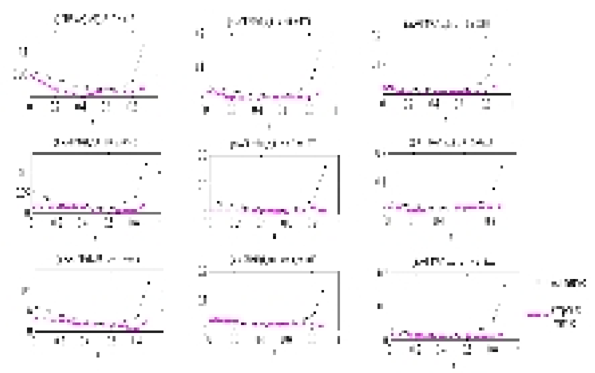

Fig. 5: Difference in average RMSE and the average of the minimum RMSE of the proposed method and the MSE optimal bandwidth method for ARFIMA (1, d, 0) with various parameter combinations for sample sizes of n = 512 , n = 1000 and n = 2000.

5. Summary

Literature shows that the BRLP estimator $\hat{d}_r$ performs well relative to the standard log-period gram regression estimator. However, its performance is largely dependent on the bandwidth selection m. The MSE optimal bandwidth selection (AG method) is not ideal as it depends on the estimation of an unknown which is a non-parametric problem. Working on the notion that spectrum is the variance per unit frequency, where a peak in the spectrum represents relatively high variance in the respective frequency band, this paper proposed a method to determine the optimum bandwidth based on the magnitude of the modified period gram. Based on the simulation results, it is observed that discarding the high frequencies that have less than 5% of the long memory information gives the optimum bandwidth in terms of minimum RMSE and better CI coverage probability, but processes with close-to-unit-root AR require a strident truncation to the bandwidth size in order to avoid the short memory contamination. This paper proposes some markers to identify the characteristics of the short and long range dependence using the plots of ACF and modified period gram. The Monte Carlo simulation results verify that by including an appropriate number of frequencies in the BLRP model, an estimator $\hat{d}_r$ with small bias yet not causing much increase in the variance can be obtained. We believe that this technique has potential applications wherever the long memory parameter estimation is of interest. The future work may include setting the optimum bandwidth for the long memory process that has an element of moving average.

Acknowledgment

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.

Andrews, D. W. K., & Guggenberger, P. (2003). A bias-reduced log-period gram regression estimator for the long-memory parameter. *Econometrica*, *71*(2), 675–712.

Arteche, J. (2004). Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. *Journal of Econometrics*, *119*, 131–154.

Arteche, J., & Orbe, J. (2009). Bootstrap-based bandwidth choice for log-period gram regression. *Journal of Time Series Analysis*, *30*(6), 591 – 617.

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.

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.

Chatfield, C. (2004). *The Analysis of Time Series, An Introduction* (6th ed.). New York: Chapman & Hall.

Delgado, M. A., & Robinson, P. M. (1996). Optimal spectral bandwidth for long memory. *Statistica Sinica*, *6*, 97–112.

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.

Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long-memory time series models. *Journal of Time Series Analysis*, *4*, 221–237.

Giraitis, L., Robinson, P. M., & Samarov, A. (2000). Adaptive semiparametric estimation of the memory parameter. *Journal of Multivariate Analysis*, *72*, 183 – 207.

Graf H.P. (1983). *Long-range correlations and estimation of the self-similarity parameter*. ETH Zürich.

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.

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., & 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.

Qu, Z. (2011). A test against spurious long memory. *Journal of Business and Economics Statistics*, *29*, 423–438.

Robinson, P. M. (1994). Semiparametric analysis of long-memory time series. *The Annals of Statistics*, *22*(1), 515–539.

Robinson, P. M. (1995). Log-period gram regression of time series with long-range dependence. *Annals of Statistics*, (23), 1048 – 1072.

Robinson, P. M. (2003). Long-memory Time Series. In P. M. Robinson (Ed.), *Time Series with Long Memory* (pp. 4 – 32). Oxford: Oxford University Press.

Shenoi, B. A. (2006). *Introduction to Digital Signal Processing and Filter Design*. Hoboken, New Jersey: John Wiley & Sons Inc.

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.

Toussoun, O. (1925). Memoire sur l'Histoire du Nil. *Memoires a l'Institut d'Egypte*, *18*, 366–404.

Van de Vegte, J. (2002). *Fundamentals of Digital Signal Processing*. Upper Saddle River, New Jersey: Prentice Hall.

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 period grams. *IEEE Trans, Audio and Electroacoust.*, *Au-15*, 70 – 73.