

# MODELING NON-STATIONARY LONG-MEMORY SIGNALS WITH LARGE AMOUNTS OF DATA

*Li Song and Pascal Bondon*

CNRS UMR 8506, Université Paris-Sud  
91192, Gif-sur-Yvette, France

phone: +33 1 69 85 17 31, fax: +33 1 69 85 17 65, email: song@lss.supelec.fr, bondon@lss.supelec.fr

## ABSTRACT

We consider the problem of modeling long-memory signals using piecewise fractional autoregressive integrated moving average processes. The signals considered here can be segmented into stationary regimes separated by occasional structural break points. The number as well as the locations of the break points and the parameters of each regime are assumed to be unknown. An efficient estimation method which can manage large amounts of data is proposed. This method uses information criteria to select the number of structural breaks. Its effectiveness is illustrated by Monte Carlo simulations.

## 1. INTRODUCTION

Stationarity plays an important role in modeling time series. However, many signals in the real life are non-stationary. Models with structural changes have been of interest to many researchers, especially in the case of detecting and estimating a single break, see for instance [1]. Studies related to multiple structural changes have received relatively less attention, see [10] for a review. A natural method for fitting a piecewise parametric model with multiple structural changes to data consists in minimizing some criterion based on the likelihood function or on the model fitting residuals. The arguments in such a criterion are the number and the locations of structural break points (BPs) as well as the model coefficients of each stationary regime. Recent works include [2] and [11] where the authors have addressed the multiple structural changes problem in a linear regression model using least-squares criterion, and establish the consistency and the rates of convergence of the estimates of the BPs and the regression coefficients. However, the criterion based approach may encounter some practical difficulties when the signal has a large amount of data so that the search space for the optimization of the criterion is huge. This happens for instance with Internet traffic data. Therefore, other methodologies are needed in this case.

Piecewise long-memory (LM) models have been proved to be attractive for instance in telecommunications, see [17] and [16], and in economics and hydrology, see [12]. However, the literature addressing the issue of structural changes in LM models is relatively sparse, partly because these two phenomena are easy to confuse, see [4] and [7]. Due to this difficulty, studies addressing piecewise LM processes consider partial structural change models where only some coefficients are allowed to vary. Then, the BP number is often assumed to be known which simplifies the model fitting method. For

example, [8] has derived the consistency and the rate of convergence of the least-squares BP estimate for a mean change LM process whose BP number is known; [5] and [12] have considered a piecewise LM process with known and constant autoregressive and moving average orders; [6] has estimated the BP locations for a piecewise fractionally integrated process with a known number of breaks.

From a practical point of view, partial structural changes models may be unrealistic. Therefore, a model in which all parameters are allowed to change as well as a corresponding estimation procedure seem to be needed. In [16], a procedure has been proposed for fitting a piecewise fractional autoregressive integrated moving average (FARIMA) process to a non-stationary LM signal. The main idea of this procedure is to estimate the BPs accurately by detecting the changes between the parameter estimates of each regimes. This method is simple and effective, and has the advantage to deal with many thousands of data with occasional BPs, like Internet traffic data. However, this methods has two main limitations. First, it is designed for non-stationary models with at least one BP, and therefore it is not able to discriminate between a stationary and a non-stationary LM signal. Hence, preliminary tests are needed to reject the null hypothesis of stationarity before using this procedure. The second and more serious limitation is that the number of BPs needs to be known. Although an ad-hoc solution to estimate the BP number has been proposed in [16], this solution works only under some restrictions on the BP locations. In this paper, we present a new method which does not suffer from the two limitations above. Our method can be applied to long time series with many thousands of data and BPs satisfying assumptions (A1) and (A2) in Section 3, and we concentrate on simulations results.

The rest of this paper is organized as follows. The piecewise LM model is described in Section 2. In Section 3, our fitting procedure is presented with different BP number selection criteria. Numerical simulation results are discussed in Section 4 and Section 5 concludes the paper.

## 2. MODEL DESCRIPTION

We suppose that the non-stationary process  $\{Y_t\}$ ,  $t = 1, \dots, n$ , can be segmented into  $m+1$  blocks of fractional FARIMA processes. For  $j = 1, \dots, m$ , denote the BP between the  $j$ th and  $(j+1)$ th FARIMA processes as  $\tau_j$ , and set  $\tau_0 = 1$  and  $\tau_{m+1} = n + 1$ . For  $j = 1, \dots, m + 1$ ,

the  $j$ th block of  $\{Y_t\}$  is modeled by

$$Y_t = X_{t+1-\tau_{j-1},j}, \quad \tau_{j-1} \leq t < \tau_j, \quad (1)$$

where  $\{X_{t,j}\}$ ,  $t \in \mathbb{Z}$ , is the FARIMA( $p_j, d_j, q_j$ ) process defined by the difference equation

$$\Phi_j(B)X_{t,j} = \Theta_j(B)(1-B)^{-d_j}\epsilon_{t,j}, \quad (2)$$

$B$  is the backward operator  $BX_t = X_{t-1}$ ,  $\{\epsilon_{t,j}\}$ ,  $t \in \mathbb{Z}$ ,  $j = 1, \dots, m+1$ , is a sequence of iid zero-mean random variables with finite variance,  $d_j \in (0, 1/2)$ , and the polynomials  $\Phi_j(z) = 1 - \phi_{j,1}z - \dots - \phi_{j,p_j}z^{p_j}$  and  $\Theta_j(z) = 1 + \theta_{j,1}z + \dots + \theta_{j,q_j}z^{q_j}$  with real coefficients have no common zeros and neither  $\Phi_j(z)$  nor  $\Theta_j(z)$  has zeros in the closed unit disk  $\{z \in \mathbb{C} : |z| \leq 1\}$ . The process  $(1-B)^{-d_j}\epsilon_{t,j}$  is defined by

$$(1-B)^{-d_j}\epsilon_{t,j} = \sum_{k=0}^{\infty} \varphi_k(d_j)\epsilon_{t-k,j}, \quad (3)$$

where  $\varphi_0(d_j) = 1$  and  $\varphi_k(d_j) = \prod_{s=1}^k \frac{d_j+s-1}{s}$  for  $k \geq 1$ . Since  $d_j < 1/2$ ,  $\sum_{k=0}^{\infty} \varphi_k(d_j)^2 < \infty$  and the series in (3) converges in the mean square sense. Since the sequence  $\{\epsilon_{t,j}\}$ ,  $t \in \mathbb{Z}$ , is zero-mean and iid, the series in (3) converges also almost surely.

Let  $p \geq \max(p_j)$ ,  $q \geq \max(q_j)$ ,  $\alpha_j = (d_j; \phi_{j,1}, \dots, \phi_{j,p}; \theta_{j,1}, \dots, \theta_{j,q})$  where  $\phi_{j,k} = 0$  for  $k > p_j$  and  $\theta_{j,k} = 0$  for  $k > q_j$ . Vector  $\alpha_j$  contains the parameters of the  $j$ th model defined in  $[\tau_{j-1}, \tau_j)$ . The piecewise FARIMA process  $\{Y_t\}$  is characterized by the BPs  $\tau_j$  and the parameters  $\alpha_j$  for  $j = 1, \dots, m+1$ .

### 3. ESTIMATION PROCEDURE

The problem of fitting model (1)–(2) to data consists in finding  $(m, \tau_1, \dots, \tau_m, \alpha_1, \dots, \alpha_{m+1})$ . Following [16], we divide the original time series  $\{Y_t\}$  into a set of elementary sub-series of length  $E$  and use the data in a same sub-series to get a local parameter estimation. Then the differences between the parameter estimates in the elementary intervals can be used to search the BPs which are dispersed into a few intervals.

In the following,  $K$  is the integer part of  $n/E$ , i.e.  $K = \lfloor n/E \rfloor$  and we introduce the elementary intervals  $I_k = ((k-1)E, kE]$  for  $k = 1, \dots, K-1$  and  $I_K = ((K-1)E, n]$ . We make the following assumptions:

(A1) There is no BP neither in  $I_1$  nor in  $I_K$ .

(A2) At least  $(2+\delta)E$  data separate two consecutive BPs for some  $\delta > 0$ .

Our new estimation procedure consists in the following steps.

**Step 1 : Local estimation.** For each interval  $I_k$ ,  $k = 1, \dots, K$ , the model's parameters  $\tilde{\alpha}_k$  are chosen by quasi Gaussian maximum-likelihood estimation (QMLE), see e.g. [3], and a pair  $(\tilde{p}_k, \tilde{q}_k)$  is selected with the Bayes information criterion (BIC) as suggested by [19]. To catch the parameters changes with a comparatively small  $E$ , we choose QMLE since these estimates perform better than the two others popular estimates, namely the wavelet estimates, see e.g. [17], and the Whittle estimates, see e.g. [18], when the data length is not long.

**Step 2 :** Choose  $0 < \eta < \min\{0.5, \delta\}$  and for  $m = 1, \dots, \lfloor \frac{(K-2)E-1}{(2+\eta)E} \rfloor + 1$ , do Steps 2a, b, c.

**Step 2a :** Selection of the intervals with a BP. If model (1)–(2) is suitable for the data, one expects that  $\tilde{\alpha}_k$  is close to the true values of the parameters when there is no BP in the interval  $I_k$ . Now, if there is a BP in  $I_k$  and no BP in  $I_{k-1}$  and  $I_{k+1}$ ,  $\tilde{\alpha}_k$  should be significantly different from both  $\tilde{\alpha}_{k-1}$  and  $\tilde{\alpha}_{k+1}$ . Then, let  $k_0 = 0$ ,  $k_{m+1} = K+1$ , and

$$(\hat{k}_1, \dots, \hat{k}_m) = \operatorname{argmin}_{\{k_1, \dots, k_m\}} \sum_{j=1}^{m+1} \sum_{k=k_{j-1}+1}^{k_j-1} \left( \psi_1(|\tilde{\alpha}_k - \tilde{\alpha}_j|) + \psi_2(|\tilde{p}_k - \tilde{p}_j| + |\tilde{q}_k - \tilde{q}_j|) \right), \quad (4)$$

where the minimum is taken over all possible  $m$ -tuples  $(k_1, \dots, k_m)$  satisfying  $1 < k_1 < \dots < k_m < K$  and assumption (A2) where  $\delta$  is replaced by  $\eta$ , for any vector  $u$  with components  $u_i$ 's,  $|u| = \sum_i |u_i|$ ,  $\tilde{\alpha}_j = \frac{1}{k_j - k_{j-1} - 1} \sum_{k=k_{j-1}+1}^{k_j-1} \tilde{\alpha}_k$ ,  $\tilde{p}_j$  (resp.  $\tilde{q}_j$ ) is the order which is the most frequently selected among the orders  $\tilde{p}_k$  (resp.  $\tilde{q}_k$ ) for  $k = k_{j-1} + 1, \dots, k_j - 1$ . In the case where  $\tilde{p}_j$  (resp.  $\tilde{q}_j$ ) is not unique, the lowest order is chosen. In (4), we take  $\psi_1(x) = \ln(1+x)$  and  $\psi_2(x) = x^{1/2}$ .

When a BP is located close to the upper bound of an elementary interval, minimizing (4) might lead to select this interval or the next one. A similar problem appears when a BP is close to the lower bound of an elementary interval. For this reason, we define the intervals containing a BP as being  $(J_{\hat{k}_1}, \dots, J_{\hat{k}_m})$  where  $J_2 = (E, (2+\eta)E]$ ,  $J_k = ((k-1-\eta)E, (k+\eta)E]$  for  $k = 3, \dots, K-2$  and  $J_{K-1} = ((K-2-\eta)E, (K-1)E]$ . The role of  $\eta > 0$  is to avoid the possible erroneous selection when a BP is close to the limits of an elementary interval.

**Step 2b :** Estimation of the BPs. Suppose that all the intervals  $J_{\hat{k}_j}$  are selected properly, i.e.,  $\tau_j \in J_{\hat{k}_j}$ . Therefore, for any fixed  $j$ , there is no BP in the “previous” block between  $J_{\hat{k}_{j-1}}$  and  $J_{\hat{k}_j}$ , viz.  $((\hat{k}_{j-1} + \eta)E, (\hat{k}_j - 1 - \eta)E]$  where we set  $\hat{k}_0 + \eta = 0$ , and we define  $\hat{\alpha}_p$  as the QMLE of  $\alpha_j$  based on the data in this block where the orders  $(p_p, q_p)$  are selected by BIC. In the same way, let  $\hat{\alpha}_n$  be the QMLE of  $\alpha_{j+1}$  based on the data in the “next” block between  $J_{\hat{k}_j}$  and  $J_{\hat{k}_{j+1}}$ , viz.  $((\hat{k}_j + \eta)E, (\hat{k}_{j+1} - 1 - \eta)E]$  where we set  $(\hat{k}_{m+1} - 1 - \eta)E = n$ , and  $(p_n, q_n)$  be the orders selected by BIC. We treat  $\hat{\alpha}_p$  and  $\hat{\alpha}_n$  as two benchmarks. These estimates are more precise than any local estimate calculated in Step 2 since they involve more data. Suppose that  $l \in J_{\hat{k}_j}$  is the BP  $\tau_j$ . Then we can calculate the QMLE  $\hat{\alpha}_{l_p}$  of  $\alpha_j$  using the orders  $(p_p, q_p)$  and the QMLE  $\hat{\alpha}_{l_n}$  of  $\alpha_{j+1}$  using the order  $(p_n, q_n)$  based respectively on  $((\hat{k}_{j-1} + \eta)E, l]$  and  $(l, (\hat{k}_{j+1} - 1 - \eta)E]$ . These estimates should be close to benchmarks  $\hat{\alpha}_p$  and  $\hat{\alpha}_n$ , respectively. Hence, our choice of the BP estimate  $\hat{\tau}_j$  is based on the following criterion

$$\hat{\tau}_j = \operatorname{argmin}_{l \in J_{\hat{k}_j}} \left( \psi_1(|\hat{\alpha}_{l_p} - \hat{\alpha}_p|) + \psi_1(|\hat{\alpha}_{l_n} - \hat{\alpha}_n|) \right). \quad (5)$$

**Step 2c :** *Estimation of the parameters of each stationary block.* Once  $(\hat{\tau}_1, \dots, \hat{\tau}_m)$  are obtained, the parameters  $\alpha_j$  of the stationary sequence  $X_{t,j}$  for  $j = 1, \dots, m+1$ , can be estimated by QMLE and BIC, on the basis on the data in  $(\hat{\tau}_{j-1}, \hat{\tau}_j]$ , where  $\hat{\tau}_0 = 1$  and  $\hat{\tau}_{m+1} = n$ . We denote by  $\hat{\alpha}_j$  and  $(\hat{p}_j, \hat{q}_j)$  the corresponding parameters and orders.

**Step 3 :**  $m = 0$ . We fit a FARIMA model to the series  $\{Y_t\}$ ,  $t = 1, \dots, n$ , estimating the model's parameters by QMLE and selecting the ARMA orders by BIC.

**Step 4 :** *Selection of the BP number.* For  $m = 0, \dots, \lfloor \frac{(K-2)E-1}{(2+\eta)E} \rfloor + 1$ , we compute the sum of squared residuals of the fitted model (1)–(2) with  $m$  BPs,  $S_n(\hat{\tau}_1, \dots, \hat{\tau}_m)$ , and the Gaussian  $\log_2$ -likelihood  $\log_2 \mathcal{L}_j$  of the  $j$ th segment for  $j = 1, \dots, m+1$ . For  $m = 0$ , we take the model fitted in Step 3 and for  $m > 0$ , we employ the models fitted in Step 2c. These quantities are useful for calculating the following BP number selection criteria. The first three criteria are based on the Schwarz criterion [14] and differ in the severity of their penalty for over-specification. The last criterion is based on the minimum description length (MDL) principle introduced by [13]. Following [20], the selected number of BPs minimizes

$$C_1(m) = \ln [S_n(\hat{\tau}_1, \dots, \hat{\tau}_m)/n] + p^* \frac{\ln n}{n}, \quad (6)$$

where  $p^* = \sum_{i=1}^{m+1} (p_i + q_i) + 2m + 1$  is the total number of parameters. A criterion proposed by [21] takes the form

$$C_2(m) = \ln [S_n(\hat{\tau}_1, \dots, \hat{\tau}_m)/n] + m \frac{C_n}{n}, \quad (7)$$

where  $C_n$  satisfies some constraints, and [9] introduced criterion

$$C_3(m) = \ln [S_n(\hat{\tau}_1, \dots, \hat{\tau}_m)/(n - p^*)] + p^* \frac{c_0 (\ln n)^{2+\gamma_0}}{n}, \quad (8)$$

where  $c_0 > 0$  and  $\gamma_0 > 0$ . Finally, following [15] amounts to minimize

$$C_4(m) = L(m) + \sum_{j=1}^{m+1} \{L(\hat{n}_j) + L(\hat{p}_j) + L(\hat{q}_j) + \frac{\hat{p}_j + \hat{q}_j + 2}{2} \log_2 \hat{n}_j - \log_2 \mathcal{L}_j\}, \quad (9)$$

where  $\hat{n}_j = \hat{\tau}_j - \hat{\tau}_{j-1}$  and for any nonnegative integer  $x$ ,

$$L(x) = \begin{cases} \log_2 c + \log_2 x + \log_2 \log_2 x + \dots & \text{if } x > 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Here  $c$  is a constant approximately equal to 2.865 and the sum involves only the nonnegative terms, whose number is clearly finite. We compare these four criteria in Section 4.

**Remark 1.** Predefining a suitable length  $E$  for the elementary sub-series is not always an easy task: on the one hand, due to LRD, a reasonable number of observations are needed to obtain precise parameter estimates, and then  $E$  can't be too short; on the other, the probability of meeting a BP increases as  $E$  grows. Hence some restriction should be put on  $E$ , and  $E$  is chosen by empirical experience.

**Remark 2.** To reduce the complexity,  $\hat{\alpha}_{l_p}$  and  $\hat{\alpha}_{l_n}$  in Step 2b are calculated using the data in  $(l - E, l)$  and  $(l, l + E)$ , respectively, and this gives good results in practice as shown in Section 4.

## 4. MONTE CARLO SIMULATIONS

Our simulations are based on 1000 replications of two kind of cases where the procedure in [16] does not work. First we consider the case of a stationary process and then we study a case where the ad-hoc solution proposed in [16] to estimate the BP number fails. In our simulations, we take  $E = 2000$ ,  $\eta = 0.1$ , and the maximum value of the ARMA orders considered in BIC is 7. Since there is no general guideline for deciding the values of  $C_n$ ,  $c_0$  and  $\gamma_0$  in the case of piecewise LM series, we take  $C_n = c_1 n^{0.9}$  in (7),  $\gamma_0 = 2$  in (8) and we choose  $c_0$  and  $c_1$  to get the same penalty in (6), (7) and (8) for  $n = 2000$ . This gives  $c_0 = (\ln 2000)^{-3}$ , but since  $m, p_i, q_i$  are unknown, we take arbitrarily  $2 \ln 2000 = c_1 2000^{0.9}$ . These values work properly in the following experiments.

### 4.1 Stationary processes

Here we consider a stationary FARIMA(1,  $d$ , 1) model of length  $n = 40000$ , where  $\{\epsilon_{t,j}\}$ ,  $t \in \mathbb{Z}$ ,  $j = 1, \dots, m+1$ , is a Gaussian sequence,  $\phi = 0.5$ ,  $\theta = -0.7$  and  $d$  varies in  $\{0.15, 0.3, 0.45\}$ .

Table 1 below displays the number of times the different criteria in Step 4 select  $m = 0$  in the 1000 replications. For all  $d$  values and all criteria except  $C_1$ ,  $m = 0$  is properly selected in each experiment. Furthermore, the percentage error of  $C_1$  does not exceed 1%. These excellent performances can be explained as follows. When the series is stationary, the estimated parameters of each stationary block in Step 2c are almost the same and coincide with the parameters obtained in Step 3. Therefore,  $S_n(\hat{\tau}_1, \dots, \hat{\tau}_m)/n$  and  $\sum_{j=1}^{m+1} \log_2 \mathcal{L}_j$  do not vary too much with  $m$ . Hence, the four selection criteria in Step 4 are minimum for the true number of BPs  $m = 0$ .

Criterion	$d$		
	0.15	0.3	0.45
$C_1$	991	987	989
$C_2$	1000	1000	1000
$C_3$	1000	1000	1000
$C_4$	1000	1000	1000

Table 1: Selection of  $m = 0$  for a stationary process.

### 4.2 Multiple BPs processes

Here we consider two piecewise FARIMA models M1 and M2 of length  $n = 40000$  with 4 BPs, where  $\{\epsilon_{t,j}\}$ ,  $t \in \mathbb{Z}$ ,  $j = 1, \dots, m+1$ , is a Gaussian sequence and the parameters of each block are given in Table 2. The four BPs are located at  $\lambda_1 = 6100$ ,  $\lambda_2 = 11000$ ,  $\lambda_3 = 24200$  and  $\lambda_4 = 29900$  for both models. From now on, we use the standardized break fraction  $\lambda_j = \tau_j/n$ , and then  $\lambda_1 = 0.1525$ ,  $\lambda_2 = 0.2750$ ,  $\lambda_3 = 0.6050$  and  $\lambda_4 = 0.7475$ . In model M1, each block is a FARIMA(1,  $d$ , 1) process,

while the orders  $(p_j, q_j)$  of each block are different in model M2. Since (4) and (5) measure the differences between the parameter estimates of the different regimes  $X_{t,j}$  in (2), we test the performance of our method when the parameters of the blocks are very different and when they are close in a piecewise model. The parameters of the first three regimes in model M1 can be considered as very different, while the parameters of the last three regimes are close.

	M1	M2
$\alpha_1$	(0.15; 0.8; -0.5)	(0.45; 0; 0)
$\alpha_2$	(0.4; 0.8; 0.6)	(0.15; -0.8; 0)
$\alpha_3$	(0.2; -0.7; 0.4)	(0.2; 0.7; (-0.7,-0.4))
$\alpha_4$	(0.35; -0.3; 0.5)	(0.10; 0; 0.8)
$\alpha_5$	(0.2; -0.4; 0.8)	(0.35; 0.6; -0.7)

Table 2: Model parameter values.

In Table 3, we present the BP number selection in Step 4 for the two models. For M1, we see that  $C_2$  works better than the other criteria and select the true BP number in more than 95% of the cases. Criteria  $C_3$  and  $C_4$  give the right answer in 82% and 90% of the cases, respectively, while  $C_1$  performs relatively poorly and chooses  $m = 4$  only in half of the cases. No criteria underestimate the BP number. For M2, criteria  $C_2$ ,  $C_3$  and  $C_4$  still work better than  $C_1$ . They select smaller BP numbers than  $C_1$  does, and they find the true number in at least 70% of the cases. Observe that  $C_1$  and  $C_2$  underestimate the BP number in some experiments, which is not the case with  $C_3$  and  $C_4$ .  $C_2$  and  $C_4$  seem to outperform the other two criteria for both models M1 and M2. Observe that  $C_1$  and  $C_4$  given respectively by (6) and (9), are free of any tuning parameter which is not the case of  $C_2$  and  $C_3$  given by (7) and (8), respectively. As a result, taking into consideration simplicity and performance stability,  $C_4$  is more attractive than  $C_2$ .

	$m$	$C_1$	$C_2$	$C_3$	$C_4$
M1	$< 4$	0	0	0	0
	4	544	981	828	907
	$> 4$	456	81	172	93
M2	$< 4$	85	57	0	0
	4	607	786	731	802
	$> 4$	308	157	269	198

Table 3: BP number selection in Step 4.

Table 4 shows the sample mean  $\hat{\mu}(\hat{\lambda}_j)$ , the standard error  $\hat{\sigma}(\hat{\lambda}_j)$  and the mean-squared error (MSE) of the estimation of the BPs in Step 2b when  $m = 4$  for M1 and M2. For both models, the estimated BPs are close to the true BPs, and standard errors as well as MSEs are quite small, especially for model M1. We observe that the estimations of the third and the fourth BP in M1 are not as good as the estimations of the first two BPs since the MSEs are larger. The reason is that the selection of the intervals with a BP in Step 2a encounters more

difficulties when the parameters of the regimes are close.

	$\lambda_j$	0.1525	0.2750	0.6050	0.7475
M1	$\hat{\mu}(\hat{\lambda}_j)$	0.1521	0.2751	0.6073	0.7493
	$\hat{\sigma}(\hat{\lambda}_j)$	0.0019	0.0008	0.0065	0.0055
	MSE	3.82e-6	6.077e-7	4.835e-5	3.353e-5
M2	$\hat{\mu}(\hat{\lambda}_j)$	0.1553	0.2792	0.6059	0.7487
	$\hat{\sigma}(\hat{\lambda}_j)$	0.0222	0.0409	0.0165	0.0161
	MSE	5.103e-4	1.726e-3	2.736e-4	2.619e-4

Table 4: Estimated BPs in Step 2b.

Table 5 gives the number of right model order selection in Step 2c for each stationary regime identified in Step 2b, in the cases where  $m = 4$  is selected. We see that the true orders are well identified for both models. Slightly better results are obtained for M1 since the estimation of the BP locations is more precise in the case of M1.

	1	2	3	4	5
M1	979	951	973	923	912
M2	950	928	960	907	906

Table 5: Right model order selection for each regime in Step 2c.

Table 6 displays the model coefficient estimation results for each regime when  $m = 4$  and the right model orders are selected for all regimes. We see that since the BP locations and the model orders are well recognized, the estimates of the coefficients are quite precise.

## 5. CONCLUSION

In this article, we have proposed a new method to fit piecewise FARIMA models to both stationary and non-stationary LM data. The method can deal with many thousands of data and consists in a four-steps procedure designed to estimate both the BPs and the parameters. Four information criteria are used in the last step of this procedure to determine the BP number. Monte Carlo simulations have demonstrated the good performances of this fitting method.

## REFERENCES

- [1] D. W. K. Andrews. Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4):821–856, 1993.
- [2] J. Bai and P. Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78, 1998.
- [3] J. Beran. *Statistics for long-memory processes*, volume 61 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, New York, 1994.
- [4] R. N. Bhattacharya, V. K. Gupta, and E. Waymire. The Hurst effect under trends. *J. Appl. Probab.*, 20(3):649–662, 1983.

$\hat{\alpha}_i$	M1
$\hat{\mu}(\hat{\alpha}_1)$	(0.144; 0.802; -0.497)
$\hat{\sigma}(\hat{\alpha}_1)$	(0.041; 0.028; 0.018)
MSE	(0.0018; 0.0008; 0.0003)
$\hat{\mu}(\hat{\alpha}_2)$	(0.374; 0.807; 0.580)
$\hat{\sigma}(\hat{\alpha}_2)$	(0.067; 0.031; 0.051)
MSE	(0.0052; 0.0010; 0.0031)
$\hat{\mu}(\hat{\alpha}_3)$	(0.208; -0.699; 0.403)
$\hat{\sigma}(\hat{\alpha}_3)$	(0.031; 0.008; 0.028)
MSE	(0.0010; 0.0001; 0.0008)
$\hat{\mu}(\hat{\alpha}_4)$	(0.344; -0.309; 0.496)
$\hat{\sigma}(\hat{\alpha}_4)$	(0.036; 0.023; 0.045)
MSE	(0.0013; 0.0006; 0.0020)
$\hat{\mu}(\hat{\alpha}_5)$	(0.202; -0.400; 0.799)
$\hat{\sigma}(\hat{\alpha}_5)$	(0.037; 0.014; 0.025)
MSE	(0.0014; 0.0002; 0.0006)
$\hat{\alpha}_i$	M2
$\hat{\mu}(\hat{\alpha}_1)$	(0.4504; 0; 0)
$\hat{\sigma}(\hat{\alpha}_1)$	(0.0106; 0; 0)
MSE	(0.0001; 0; 0)
$\hat{\mu}(\hat{\alpha}_2)$	(0.157; -0.800; 0)
$\hat{\sigma}(\hat{\alpha}_2)$	(0.021; 0.009; 0)
MSE	(0.0005; 0.0001; 0)
$\hat{\mu}(\hat{\alpha}_3)$	(0.199; 0.700; (-0.694, -0.398))
$\hat{\sigma}(\hat{\alpha}_3)$	(0.032; 0.028; (0.0502, 0.012))
MSE	(0.0010; 0.0008; (0.0025, 0.0002))
$\hat{\mu}(\hat{\alpha}_4)$	(0.091; 0; 0.784)
$\hat{\sigma}(\hat{\alpha}_4)$	(0.046; 0; 0.037)
MSE	(0.0021; 0; 0.0016)
$\hat{\mu}(\hat{\alpha}_5)$	(0.359; 0.591; -0.697)
$\hat{\sigma}(\hat{\alpha}_5)$	(0.031; 0.031; 0.010)
MSE	(0.0011; 0.0011; 0.0001)

Table 6: Estimated parameters in Step 2c when  $m = 4$  and the right model orders are selected.

[5] M. Coulon and A. Swami. Least squares detection of multiple changes in fractional ARIMA processes. In *Proceedings of the International Conference on Acoustic, Speech and Signal*, volume 5, pages 3177–3180, 2001.

[6] L. A. Gil-Alana. Fractional integration and structural breaks at unknown periods of time. *J. Time Ser. Anal.*, 29(1):163–185, 2008.

[7] C. M. Kuan and C. C. Hsu. Change-Point Estimation of Fractionally Integrated Processes. *J. Time Ser. Anal.*, 19(6):693–708, 1998.

[8] M. Lavielle and E. Moulines. Least-squares estimation of an unknown number of shifts in a time series. *J. Time Ser. Anal.*, 21(1):33–59, 2000.

[9] J. Liu, S. Wu, and J. V. Zidek. On segmented multivariate regressions. *Statistica Sinica*, 7(2):497–525, 1997.

[10] P. Perron. Dealing with structural breaks. In T. C. Mills and K. Patterson, editors, *Palgrave handbook of econometrics: volume 1: econometric theory*. Palgrave Macmillan, 2006.

[11] P. Perron and Z. Qu. Estimating restricted structural change models. *J. Econometrics*, 134(2):373–399, 2006.

[12] B. K. Ray and R. S. Tsay. Bayesian methods for change-point detection in long-range dependent processes. *J. Time Ser. Anal.*, 23(6):687–705, 2002.

[13] J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978.

[14] G. Schwarz. Estimating the dimension of a model. *Ann. Statist.*, 6(2):461–464, 1978.

[15] L. Song. *Piecewise models for long memory time series*. PhD thesis, Université Paris XI, September 2010.

[16] L. Song and P. Bondon. A local stationary long-memory model for internet traffic. In *Proceedings of the European Signal Processing Conference*, pages 1067–1071, Glasgow, Scotland, 2009.

[17] S. Stoev, M. S. Taqqu, C. Park, G. Michailidis, and J. S. Marron. LASS: a tool for the local analysis of self-similarity. *Comput. Statist. Data Anal.*, 50(9):2447–2471, 2006.

[18] M. S. Taqqu and V. Teverovsky. Robustness of Whittle-type estimators for time series with long-range dependence. *Comm. Statist. Stochastic Models*, 13(4):723–757, 1997.

[19] K. Torre, D. Delignières, and L. Lemoine. Detection of long-range dependence and estimation of fractal exponents through ARFIMA modelling. *British Journal of Mathematical and Statistical Psychology*, 60(1):85–106, 2007.

[20] Y.-C. Yao. Estimating the number of change-points via Schwarz’ criterion. *Statistics & Probability Letters*, 6(3):181–189, 1988.

[21] Y.-C. Yao and S. T. Au. Least-squares estimation of a step function. *Sankhyā Ser. A*, 51(3):370–381, 1989.