

**Prediction of Long Memory Time
Series Models: A Simulation Study
and an Application**

V. Reisen and B. Abraham
Univeristy of Waterloo

RR-98-04
April 1998

Prediction of long memory time series models: A Simulation Study and an Application.

Valdério Anselmo Reisen¹
Departamento de Estatística, CCE
Universidade Federal do Espírito Santo
Brazil

Bovas Abraham
Institute for Improvement in
Quality and Productivity
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

Abstract:

In this paper we consider forecasts from long memory time series using the ARFIMA(p, d, q) model with $d \in (0.0, 0.5)$. We also investigate through simulations, the bias in the estimate of the variance of the k -step ahead forecast errors. The ARFIMA model is also used to analyse and forecast a set of wind speed data and these forecasts are compared with those from an ARIMA model.

Key words: long memory, fractional differencing, forecasting, smoothed and periodogram regressions.

1. Introduction

An autoregressive integrated moving average (ARIMA(p, d, q)) process in which the parameter d (the degree of differencing) takes non-integer real values is sometimes called an ARFIMA(p, d, q) process. The ARFIMA model has recently become a useful tool in the analyses of time series in different fields such as, astronomy, hydrology, mathematics, and computer science. It

¹Work was done while visiting the University of Waterloo, Canada.

can characterize “long-range” dependence when $d \in (0.0, 0.5)$, and “short-range” dependence when $d \in (-0.5, 0.0)$. A good review of long memory process may be found in Beran (1994).

The long-range dependence or “persistence” as defined by some authors is characterized by the presence of significant dependence between observations separated by a long time interval. The characteristics of long memory or short memory properties of ARFIMA(p, d, q) processes when $d \in (-0.5, 0.5)$ can be seen in the shapes of the spectral density and autocorrelation functions. For $d \in (0.0, 0.5)$, the process is long memory because $\sum_j |\rho_j|$ diverges, ρ_j being the autocorrelation function of the process, and as the frequency w goes to 0 the spectral density becomes unbounded. Thus the type of dependence between observations is determined essentially by the fractional parameter d . For some aspects of the estimation of d see Geweke and Porter-Hudak (1983), Hassler (1993), Reisen (1993, 1994) and Chen, et al. (1994).

This paper is concerned with the problem of forecasting a time series with possible long memory features. The outline of this paper is as follows: in Section 2, we summarize some results related to the ARFIMA(p, d, q) model and the estimation of the fractional differencing parameter d . Section 3 describes some forecasting issues in long memory models and use simulated results to evaluate the bias in the estimate of the variance of the k -step ahead forecast are presented. In section 4, long memory and short memory models are used to analyse a set of wind speed data. Some concluding remarks are given in section 5.

2. The ARFIMA(p,d,q) model

Hosking (1981) and Reisen (1994) describe ARFIMA models in detail. Here we summarize some results.

Let $\{\epsilon_t\}$ be a white noise process with $E(\epsilon_t) = 0$, $V(\epsilon_t) = \sigma_\epsilon^2$ and denote the back-shift operator, B , such that $BX_t = X_{t-1}$. Let $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ and $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ be polynomials of orders p , and q respectively with roots outside of the unit circle. If $\{X_t\}$ is a linear process satisfying

$$\Phi(B)(1 - B)^d X_t = \Theta(B)\epsilon_t, \quad d \in (-0.5, 0.5), \quad (2.1)$$

then $\{X_t\}$ is called an ARFIMA(p, d, q) process where d is the degree of differencing.

The process defined in (2.1) is stationary and invertible, and its spectral density, $f(w)$, is given by

$$f(w) = f_u(w)(2\sin(w/2))^{-2d}, \quad w \in [-\pi, \pi] \quad (2.2)$$

where the function $f_u(w)$ is the spectral density of an ARMA(p, q) process. The process in (2.1) can be written in the form

$$X_t = (1 - B)^{-d} \frac{\Theta(B)}{\Phi(B)} \epsilon_t = \Psi(B)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j},$$

and Hassler (1993) showed that

$$\psi_j \cong j^{d-1} \frac{b}{(d-1)!} \quad \text{as } j \rightarrow \infty, \quad (2.3)$$

where b is a constant.

2.1. Estimates of d

Consider the set of harmonic frequencies $w_j = \frac{2\pi j}{n}$, $j = 0, 1, \dots, [n/2]$ where n is the sample size. Let $\{X_t\}$ be an ARFIMA(p, d, q) process with $d \in (-0.5, 0.5)$. The logarithm of the spectral density may be written as

$$\ln f(w_j) = \ln f_u(0) - d \ln (2\sin(w_j/2))^2 + \ln\{f_u(w_j)/f_u(0)\}. \quad (2.4)$$

We now consider two estimators for d which are obtained through the regression equations constructed from equation (2.4). The first estimate, denoted by \hat{d}_p , uses the periodogram function, and the second, \hat{d}_{sp} , uses a smoothed periodogram function. Geweke and Porter-Hudak (1983) showed that \hat{d}_p is asymptotically normally distributed with $E(\hat{d}_p) = d$ and $Var(\hat{d}_p) = \frac{\pi^2}{6 \sum_{i=1}^{g(n)} (X_i - \bar{X})^2}$, where $g(n)$ is a function of n and $x_j = \ln(2\sin(w_j/2))^2$.

The estimator \hat{d}_{sp} is obtained by replacing the spectral function in equation (2.4) by the smoothed periodogram function with the Parzen lag window. Reisen(1994) showed that \hat{d}_{sp} is asymptotically normally distributed with $E(\hat{d}_{sp}) = d$ and $Var(\hat{d}_{sp}) \approx 0.539285 \frac{m}{n \sum_{i=1}^{g(n)} (X_i - \bar{X})^2}$, where m is a function of n and usually referred to as the truncation point in the Parzen lag window ($m = n^\beta, 0 < \beta < 1$). Since the autocorrelation function of the ARFIMA(p, d, q) process is not summable for d in $(0.0, 0.5)$, the theoretical results relating to both the estimates through regression hold only in the case where d is negative. However, simulations have shown that these estimators can also be applied in the case $d > 0$ (see, for instance, Reisen (1994)).

3. Forecasting the ARFIMA process

Since the ARFIMA process in (2.1) is invertible for $d > -.5$ forecasts from this process can be obtained in the same way as those from an ARIMA process. Similarly, since the ARFIMA process is stationary for $d < .5$, the variance of the k -step ahead forecast error can be obtained using the same procedure as in the ARIMA case. The ARFIMA(p, d, q) process may be written in either infinite *AR* or *MA* representations which at time $t + k$ are:

$$AR: X_{t+k} = \sum_{j=1}^{\infty} \pi_j X_{t+k-j} + \epsilon_{t+k}, \quad (3.3)$$

$$MA: X_{t+k} = \sum_{j=0}^{\infty} \psi_j \epsilon_{t+k-j} \quad (3.4)$$

where ψ_j and π_j are the coefficients of B^j in the expansions of $\Psi(B) = \frac{\Theta(B)}{\Phi(B)}(1 - B)^{-d}$ and $\Pi(B) = \frac{\Phi(B)}{\Theta(B)}(1 - B)^d$ respectively.

Hence given X_n, X_{n-1}, \dots we have

$$\hat{X}_n(k) = E(X_{n+1}|X_n, X_{n-1}, \dots) = \sum_{j=1}^{\infty} \pi_j \hat{X}_n(k - j), \quad \text{for } k \geq 1. \quad (3.5)$$

Note that for $j \geq 1$,

$$\hat{\epsilon}_n(j) = E(\epsilon_{n+j}|X_n, X_{n-1}, \dots) = 0, \quad \text{and } \hat{X}_n(j) = E(X_{n+j}|X_n, X_{n-1}, \dots),$$

and for $j \leq 0$, $\hat{\epsilon}_n(j) = \epsilon_{n+j}$ and $\hat{X}_n(j) = X_{n+j}$.

Now we may obtain the forecast error $e_n(k)$ using the infinite *MA* representation given by (3.4) as follows:

$$\hat{X}_n(k) = E \left(\sum_{j=0}^{\infty} \psi_j \epsilon_{n+k-j} | X_n, X_{n-1}, \dots \right) = \sum_{j=k}^{\infty} \psi_j \epsilon_{n+k-j} \quad (3.6)$$

The forecast error is

$$e_n(k) = X_{n+k} - \hat{X}_n(k) = \sum_{j=0}^{k-1} \psi_j \epsilon_{n+k-j}. \quad (3.7)$$

By (2.3) $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, and hence $Var(e_n(k)) = \sigma_\epsilon^2 \sum_{j=0}^{k-1} \psi_j^2$ is finite for all k .

Simulation study

To evaluate the forecasting properties of the ARFIMA(p,d,q) process, we simulate data from ARFIMA(p, d, q) models, with $0 < d < .5$, and $p, q = 0, 1$, using the algorithm in Hosking (1984). The MA and AR weights for the above models are needed for generation of forecasts and forecast errors, and these are as follows:

MA :

$$\psi_0 = 1.0, \quad \psi_1 = -\theta - A_1$$

$$\psi_j = \sum_{m=0}^{j-1} \psi_m A_{k-m}, \text{ for } j \geq 2, \quad \text{where } A_j = \frac{1}{\Gamma(-d)} \left[\frac{\Gamma(j-d)}{\Gamma(j+1)} - \phi \frac{\Gamma(j-d-1)}{\Gamma(j)} \right]$$

and $\Gamma(\cdot)$ is the Gamma function.

AR :

$$\pi_0 = 1, \quad \pi_j = \theta \pi_{j-1} + A_j, \quad j \geq 1$$

These weights can be calculated once d, θ and ϕ are known.

For each model with given (d, ϕ, θ) we generate 350 observations with mean zero and $\sigma_\epsilon^2 = 1$. Then we discard the first 50 observations to avoid transient initial effects and the remaining 300 observations are labelled as X_1, X_2, \dots, X_{300} . From this series the parameter estimates $\hat{\phi}, \hat{\theta}, \hat{\sigma}_\epsilon^2$ are calculated using NAG subroutines and the forecast $\hat{X}_{260+j}(k)$ for $X_{260+j+k}$ is obtained for $j = 0, 1, 2, \dots, 40 - k$ and $k = 1, 2, \dots, 7$. Thus we have the one

step ahead forecasts and corresponding forecast errors:

$$\hat{X}_{260}(1), \hat{X}_{261}(1), \dots, \hat{X}_{299}(1)$$

$$e_{260}(1), e_{261}(1), \dots, e_{299}(1)$$

the two step ahead forecasts and the corresponding forecast errors

$$\hat{X}_{260}(2), \hat{X}_{261}(2), \dots, \hat{X}_{298}(2)$$

$$e_{260}(2), e_{261}(2), \dots, e_{298}(2)$$

and up to seven steps ahead.

It should be noted that the forecasts are obtained by the truncation

$$\hat{X}_n(k) \simeq \sum_{j=1}^{n+k-1} \hat{\pi}_j \hat{X}_n(k-j) \quad (3.8)$$

Then we obtain the sample mean $\bar{e}(k) = \sum_{j=0}^{40-k} e_{260+j}(k)/(41-k)$, the sample variance,

$$s^2(k) = \sum_{j=0}^{40-k} (e_{260+j}(k) - \bar{e}(k))^2/(40-k),$$

the true variance

$$V(e_n(k)) = \sigma_\epsilon^2 \left[\sum_{j=0}^{k-1} \psi_j^2 \right]$$

and

$$\hat{V}(e_n(k)) = \hat{\sigma}_\epsilon^2 \sum_{j=0}^{k-1} \hat{\psi}_j^2 \quad k = 1, 2, \dots, 7 \quad (3.9)$$

when $\hat{\psi}_j$'s and $\hat{\sigma}_\epsilon^2$ are the estimates from the series. This process is repeated 200 times and averages over the repetitions are calculated. Then we have

$$\bar{e}(k) = \sum_{i=1}^{200} \bar{e}_i(k)/200, \quad \bar{s}^2(k) = \sum_{i=1}^{200} s_i^2(k)/200 \quad (3.10)$$

and

$$\bar{V}(e_n(k)) = \sum (\hat{V}_i(e_n(k)))/200$$

when $\bar{e}_i(k)$, $s_i^2(k)$ and $\hat{V}_i(e_n(k))$ refer to the quantities from the i th simulation ($i = 1, 2, \dots, 200$). Then we define two measures of bias for the estimated variance of the k -step forecast error.

$$b_{1k} = V(e_n(k)) - \bar{s}^2(k)$$

$$b_{2k} = V(e_n(k)) - \bar{V}(e_n(k)) \quad k = 1, 2, \dots, 7$$

We also consider the following sample variance from the 200 series:

$$s_*^2(k) = \sum_{i=1}^{200} [e_{i,260}(k) - \bar{e}_{260}(k)]^2 / 199, \quad k = 1, 2, \dots, 7$$

where $\bar{e}_{260}(k) = \sum_{i=1}^{200} e_{i,260}(k) / 200$ and $e_{i,260}(k)$ refer to the k step ahead forecast error from simulation i ($i = 1, 2, \dots, 200$).

Table 1 gives values for $\bar{e}(k)$, b_{1k} , b_{2k} for $d = .2, .3, .4$, and for some values of ϕ . To save space we show only the results for $k = 1, 3, 5, 7$. The quantities in parenthesis are the standard deviations over the 200 repetitions. From this Table we conclude:

1. The average value of $\hat{\phi}$ is very close to the true value of ϕ for all the values of d and ϕ considered.
2. $\bar{e}(k)$ is very close to zero in all cases. As expected the sample standard deviation (over the 200 repetitions shown in parenthesis) increases with k .
3. For given k , b_{1k} and b_{2k} increase with d and these increases are substantial when ϕ is close to 1 and also when d is close to .5. However, when ϕ is close to -1 the increase in b_{ik} ($i = 1, 2$) is not as large as when ϕ is positive.

Table 1. Simulation results for $\bar{e}(k)$, b_{1k} and b_{2k}

d	ϕ	$\hat{\phi}$	$k = 1$			$k = 3$			$k = 5$			$k = 7$		
			\bar{e}	$b_1 = b_2$		\bar{e}	b_1	b_2	\bar{e}	b_1	b_2	\bar{e}	b_1	b_2
0.2	-0.9	-0.895 (0.026)	0.011 (0.135)	0.099 (0.21)	0.013 (0.142)	0.181 (0.64)	0.215 (0.442)	0.018 (0.149)	0.272 (1.02)	0.296 (0.623)	0.019 (0.155)	0.339 (1.359)	0.346 (0.767)	
	-0.5	-0.494 (0.055)	-0.004 (0.150)	0.108 (0.217)	-0.001 (0.149)	0.125 (0.285)	0.126 (0.259)	0.0 (0.16)	0.128 (0.295)	0.125 (0.264)	0.005 (0.172)	0.131 (0.301)	0.125 (0.264)	
	-0.2	-0.196 (0.059)	-0.020 (0.153)	0.107 (0.200)	-0.020 (0.173)	0.110 (0.218)	0.105 (0.203)	-0.022 (0.197)	0.119 (0.221)	0.106 (0.205)	-0.021 (0.213)	0.124 (0.226)	0.106 (0.206)	
	0.0	-	(0.013)	(0.103)	0.021	0.127	0.109	0.023	0.15	0.110	0.026	0.161	0.111	
			(0.156)	(0.202)	(0.209)	(0.221)	(0.213)	(0.241)	(0.227)	(0.216)	(0.264)	(0.233)	(0.217)	
	0.2	0.200 (0.054)	-0.005 (0.155)	0.132 (0.185)	-0.005 (0.248)	0.180 (0.257)	0.154 (0.220)	-0.006 (0.289)	0.200 (0.266)	0.156 (0.225)	-0.008 (0.316)	0.223 (0.278)	0.157 (0.227)	
	0.5	0.496 (0.055)	-0.008 (0.154)	0.137 (0.195)	-0.019 (0.339)	0.317 (0.440)	0.235 (0.349)	-0.019 (0.434)	0.429 (0.533)	0.251 (0.397)	-0.016 (0.493)	0.508 (0.554)	0.254 (0.414)	
	0.9	0.899 (0.028)	-0.007 (0.141)	0.115 (0.198)	-0.014 (0.456)	0.526 (0.937)	0.400 (0.689)	-0.014 (0.775)	1.254 (1.939)	0.652 (1.204)	-0.021 (1.080)	2.15 (2.810)	0.838 (1.710)	
	-0.9	-0.895 (0.025)	0.002 (0.129)	0.270 (0.161)	0.006 (0.149)	0.489 (0.413)	0.524 (0.310)	0.004 (0.161)	0.662 (0.712)	0.698 (0.426)	0.005 (0.167)	0.770 (0.960)	0.830 (0.520)	
	-0.2	-0.200 (0.056)	0.019 (0.148)	0.253 (0.168)	0.023 (0.191)	0.284 (0.186)	0.260 (0.177)	0.027 (0.230)	0.303 (0.189)	0.266 (0.181)	0.028 (0.257)	0.316 (0.195)	0.269 (0.183)	
0.3	0.0	-	(0.026)	(0.235)	0.0198	0.291	0.266	0.015	0.32	0.274	0.016	0.341	0.279	
			(0.141)	(0.159)	(0.215)	(0.209)	(0.179)	(0.262)	(0.227)	(0.185)	(0.298)	(0.239)	(0.189)	
	0.2	0.202 (0.058)	-0.004 (0.136)	0.267 (0.160)	-0.01 (0.250)	0.404 (0.252)	0.35 (0.220)	-0.012 (0.310)	0.455 (0.295)	0.366 (0.235)	-0.018 (0.348)	0.493 (0.323)	0.375 (0.242)	
	0.5	0.500 (0.051)	-0.016 (0.138)	0.274 (0.175)	-0.036 (0.336)	0.652 (0.428)	0.531 (0.388)	-0.044 (0.460)	0.840 (0.570)	0.600 (0.480)	-0.047 (0.556)	0.958 (0.631)	0.629 (0.526)	
	0.9	0.900 (0.028)	0.007 (0.144)	0.254 (0.170)	0.024 (0.502)	1.183 (0.887)	1.028 (0.700)	0.029 (0.886)	2.440 (2.010)	1.840 (1.350)	0.025 (1.261)	3.910 (3.220)	2.570 (2.060)	

Table 1 (cont'd)

d	ϕ	$\hat{\phi}$	$k = 1$			$k = 3$			$k = 5$			$k = 7$		
			\bar{e}	$b_1 = b_2$	\bar{e}	b_1	b_2	\bar{e}	b_1	b_2	\bar{e}	b_1	b_2	\bar{e}
0.4	-0.9	-0.895 (0.028)	-0.014 (0.115)	0.531 (0.109)	-0.017 (0.143)	0.929 (0.269)	0.950 (0.199)	-0.023 (0.164)	1.202 (0.455)	1.220 (0.269)	-0.026 (0.177)	1.388 (0.610)	1.410 (0.324)	
	-0.2	-0.194 (0.053)	0.016 (0.120)	0.540 (0.100)	0.021 (0.177)	0.605 (0.118)	0.589 (0.113)	0.028 (0.221)	0.649 (0.133)	0.619 (0.119)	0.031 (0.260)	0.684 (0.142)	0.636 (0.123)	
	0.0	-	-0.003 (0.110)	0.530 (0.110)	-0.005 (0.187)	0.673 (0.155)	0.656 (0.139)	-0.004 (0.236)	0.745 (0.176)	0.702 (0.148)	-0.006 (0.274)	0.790 (0.196)	0.729 (0.154)	
0.9	0.2	0.199 (0.055)	0.007 (0.110)	0.536 (0.095)	0.014 (0.229)	0.846 (0.195)	0.814 (0.147)	0.021 (0.298)	0.953 (0.263)	0.895 (0.165)	0.027 (0.352)	1.026 (0.298)	0.940 (0.174)	
	0.5	0.497 (0.057)	-0.005 (0.104)	0.521 (0.107)	-0.014 (0.276)	1.29 (0.345)	1.220 (0.270)	-0.017 (0.395)	1.710 (0.494)	1.510 (0.365)	-0.027 (0.480)	1.960 (0.570)	1.660 (0.418)	
	0.9	0.90 (0.028)	0.002 (0.102)	0.536 (0.096)	0.005 (0.396)	2.650 (0.650)	2.560 (0.470)	-0.001 (0.735)	5.600 (1.706)	5.090 (1.010)	-0.012 (1.090)	8.940 (3.120)	7.607 (1.677)	

Table 2. Simulation Results for $\bar{e}(k)$, b_{1k} , b_{2k}

d	θ	$\hat{\theta}$	$k = 1$			$k = 3$			$k = 5$			$k = 7$		
			\bar{e}	$b_1 = b_2$		\bar{e}	b_1	b_2	\bar{e}	b_1	b_2	\bar{e}	b_1	b_2
0.2	0.2	0.200 (0.054)	0.003 (0.147)	0.099 (0.210)	0.001 (0.161)	0.100 (0.220)	0.097 (0.212)	-0.002 (0.187)	0.111 (0.223)	0.098 (0.214)	0.000 (0.200)	0.124 (0.232)	0.098 (0.215)	
	0.5	0.500 (0.050)	-0.015 (0.156)	0.093 (0.200)	-0.011 (0.120)	0.090 (0.243)	0.095 (0.226)	-0.008 (0.132)	0.091 (0.246)	0.095 (0.226)	-0.007 (0.134)	0.089 (0.258)	0.095 (0.226)	
	0.9	0.900 (0.028)	-0.008 (0.145)	0.097 (0.201)	-0.005 (0.048)	0.105 (0.39)	0.134 (0.306)	-0.003 (0.044)	0.104 (0.394)	0.139 (0.306)	0.000 (0.048)	0.100 (0.410)	0.134 (0.306)	
0.4	0.2	0.200 (0.057)	-0.006 (0.102)	0.529 (0.110)	0.01 (0.195)	0.578 (0.122)	0.568 (0.121)	0.010 (0.180)	0.610 (0.139)	0.590 (0.110)	0.010 (0.220)	0.640 (0.142)	0.610 (0.131)	
	0.5	0.495 (0.056)	0.004 (0.110)	0.510 (0.107)	0.003 (0.120)	0.520 (0.120)	0.520 (0.110)	0.005 (0.138)	0.530 (0.120)	0.503 (0.110)	0.004 (0.160)	0.540 (0.120)	0.530 (0.110)	
	0.9	0.900 (0.028)	0.001 (0.113)	0.537 (0.103)	0.002 (0.054)	0.660 (0.151)	0.670 (0.132)	0.001 (0.052)	0.660 (0.150)	0.670 (0.132)	0.001 (0.052)	0.670 (0.154)	0.670 (0.132)	

4. We note that $b_{2k} < b_{1k}$ (i.e. $\hat{V}(e_n(k)) < \bar{s}^2(k)$) and the average difference is rather small for most cases. However, the standard deviation of b_{2k} is smaller than that of b_{1k} . It should be noted that both estimators $\hat{V}(e_n(k))$ and $\bar{s}^2(k)$ underestimate $V(e_n(k))$.

Table 2 presents results similar to those in Table 1 for $\theta = .2, .5, .9$. The behaviour of $\hat{\theta}$, $\bar{e}(k)$, b_{1k} and b_{2k} are similar to those in Table 1. However b_{1k} and b_{2k} are much smaller than those in Table 1 (with $\phi \neq 0$), especially when $\phi > 0$.

In our simulation study we have also considered many other values of ϕ and θ as well as the ARFIMA(1, d , 1) process. The results are similar to those presented and hence to save space we did not include those here.

4. An Example: The Wind Speed Data

We consider a set of wind speed data (wsd). Wind speeds were collected every five minutes from 00:00 to 23:55 hours at the SILSOE Research Institute on 17/05/91 and are in units of miles per second (m/s). This series, the sample autocorrelation function (acf) and the partial acf are shown in Figures 1-3. The series appears to be stationary and the sample acf indicate that a long memory (ARFIMA) model could be considered. We adopt the following four step strategy to find an adequate ARFIMA(p, d, q) model for this data.

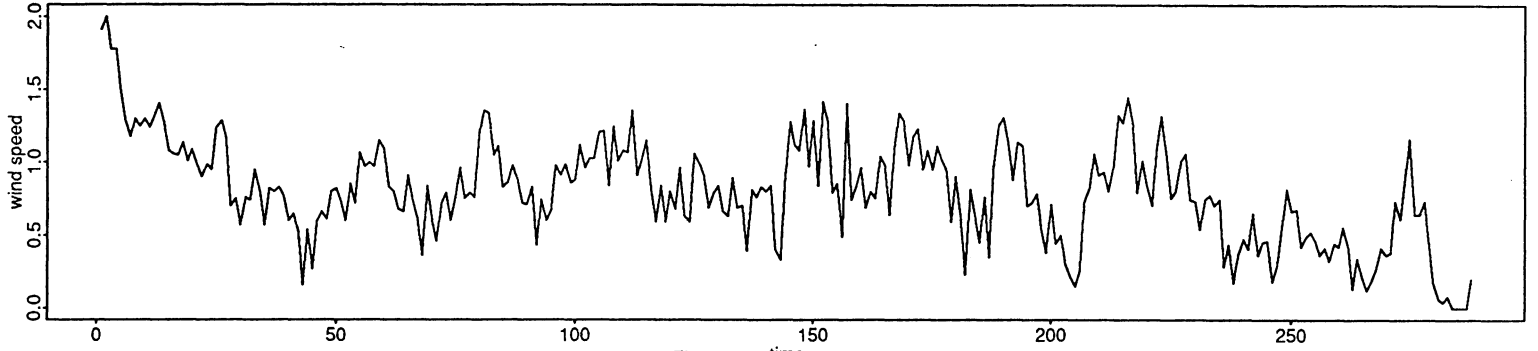


Figure 1. Wind Speed Data
Series : wsd

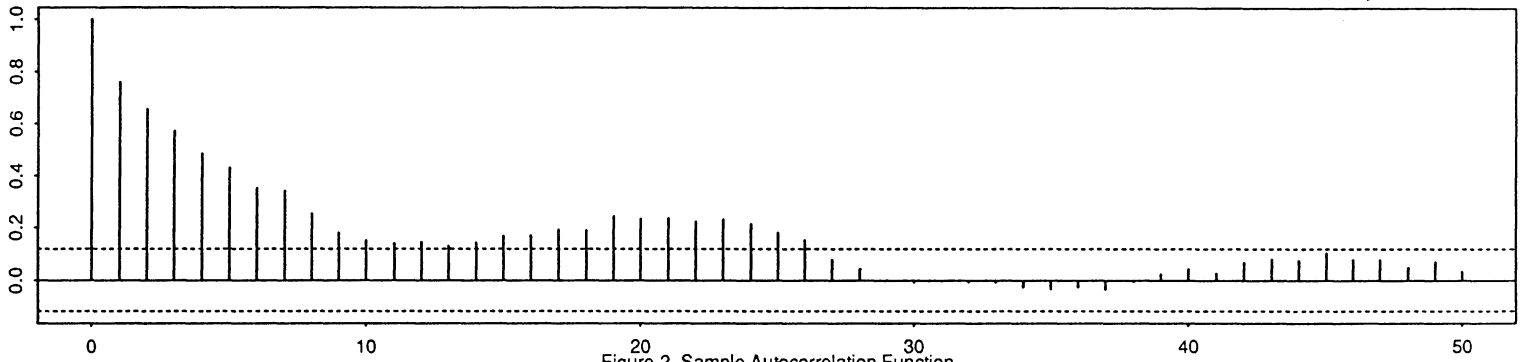


Figure 2. Sample Autocorrelation Function
Series : wsd

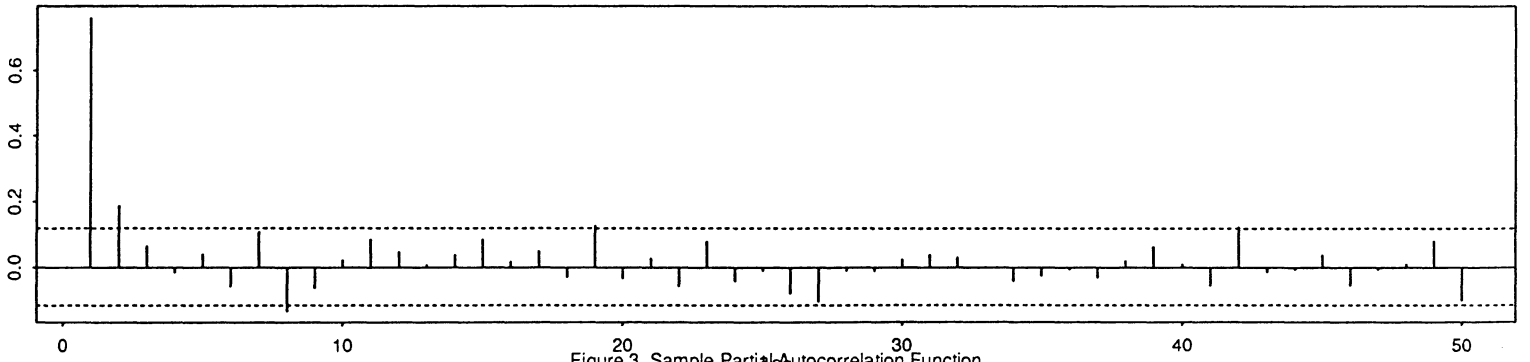


Figure 3. Sample Partial Autocorrelation Function
Series : wsd

4.1 Model building:

Step 1. Estimation of d .

As indicated in Section 2, d can be estimated in different ways. Periodogram regression gives $\hat{d}_p = 0.289$ with estimated variance $S_p^2 = .044$ obtained through regression, and the estimated asymptotic variance of d_p , $Var(\hat{d}_p) = 0.038$. The smoothed periodogram regression (with the truncation point $n^{0.5}$) leads to $\hat{d}_{sp} = 0.299$ with estimated regression variance, $S_{sp}^2 = 0.008$ and $Var(\hat{d}_{sp}) = 0.004$. The number of regression observations used is $g(n) = n^{0.5} = 16$.

Although the point estimates of d look close to each other, i.e. $\hat{d}_p \cong \hat{d}_{sp}$, the estimated variances are quite different.

Step 2. Test $H_0 : d = 0$. We consider the statistic $z = \hat{d}_i / \hat{\sigma}_i$, $i = s, sp$ for this purpose where $\hat{\sigma}_i^2 = Var(\hat{d}_i)$.

In this example $z = 3.72$ for the smoothed periodogram case and $z = 1.37$ for the periodogram estimate. The approximate 95% confidence interval for d , $(\hat{d} \pm 1.96\hat{\sigma}_{\hat{d}})$, are respectively $(0.175, 0.423)$ and $(-.09, 0.67)$. The smoothed periodogram method indicates that there is evidence against $d = 0$ while the periodogram method shows the opposite. It is shown in Reisen (1994) that for testing $H_0 : d = 0$ ($0 < d < 0.5$) the smoothed regression estimator is more powerful than the periodogram estimator. In any case we proceed with $d = 0.299$.

Step 3. Identification and estimation of ARFIMA.

We now look for an ARMA(p, q) model for $\hat{u}_t = (1 - B)^{0.299}(X_t - 0.8)$. The AIC criterion (Akaike (1973)) leads to ARFIMA(1, \hat{d} , 1) model and the estimated model is $(1 - 0.81B)(1 - B)^{0.299}X_t^* = (1 - 0.55B)\epsilon_t$ with $\hat{\sigma}_\epsilon^2 = 0.044$

and $X_t^* = X_t - 0.8$. The AIC (estimated value) is -73.37 .

As discussed in Crato and Ray (1996), the AIC criterion generally tends to underestimate the orders for long range dependent processes. Hence, we also looked for other possible criterion such as BIC to choose p and q . These all lead to the same model.

Step 4: Diagnostic checking

Several residual checks including the autocorrelation function, and the normal probability plots indicated that the model above is adequate for the data.

It was noted earlier that the periodogram estimator had suggested $d = 0$, i.e., the data may be modelled by a short memory model. Based on the AIC criterion we arrived at an ARMA (1,1) model.

$(1 - 0.89B)X_t^* = (1 - 0.35B)\epsilon_t$, $\hat{\sigma}_\epsilon^2 = 0.04325$, $AIC = -82.83$. The diagnostic checks for this model also indicated that the fitted model is adequate.

4.2 Forecast:

For comparison we generate the minimum mean square forecasts from the ARFIMA and ARMA models. Thus we re-estimate the models using the first 260 observations. This lead to

$$\text{ARFIMA: } (1 - 0.79B)(1 - B)^{0.3}X_t^* = (1 - 0.58B)\epsilon_t, \quad \hat{\sigma}_\epsilon^2 = 0.0443,$$

$$AIC = -64.14$$

$$\text{ARMA: } (1 - 0.86B)X_t^* = (1 - 0.37)\epsilon_t, \quad \hat{\sigma}_\epsilon^2 = 0.044, \quad AIC = -75.43$$

The estimate of d , ϕ and θ are similar to those obtained before. The estimate of d by periodogram function was 0.043 which is somewhat smaller than

Table 3. Wind Speed Data

mean values ($\bar{e}(k)$), sample variances ($s^2(k)$), estimated variances $\hat{V}(e_n(k))$,
 mean square error ($mse(k)$), $k = 1, 2, \dots, 15$.

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
ARMA	$\bar{e}(k)$	-0.096	-0.156	-0.204	-0.236	-0.269	-0.292	-0.306	-0.318	-0.329	-0.341	-0.347	-0.354	-0.381	-0.399	-0.442
	$s^2(k)$	0.035	0.056	0.075	0.101	0.123	0.139	0.149	0.154	0.157	0.154	0.153	0.153	0.141	0.137	0.122
	$\hat{V}(e_n(k))$	0.035	0.045	0.053	0.059	0.064	0.067	0.070	0.072	0.074	0.075	0.076	0.077	0.078	0.080	0.078
	$mse(k)$	0.045	0.081	0.116	0.156	0.195	0.224	0.242	0.255	0.265	0.271	0.274	0.278	0.287	0.298	0.318
ARFIMA	$\bar{e}(k)$	-0.069	-0.12	-0.15	-0.176	-0.201	-0.218	-0.228	-0.236	-0.243	-0.252	-0.254	-0.257	-0.281	-0.295	-0.330
	$s^2(k)$	0.036	0.058	0.077	0.105	0.129	0.147	0.159	0.165	0.168	0.164	0.159	0.156	0.139	0.131	0.114
	$\hat{V}(e_n(k))$	0.036	0.047	0.055	0.062	0.067	0.072	0.075	0.078	0.083	0.084	0.085	0.087	0.088	0.089	0.091
	$mse(k)$	0.040	0.072	0.10	0.136	0.170	0.195	0.211	0.220	0.227	0.227	0.224	0.222	0.218	0.217	0.226
$\gamma(k) - 100$	12.5%	12.5%	16%	14.7%	14.7%	14.9%	15.2%	15.9%	16.7%	19.38%	22.3%	25.2%	31.6%	37.3%	40.7%	

that obtained before. However, this value is still inside the 95% confidence interval obtained for d by the periodogram estimator.

We generate the k -step ahead forecasts, and the corresponding forecast errors using the two models, for the holdout period $t = 261, \dots, 288$ and for $k = 1, 2, \dots, 15$. Then, as in the previous simulations in Section 3 we obtained the sample average error ($\bar{e}(k)$), the sample variance ($s^2(k)$), the estimated variance $\hat{V}(e_n(k)) = \hat{\sigma}_\epsilon^2 \left[\sum_{j=0}^{k-1} \hat{\psi}_j^2 \right]$ and the mean square error of the forecast, $mse(k) = \sum_{j=0}^{28-k} e_{260+j}^2(k)/(29-k)$.

The results are shown in Table 3. We note the following

1. For all k $|\bar{e}(k)|$ is larger for the model.
2. $s^2(k) > \hat{V}(e_t(k))$ in both models.
3. To compare the forecasts from the two models we compute the ratio of the mean squares from both models:

$$\text{i.e. } \gamma(k) = \frac{\text{arma } mse(k)}{\text{arfima } mse(k)} \times 100.$$

The last row of Table 3 shows the % increase in $mse(k)$ if ARIMA is used instead of ARFIMA (i.e. $\gamma(k) - 100$). We note that this value is bigger than 12.5% and that it increases with k .

4. We also obtained

$$MSE = \sum_{k=1}^{20} e_{260}^2(k)/20$$

and

$$MPE = 100 \sum_{k=1}^{20} \left| \frac{e_{260}(k)}{X_{260+k}} \right| / 20$$

for the two models.

	MSE	MPE
ARFIMA	.084	70.28
ARMA	.113	75.64

From the analysis of the forecasts we see that the ARFIMA model has an edge on ARIMA for the wind speed data and hence it is to be preferred in this case.

5. Summary and Conclusions

We summarized some results from ARFIMA models and investigated the forecasts from this model as well as the ARIMA model. We also considered the bias in the estimate of the variance of the k -step ahead forecast error. This bias tends to increase with k and with the value of d , the fractional differencing parameter.

A set of wind speed data was analysed and forecasts from the ARFIMA and ARIMA models were generated and compared.

Acknowledgments

V.A. Reisen was partially supported by CNP, Brazil and B. Abraham was partially supported by grants from NSERC and MRCO. We like to thank Giovanni Merola for helping with some computations.

References

- Akaike, H. (1973), Maximum likelihood identification of gaussian autoregressive moving average models, *Biometrika*, 60 (2), 255-65.
- Beran, J. (1994), *Statistics for long memory process*, New York; Chapman and Hall.
- Chen G., Abraham, B. and Peires S. (1994), Lag window estimation of the degree of differencing in fractionally integrated time series models, *Journal of Time Series Analysis*, 15 (5), 473-487.
- Crato, N. and Ray, B. (1996), Model selection and forecasting for long-range dependent processes, *Journal of Forecasting*, 15, 107-125.
- Geweke, J. and Porter-Hudak, S. (1983), The estimation and application of Long memory time series model, *Journal of Time Series Analysis*, 4 (4), 221-238.
- Hassler, U. (1993), Regression of spectral estimator with fractionally integrated time series, *Journal of Time Series Analysis*, 14, 369-80, .
- Hosking, J. (1981), Fractional differencing, *Biometrika*, 68 (1), 165-176.
- Hosking, J. (1982), Some models of persistence in time series analysis : theory and practice 1, *North Holland Publishing Company*, 641-653.
- Hosking, J. (1984), Modelling persistence in hydrological time series using fractional differencing, *Water Resources Research*, 20 (12), 1898-1908.

Reisen, V. A. (1993), Long memory time series models, Thesis Ph.D. Department of Mathematics, UMIST, Manchester, U.K.

Reisen, V. A. (1994), Estimation of the fractional difference parameter in the ARIMA(p, d, q) model using the smoothed periodogram, *Journal of Time Series Analysis*, 15 (3), 335-350.