Temporal aggregation of stationary and nonstationary FARIMA$(p, d, 0)$ models

Jan Beran
Department of Mathematics and Statistics
University of Konstanz
Jan.Beran@uni-konstanz.de

and

Dirk Ocker
Department Riskmanagement
The Swiss Union of Raiffeisen Banks
dirk.ocker@raiffeisen.ch

26 July 2000

Abstract

We consider temporal aggregation of stationary and nonstationary time series with short memory, long memory and antipersistence, within the framework of fractional autoregressive processes. Asymptotically, long memory and antipersistence are preserved whereas short memory components vanish. In the case of integrated processes, the results extend Tiao's [15] to the fractional case.

Key words: temporal aggregation, differencing, long memory, short memory, antipersistence, stationarity, nonstationarity, fractional ARIMA, Box-Jenkins ARIMA, fractional Gaussian noise, white noise.

1 Introduction

Data are often available in an aggregated form (temporal aggregation). Typical examples are flow variables (e.g. industrial production) which exist only through aggregation over a certain time interval. Also, applied data analysts sometimes prefer to analyze aggregated data in order to eliminate seasonal fluctuations. Here, we consider the problem of temporal aggregation for a class of parametric time series models that includes classical nonfractional Box-Jenkins ARIMA$(p,m,0)$ models as well as stationary and nonstationary fractional autoregressive processes. Beran [3] (also see Beran, Bhansali and Ocker [4]) proposes a unified treatment by noting that these models are special cases of

$$\phi(B)(1-B)^{\delta}\{(1-B)^{m}X_t-\mu\}=\epsilon_t, \tag{1}$$

where ϵ_t are iid zero mean normal with $\sigma_\epsilon^2 = Var(\epsilon_t)$, B denotes the backshift operator such that $BX_t = X_{t-1}$. Also $\phi(x)=\sum_{j=0}^{p}\phi_j x^j$ is a polynomial with $\phi_0=1$ and roots outside the unit circle. The integer m is the number of times X_t must be differenced to achieve stationarity. The mth difference $(1-B)^m X_t$ is a stationary FARIMA$(p,\delta,0)$ process with fractional differencing parameter $\delta \in (-.5,.5)$ and expected value μ. The fractional difference $(1-B)^{\delta}$ is defined by

$$(1-B)^{\delta}=\sum_{k=0}^{\infty} b_k(\delta)B^k, \quad \text{where}$$

$$b_k(\delta)=(-1)^k\frac{\Gamma(\delta+1)}{\Gamma(k+1)\Gamma(\delta-k+1)}.$$

For $\delta = 0$, (1) reduces to an ARIMA$(p,m,0)$ model (Box & Jenkins [5]). For $m=0$, (1) reduces to a stationary FARIMA$(p,\delta,0)$ process (Granger & Joyeux [6], Hosking [7]). Definition (1) extends the concept of standard ARIMA$(p, m, 0)$ models by allowing the mth difference to be a stationary FARIMA$(p, \delta, 0)$ process with arbitrary $\delta \in (-.5, .5)$. Also, (1) is an extension of the FARIMA$(p, \delta, 0)$ models, by allowing the possibility of nonstationarity $(d > .5)$. The differencing parameter $d = m + \delta$ determines which (possibly fractional) difference has to be taken in order to obtain a stationary Box-Jenkins ARMA$(p, 0)$ process. We refer to (1) to as FARIMA$(p, d, 0)$ model.

Here, a stationary process $(1-B)^m X_t$ is said to be persistent or to have long memory (or long- range dependence) if $\sum_{k=-\infty}^{\infty} \rho(k) = \infty$, where $\rho(k) = Corr(X_t, X_{t+k})$. For (1) this is the case if $\delta \in (0, .5)$, since the correlations of a stationary FARIMA$(p, \delta, 0)$ process decay hyperbolically, i.e., $\rho(k) \sim_{|k|\to\infty} c_\rho(\phi, \delta)|k|^{2\delta-1}$, $c_\rho(\cdot) > 0$ (see e.g. Hosking [7]). For $\delta \in (-.5, 0)$, the correlations decay hyperbolically too, but $\sum_{k=-\infty}^{\infty} \rho(k) = 0$. The process is then said to be antipersitent. For $\delta = 0$, the differenced process $(1 - B)^m X_t$ becomes a standard short memory ARMA$(p, 0)$ model with summable exponentially decaying correlations. Thus (1) incorporates stationary and nonstationary, short- and long-range dependent autoregressive processes.

The impact of temporal aggregation on nonfractional ARIMA$(p, m, 0)$ models is well known (see e.g. Amemiya & Wu [1], Stram & Wei [12], Tiao [15], Wei [16]). For stationary fractional processes, related results follow directly from functional limit theorems (see Lamperti [9], Taqqu [13] and others). Nothing is however known about the effects upon nonstationary FARIMA$(p, d, 0)$ models. Many time series, in particular in economics, are nonstationary and appear to have long-range correlations or antipersistence. It is therefore important to know how a process changes when it is aggregated.

The purpose of this article is to derive the asymptotic behaviour of temporal aggregates of an observed time series x_t generated by the FARIMA$(p, d, 0)$ model (1). All results on temporal aggregation of non-fractional ARIMA$(p, m, 0)$, stationary fractional FARIMA$(p, \delta, 0)$ as well as nonstationary fractional FARIMA$(p, d, 0)$ models can be derived using one unified approach. The following definition of temporal aggregation is used.

Definition 1 *Let $t = sT$ and $s \geq 2$, then the series*

$$z_T = \left(\sum_{i=0}^{s-1} B^i \right) x_{sT},$$

represents the s-period nonoverlapping aggregates of x_t.

Tiao [15] has shown for Box-Jenkins ARIMA processes that, as the degree of aggregation s tends to infinity, the series z_T approaches an ARIMA$(0, m, m)$ model, irrespective of the ARMA order and the size of the coefficients. Thus the integer differencing parameter m determines the asymptotic process. The question arises whether a similar result holds in connection with fractional differencing.

The article is organized as follows. The main theoretical results are obtained in section 2. The results are illustrated by simulations in section 3. Final remarks in section 4 conclude the paper. Proofs are given in the appendix.

2 Asymptotic results

To simplify presentation, the moving average order q is chosen to be equal to zero. The same results hold for $q \neq 0$. Also, without loss of generality, we assume μ to be known and equal to zero. Since a FARIMA$(p, d, 0)$ process has an infinite moving-average representation for $\delta < .5$ (Hosking [7]), we can express (1) by

$$(1-B)^m x_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}, \tag{2}$$

where ψ_k is obtained by inverting (1). In order to derive the process for the aggregates z_T, we adapt Telser's [14] technique and multiply (2) by $(\sum_{i=0}^{s-1} B^i)^{m+1}$. Using progression we get

$$(1-B^s)^m \left(\sum_{i=0}^{s-1} B^i\right) x_t = \left(\sum_{i=0}^{s-1} B^i\right)^{m+1} u_t, \tag{3}$$

where $u_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}$. Now let

$$y_T = (1-B^s)^m \left(\sum_{i=0}^{s-1} B^i\right) x_{sT} = (1-B)^m z_T, \tag{4}$$

where B operates on T by $Bz_T = z_{T-1}$. That is, y_T is the mth difference of the aggregate z_T. The covariances $\gamma_y(k) = Cov(y_T, y_{T+k})$ of the process y_T are specified in the following lemma for $m = 0$ and 1.

Lemma 1 *Let $\gamma_y(k)$ be the autocovariance of the series $y_T = (1-B)^m z_T$ at lag k. Then*

(i) $\gamma_y(k) = \sum_{j,l=0}^{s-1} \gamma_u(j-l-sk)$, *if* $m = 0$,

(ii) $\gamma_y(k) = \sum_{j,l=-(s-1)}^{s-1} (s-|j|)(s-|l|)\gamma_u(j-l-sk)$, *if* $m = 1$,

where $\gamma_u(j-l-sk) = Cov(u_{i(s,T)+j}, u_{i(s,T+k)+l})$, *with* $i = i(s,T) = s(T-1)+1$, *for* $m = 1$, *and* $\gamma_u(j-l-sk) = Cov(u_{sT-l}, u_{s(T+k)-j})$, *for* $m = 0$, *respectively.*

Lemma 1 enables us to derive explicit formulas for the autocorrelations. In particular in the following theorem it is shown that the correlations $\rho_y(k) = \gamma_y(k)/\gamma_y(0)$ of y_T depend on the differencing parameter $d = m + \delta$ only. This is a generalization of Tiao's [15] result.

Theorem 1 *Let $\gamma_y(k)$ be the autocovariance and $\rho_y(k)$ the autocorrelation of $y_T = (1-B)^m z_T$ at lag k. Then, as s tends to infinity, the following holds:*

(i) If $d = m = 0$, then $\gamma_y(0) = O(s)$, $\gamma_y(k) = o(s)$ $(k \geq 1)$, and $\lim_{s\to\infty} \rho_y(k) = 0$;

(ii) If $d = m = 1$, then $\gamma_y(k) = O(s^3)$ $(k = 0, 1)$, $\gamma_y(k) = o(s)$ $(k \geq 2)$, $\lim_{s\to\infty} \rho_y(1) = .25$ and $\lim_{s\to\infty} \rho_y(k) = 0$ $(k \geq 2)$;

(iii) if $d = \delta \in (-.5, .5) \setminus \{0\}$, then $\gamma_y(k) = O(s^{1+2\delta})$ for all k and $\lim_{s\to\infty} \rho_y(k) = \frac{1}{2}\left\{(k+1)^{2\delta+1} - 2k^{2\delta+1} + (k-1)^{2\delta+1}\right\}$. This implies in particular

$$\rho_y(k) \sim_{k\to\infty} \delta(2\delta+1)k^{2\delta-1}$$

and, for $-\frac{1}{2} < \delta < 0$, $\sum_{k=-\infty}^{\infty} \rho_y(k) = 0$;

(iv) If $d = m + \delta \in (.5, 1.5) \setminus \{1\}$, then $\gamma_y(k) = O(s^{3+2\delta})$ for all k,

$$\lim_{s\to\infty} \rho_y(1) = \frac{7 + 3^{2\delta+3} - 2^{2\delta+5}}{2^{2\delta+4} - 8}$$

and for $k \geq 2$,

$$\lim_{s\to\infty} \rho_y(k) = \frac{(k+2)^{2\delta+3} - 4(k+1)^{2\delta+3} + 6k^{2\delta+3} - 4(k-1)^{2\delta+3} + (k-2)^{2\delta+3}}{2^{2\delta+4} - 8}.$$

This implies in particular

$$\rho_y(k) \sim_{k\to\infty} \frac{\delta(2\delta+1)(2\delta+2)(2\delta+3)}{(2^{2\delta+3} - 4)} k^{2\delta-1}$$

and, for $-\frac{1}{2} < \delta < 0$, $\sum_{k=-\infty}^{\infty} \rho_y(k) = 0$.

Remarks:

1. The limiting model of the aggregate y_T (equation (4)) of a nonfractional ARIMA process (1) (i.e. $\delta = 0$) becomes white noise, for $d = m = 0$, and approaches an MA(1) process, for $d = m = 1$, respectively. This result was first obtained by Tiao [15] using a different approach.

2. If the original series x_t is a stationary FARIMA model (1), with $d = \delta \in (-.5, .5) \setminus \{0\}$, then the asymptotic aggregate process y_T is fractional Gaussian noise (Mandelbrot & van Ness [10], Mandelbrot & Wallis [11]). Thus, long memory and antipersistence respectively remain asymptotically, whereas short memory model components vanish. This result can also be proved directly by using functional limit theorems for stationary fractional processes (see Lamperti [9], Taqqu [13] and others).

3. When the original series x_t is a nonstationary FARIMA$(p, d, 0)$ process, with $d = m + \delta \in (.5, 1.5) \setminus \{1\}$, then the limit of the aggregated series y_T is such that its first difference is a stationary Gaussian process with the same long-memory parameter δ as $x_t - x_{t-1}$. Thus, temporal aggregation of a nonstationary fractional autoregressive process (1) does not change the property of long-range dependence. Similarily, antipersistence is preserved. However, first difference of the process does not converge to fractional Gaussian noise. This is similar to standard integrated ARIMA processes ($\delta = 0$, $m = 1$), where the first difference of the limiting process does not converge to iid noise (which is fractional Gaussian noise with $d = 0$) but instead has a lag-1 correlation of .25. Note also that the correlation formulas exhibit a continuous transition between $\delta \neq 0$ and $\delta = 0$, since $\lim_{\delta \to 0} \rho_y(1) = \frac{1}{4}$ and $\lim_{\delta \to 0} \rho_y(k) = 0$ for $k \geq 2$.

3 Simulations

For d equal to $-.3, 0, .3, .7, 1.0, 1.3$, and sample size $n = 40000$, one hundred series of each of the following four models were simulated:

- Model a: $(1 + .9B)(1 - B)^{\delta}(1 - B)^m X_t = \epsilon_t$;
- Model b: $(1 - B)^{\delta}(1 - B)^m X_t = \epsilon_t$;
- Model c: $(1 - .9B)(1 - B)^{\delta}(1 - B)^m X_t = \epsilon_t$;
- Model d: $(1 - 1.42B + .73B^2)(1 - B)^{\delta}(1 - B)^m X_t = \epsilon_t$.

Model c has a very strong positive short memory and may converge quite slowly to the limiting model with increasing degree of temporal aggregation s. In contrast, Model b is expected to converge quite fast under temporal aggregation. Model a exhibits short-range correlations with alternating signs. Model d has a local maximum at a nonzero frequency implying random short-range periodicities.

Note that simulation by the S-Plus function *arima.fracdiff.sim* poses computational problems for large sample sizes. Therefore, the series were simulated combining the *simARMA0* function in Beran [2] (to obtained a fractional ARIMA$(0, \delta, 0)$ series) with the linear filter function *arima.sim*, and *cumsum* for $d \geq .5$.

For each simulated series ($i = 1, ..., 100$), and each aggregation level $s = 1, 2, 5, 10, 20, ..., 100$, the sample correlations $\hat{\rho}_{y,s;i}(k)$ ($k = 1, 2$) were calculated. Figures 1 and 2 display, for models 1 to 4, and k equal to 1 and 2 respectively, the simulated average sample correlations $\bar{\rho}_{y,s}(k) = 100^{-1} \sum_{i=1}^{100} \hat{\rho}_{y,s;i}(k)$, plotted against s. The straight lines mark the corresponding limiting correlation according to theorem 1. Note that, sample autocorrelations of fractional models are biased (Hosking [8]), and in the case of lon memory the bias decays very slowly as the sample size increases. In order that the effect of aggregation is not confounded with this bias, series of the same length $n = 400$ were used for each aggregation level s to calculate $\hat{\rho}_{y,s;i}(k)$. For Model c, we can observe a slow convergence to the corresponding limiting values, whereas convergence is fast for the other models. Moreover, there is a systematic negative bias in the long memory cases $d = .3$ and $d = 1.3$. This is, however, consistent with the bias expected theoretically (Hosking [8]). To illustrate this, table 3 lists 1. the theoretical limiting values of $\rho_y(k)$ ($k = 1, 2$), 2. the observed average sample correlation $\bar{\rho}_{y,s}(k)$ for the models 1 though 4, given an aggregation degree of $s = 100$, 3. the corresponding empirical bias $\hat{b} = \bar{\rho}_{y,s}(k) - \rho_y(k)$ and 4. the theoretical bias, for $n = 400$. Recall that the theoretical bias given in Hosking [8] is

$$b = \begin{cases} (\rho_y(k) - 1) n^{2\delta - 1}, & m = 0 \\ (\rho_y(k) - 1)(2\delta + 2)(2\delta + 3)(2^{2\delta+3} - 4)^{-1} n^{2\delta - 1}, & m = 1 \end{cases} . \quad (5)$$

The empirical bias is always negative, as is expected by equation (5), and slightly larger than the theoretical one. The difference is, however, quite small as can be seen from the ratio $b/\hat{b}$. Overall, around 80% of the observed bias are explained by equation (5), supporting the practical relevance of theorem 1.

Table 1: Empirical vs. theoretical bias.

	$d = .3$					$d = 1.3$				
	$\rho_y(k)$	$\bar{\rho}_y(k)$	b	$\hat{b}$	$b/\hat{b}$	$\rho_y(k)$	$\bar{\rho}_y(k)$	b	$\hat{b}$	$b/\hat{b}$
Model a										
$k = 1$	.516	.457	-.044	-.059	.746	.658	.618	-.036	-.040	.900
2	.368	.292	-.058	-.076	.763	.431	.359	-.060	-.072	.833
5	.253	.172	-.068	-.081	.840	.292	.205	-.074	-.087	.851
10	.191	.120	-.074	-.071	1.042	.220	.123	-.082	-.097	.845
15	.163	.080	-.076	-.083	.916	.187	.085	-.085	-.102	.833
20	.145	.062	-.078	-.083	.940	.167	.053	-.087	-.114	.763
Model b										
$k = 1$	.516	.466	-.044	-.050	.880	.658	.603	-.036	-.055	.655
2	.368	.303	-.058	-.065	.892	.431	.342	-.060	-.089	.674
5	.253	.177	-.068	-.076	.895	.292	.185	-.074	-.107	.692
10	.191	.112	-.074	-.079	.937	.220	.109	-.082	-.111	.739
15	.163	.079	-.076	-.084	.905	.187	.076	-.085	-.111	.766
20	.145	.052	-.078	-.093	.839	.167	.043	-.087	-.124	.702
Model c										
$k = 1$	.516	.486	-.044	-.030	1.467	.658	.622	-.036	-.036	1.000
2	.368	.308	-.058	-.060	.967	.431	.366	-.060	-.065	.923
5	.253	.176	-.068	-.077	.883	.292	.206	-.074	-.086	.861
10	.191	.105	-.074	-.086	.861	.220	.129	-.082	-.091	.901
15	.163	.075	-.076	-.088	.864	.187	.092	-.085	-.095	.895
20	.145	.057	-.078	-.088	.886	.167	.066	-.087	-.101	.861
Model d										
$k = 1$	.516	.464	-.044	-.052	.846	.658	.604	-.036	-.054	.667
2	.368	.306	-.058	-.062	.936	.431	.346	-.060	-.085	.706
5	.253	.179	-.068	-.074	.919	.292	.192	-.074	-.100	.740
10	.191	.109	-.074	-.082	.902	.220	.118	-.082	-.102	.804
15	.163	.085	-.076	-.078	.974	.187	.074	-.085	-.113	.752
20	.145	.055	-.078	-.090	.867	.167	.052	-.087	-.115	.756

4 Final remarks

In this paper we investigated the effects of temporal aggregation on nonstationary and stationary short- and long-range dependent autoregressive processes. Long memory and antipersistence are robust with respect to temporal aggregation, whereas short memory is not. For stationary processes, aggregation leads asymptotically to fractional Gaussian noise. For integrated processes, a more complicated process is obtained in the limit which retains the long-memory properties of the original series. This extends and includes the case of an integrated short-memory ARMA process (i.e. $\delta = 0$) treated in Tiao's [15]. The assumption of Gaussian innovation is not essential for the derivation of the asymptotic correlation structure of the aggregated process y_T. However, an essential assumption was that the original process x_t is linear. In view of observed nonlinearties in many time series, the effect of aggregation will also need to be studied for nonlinear processes. In the case of long memory, this is likely to lead to nonstandard limit theorems similar to Taqqu [13].

5 Acknowledgements

This paper was supported in part by the Center of Finance and Econometrics at the University of Konstanz. Parts of this paper are based on the PhD thesis of the second author.

6 Appendix

Proof of lemma 1:

(i) Noting that $y_T = \sum_{l=0}^{s-1} u_{sT-l}$ the covariance is given by straightforward calculation as $\gamma_y(k) = \sum_{l,j=0}^{s-1} \gamma_u(j-l-sk)$.

(ii) For the case of $m = 1$, consider

$$y_T = \sum_{j=0}^{s-1} B^j \sum_{l=0}^{s-1} B^l u_{sT} = \sum_{j=s(T-2)+2}^{sT} (s - |j - s(T-1) - 1|) u_j$$

$$= \sum_{j=i(s,T)-(s-1)}^{i(s,T)+(s-1)} (s - |j - i(s,T)|) u_j = \sum_{j=-(s-1)}^{s-1} (s - |j|) u_{i(s,T)+j}.$$

Hence, the covariance equals

$$\begin{aligned} \gamma_y(k) &= Cov\left(\sum_{j=-(s-1)}^{s-1} (s-|j|) u_{i(s,T)+j}, \sum_{l=-(s-1)}^{s-1} (s-|l|) u_{i(s,T+k)+l} \right), \\ &= \sum_{j,l=-(s-1)}^{s-1} (s-|j|)(s-|l|)\gamma_u(j-l-sk), \end{aligned}$$

where $i(s,T) = s(T-1)+1$.

Proof of theorem 1:

First note that lemma 1 may be written as

(i)

$$\gamma_y(k) = \sum_{j=0}^{s-1} \sum_{v=j-s(1+k)+1}^{j-sk} \gamma_u(v); \tag{6}$$

(ii)

$$\begin{aligned}\gamma_y(k) \quad &= \quad \sum_{j=1-s}^{s-1} (s-|j|) \left\{ \sum_{v=j-sk}^{j-s(k-1)-1} [s(1-k)+j-v]\, \gamma_u(v) + \right. \\ &\quad \left. + \sum_{v=j-s(k+1)+1}^{j-sk-1} [s(k+1)-j+v]\, \gamma_u(v) \right\}. \end{aligned} \tag{7}$$

Proof of (i) Since the covariances $\gamma_u(\cdot)$ of an ARMA process decay exponentially in the sense that there is an upper bound $|\gamma_u(h)| \leq ca^{|h|}$, where $0 < c < \infty, 0 < a < 1$ are constants, we have $|\gamma_y(k)| \leq c \sum_{j=0}^{s-1} \sum_{v=j-s(1+1)+1}^{j-sk} a^{|v|}$ so that

$$|\gamma_y(0)| \leq c \sum_{j=0}^{s-1} \sum_{v=j-s+1}^{j} a^{|v|} \leq c \sum_{j=0}^{s-1} \left\{ \sum_{v=j-s+1}^{-1} a^{-v} + \sum_{v=0}^{j} a^{v} \right\}$$

$$\leq c \sum_{j=0}^{s-1} \left\{ \left(\frac{1-a^{s-j}}{1-a} \right) + \left(\frac{1-a^j}{1-a} \right) \right\} \leq \frac{c}{1-a} \left\{ \left(s - \frac{a^s-1}{1-a^{-1}} \right) + \left(s - \frac{1-a^s}{1-a} \right) \right\}$$

$$\leq s \frac{c}{1-a} = O(s)$$

and

$$|\gamma_y(k)| \leq c \sum_{j=0}^{s-1} \sum_{v=j-s(1+k)+1}^{j-sk} a^{|v|} \leq c \sum_{j=0}^{s-1} \sum_{v=sk}^{s(1+k)-1} a^{v-j} \leq ca^{sk} \left(\frac{1-a^{-s}}{1-a^{-1}} \right) \left(\frac{1-a^s}{1-a} \right)$$

$$\leq const \cdot a^{s(k-1)} = o(s), \forall k \geq 1.$$

Hence, $\rho_y(k) = 0, \forall k \geq 1$.

Proof of (ii) Suppose first that the ARMA covariances are such that $\gamma_u(v) = 0$ for $|v| > k_o$. Then from (7), we have

$$\gamma_y(0) \approx \sum_{j=1-s}^{s-1} (s-|j|) \left\{ (s+j) \sum_{v=j}^{j+s-1} \gamma_u(v) + (s-j) \sum_{v=j-s+1}^{j-1} \gamma_u(v) \right\}$$

$$\approx \sum_{j=-(s-1)}^{s-1} (s-|j|)^2 \sum_{v=-k_o}^{k_o} \gamma_u(v) \approx 2s^3 \int_0^1 (1-x)^2 dx \sum_{v=-k_o}^{k_o} \gamma_u(v)$$

$$\approx (2s^3/3) \sum_{v=-k_o}^{k_o} \gamma_u(v) = O(s^3)$$

and

$$\gamma_y(1) \approx \sum_{j=1-s}^{s-1} (s-|j|) \left\{ j \sum_{v=j-s}^{j-1} \gamma_u(v) + (2s-j) \sum_{v=j-2s+1}^{j-s-1} \gamma_u(v) \right\}$$

$$\approx \sum_{j=1}^{s} (s-|j|) j \sum_{v=-k_o}^{k_o} \gamma_u(v) \approx s^3 \int_0^1 (1-x) x dx \sum_{v=-k_o}^{k_o} \gamma_u(v)$$

$$\approx (s^3/6) \sum_{v=-k_o}^{k_o} \gamma_u(v) = O(s^3)$$

In addition we can show that $\gamma_y(k) = o(s^3), \forall k \geq 2$. Thus, we obtain $\rho_y(1) = .25$ and $\rho_y(k) = 0, \forall k \geq 2$. The general case, with $\gamma_u(v)$ not necessarily zero for $|v| > k_o$, follows by approximating the autocovariances by covariances that vanish for $|v| > k_o$, with a suitably chosen k_o.

Proof of (iii) Consider first the long memory case $\delta \in (0, .5)$. Since $\gamma_u(h) \sim_{|h| \to \infty} c_\gamma(\delta, \psi) |h|^{2\delta - 1}$, where $c_\gamma(\cdot) \neq 0$, (6) may be written as

$$\gamma_y(k) \sim c_\gamma \sum_{j=0}^{s-1} \sum_{v=j-s(1+k)+1}^{j-sk} |v|^{2\delta-1} \sim c_\gamma s^{2\delta+1} \int_0^1 \int_{x-1-k}^{x-k} |v|^{2\delta-1} dx dy = O(s^{2\delta+1}),$$

which yields

$$\gamma_y(0) \sim \frac{2 c_\gamma s^{2\delta+1}}{2\delta(2\delta+1)}$$

and

$$\gamma_y(k) \sim \frac{c_\gamma s^{2\delta+1}}{2\delta(2\delta+1)} \left\{ (k+1)^{2\delta+1} - 2k^{2\delta+1} + (k-1)^{2\delta+1} \right\}, \forall k \geq 1.$$

The correlations are

$$\rho_y(k) = \frac{1}{2} \left\{ (k+1)^{2\delta+1} - 2k^{2\delta+1} + (k-1)^{2\delta+1} \right\}, \forall k \geq 1.$$

For $\delta \in (-.5, 0)$,

$$\gamma_y(0) = \sum_{j=0}^{s-1} \sum_{v=j-s+1}^{j} \gamma_u(v) = -\sum_{j=0}^{s-1} \left\{ \sum_{v=j+1}^{\infty} \gamma_u(v) + \sum_{v=-\infty}^{j-s} \gamma_u(v) \right\}$$

$$\sim \frac{2 s^{1+2\delta} c_\gamma}{2\delta(2\delta+1)} = O(s^{2\delta+1})$$

and

$$\gamma_y(k) = \sum_{j=0}^{s-1} \sum_{v=j-s(1+k)+1}^{j-sk} \gamma_u(v) = \sum_{j=0}^{s-1} \left\{ \sum_{v=-\infty}^{j-sk} \gamma_u(v) - \sum_{v=-\infty}^{j-s(1+k)} \gamma_u(v) \right\}$$

$$\sim c_\gamma \sum_{j=0}^{s-1} \left\{ \sum_{v=-\infty}^{j-sk} (-v)^{2\delta-1} - \sum_{v=-\infty}^{j-s(1+k)} (-v)^{2\delta-1} \right\}$$

$$\sim s^{2\delta+1} c_\gamma \int_0^1 \left\{ \int_{-\infty}^{x-k} (-y)^{2\delta-1} dy - \int_{-\infty}^{x-1-k} (-y)^{2\delta-1} dy \right\} dx$$

$$\sim \frac{s^{1+2\delta} c_\gamma}{2\delta(2\delta+1)} \left\{ (k+1)^{2\delta+1} - 2k^{2\delta+1} + (k-1)^{2\delta+1} \right\} = O(s^{2\delta+1}), \forall k \geq 1.$$

The result then follows.

Proof of (iv) Again, we first consider the long memory case $\delta \in (0, .5)$. Lemma 1(ii) may be written as

$$\gamma_y(k) \sim c_\gamma s^{2\delta+3} \int_{-1}^{1} \int_{-1}^{1} (1-|x|)(1-|y|)|x-y-k|^{2\delta-1} dx dy = O(s^{2\delta+3}),$$

so that

$$\gamma_y(0) \sim \frac{c_\gamma s^{2\delta+3}(8 - 2^{2\delta+4})}{2\delta(2\delta+1)(2\delta+2)(2\delta+3)},$$

$$\gamma_y(1) \sim \frac{c_\gamma s^{2\delta+3}(2^{2\delta+5} - 7 - 3^{2\delta+3})}{2\delta(2\delta+1)(2\delta+2)(2\delta+3)},$$

$$\gamma_y(k) \sim \frac{-(k+2)^{2\delta+3} + 4(k+1)^{2\delta+3} - 6k^{2\delta+3} + 4(k-1)^{2\delta+3} - (k-2)^{2\delta+3}}{2\delta(2\delta+1)(2\delta+2)(2\delta+3)(c_\gamma s^{2\delta+3})^{-1}},$$

$\forall k \geq 2$. Thus, the correlations are

$$\rho_y(1) = \frac{2^{2\delta+5} - 7 - 3^{2\delta+3}}{8 - 2^{2\delta+4}},$$

and

$$\rho_y(k) = \frac{-(k+2)^{2\delta+3} + 4(k+1)^{2\delta+3} - 6k^{2\delta+3} + 4(k-1)^{2\delta+3} - (k-2)^{2\delta+3}}{8 - 2^{2\delta+4}},$$

$\forall k \geq 2$, respectively.

The formulas for the correlations in the case $\delta \in (-.5, 0)$ follow in a similar fashion, using (7) and the equality

$$\sum_{i=1}^{s} \gamma(i) = -\left\{ \sum_{i=s+1}^{\infty} \gamma(i) + \sum_{i=-\infty}^{0} \gamma(i) \right\} = \left\{ \sum_{i=1}^{\infty} \gamma(i) - \sum_{i=s+1}^{\infty} \gamma(i) \right\}.$$

References

[1] Amemiya, T., and Wu, R.Y., The effect of aggregation on prediction in the autoregressive model, *Journal of the American Statistical Association*, 67, 628-632, 1972.

[2] Beran, J., *Statistics for long-memory processes*, Chapman & Hall, New York, 1994.

[3] Beran, J., Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models, *Journal of the Royal Statistical Society*, 57, 659-672, 1995.

[4] Beran, J., Bhansali, R.J., and Ocker, D., On unified model selection for stationary and non- stationary short- and long-memory autoregressive processes, *Biometrika*, 85, 921-934, 1998.

[5] Box, G.E.P., and Jenkins, G.M., *Time series analysis: forecasting and control*, Holden Day, San Francisco, 1976.

[6] Granger, C.W.J., Joyeux, R., An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1, 15-29, 1980.

[7] Hosking, J.R.M., Fractional differencing, *Biometrika*, 68, 165-176, 1981.

[8] Hosking, J.R.M., Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series, *Journal of Econometrics*, 73, 261-284, 1996.

[9] Lamperti, J., Semi-stable stochastic processes, *Z. Wahrsch. verw. Geb.*, 22, 205-225, 1972.

[10] Mandelbrot, B.B., and van Ness, J.W., Fractional brownian motions, fractional noises and applications, *SIAM Review*, 10, 422-437, 1968.

[11] Mandelbrot, B.B., and Wallis, J.R., Computer experiments with fractional gaussian noises, *Water Resources Research*, 5, 228-267, 1969.

[12] Stram, D.O., and Wei, W.W.S., Temporal aggregation in the ARIMA process, *Journal of Time Series Analysis*, 7, 279-292, 1986.

[13] Taqqu, M.S., Weak convergence to fractional brownian motion and to the Rosenblatt process, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 31, 287-302, 1975.

[14] Telser, L.G., Discrete samples and moving sums in stationary stochastic processes, *Journal of the American Statistical Association*, 67, 484-499, 1967.

[15] Tiao, G.C., Temporal aggregation of time series, *Biometrika*, 59, 525-531, 1972.

[16] Wei, W.W.S., *Time series analysis: univariate and multivariate methods*, Addison-Wesley, New York et.al., 1990.

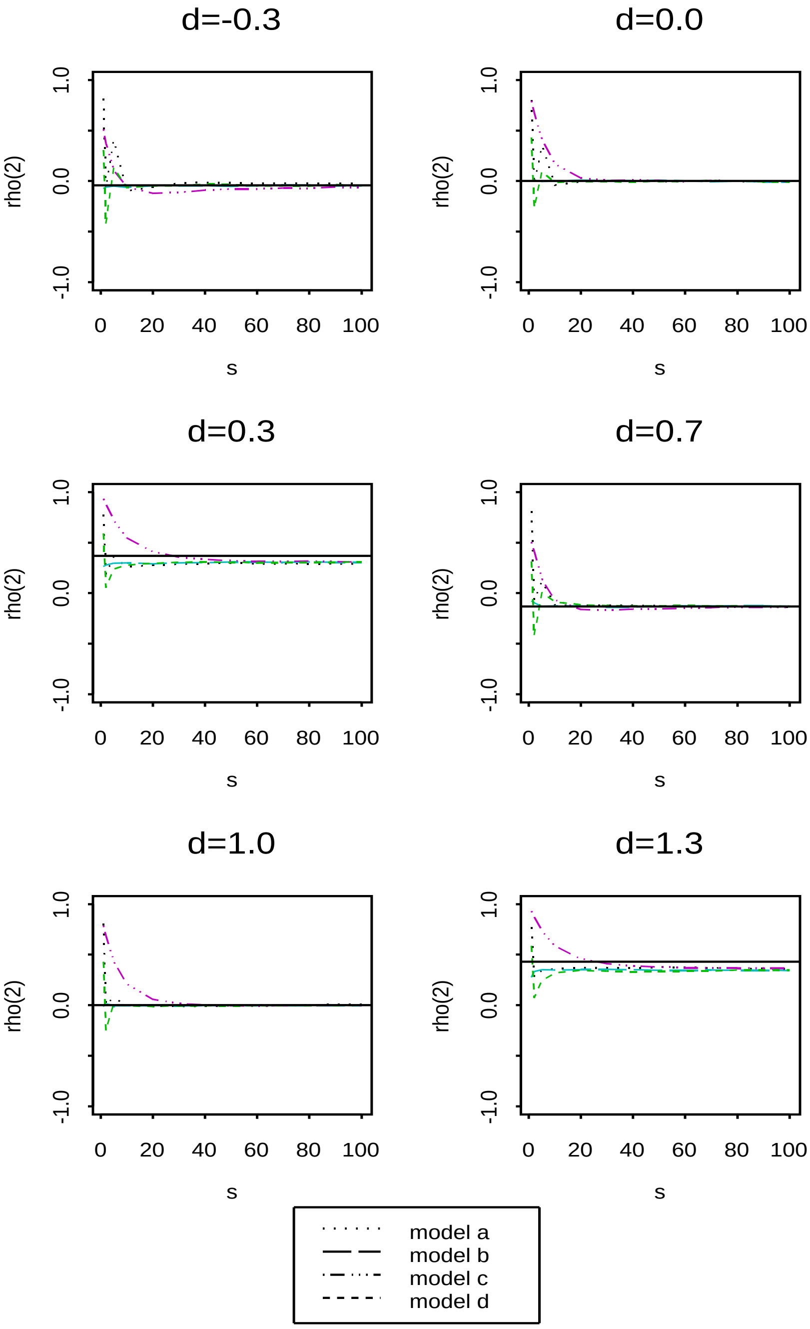

Figure 1: Lag-1 autocorrelations under temporal aggregation

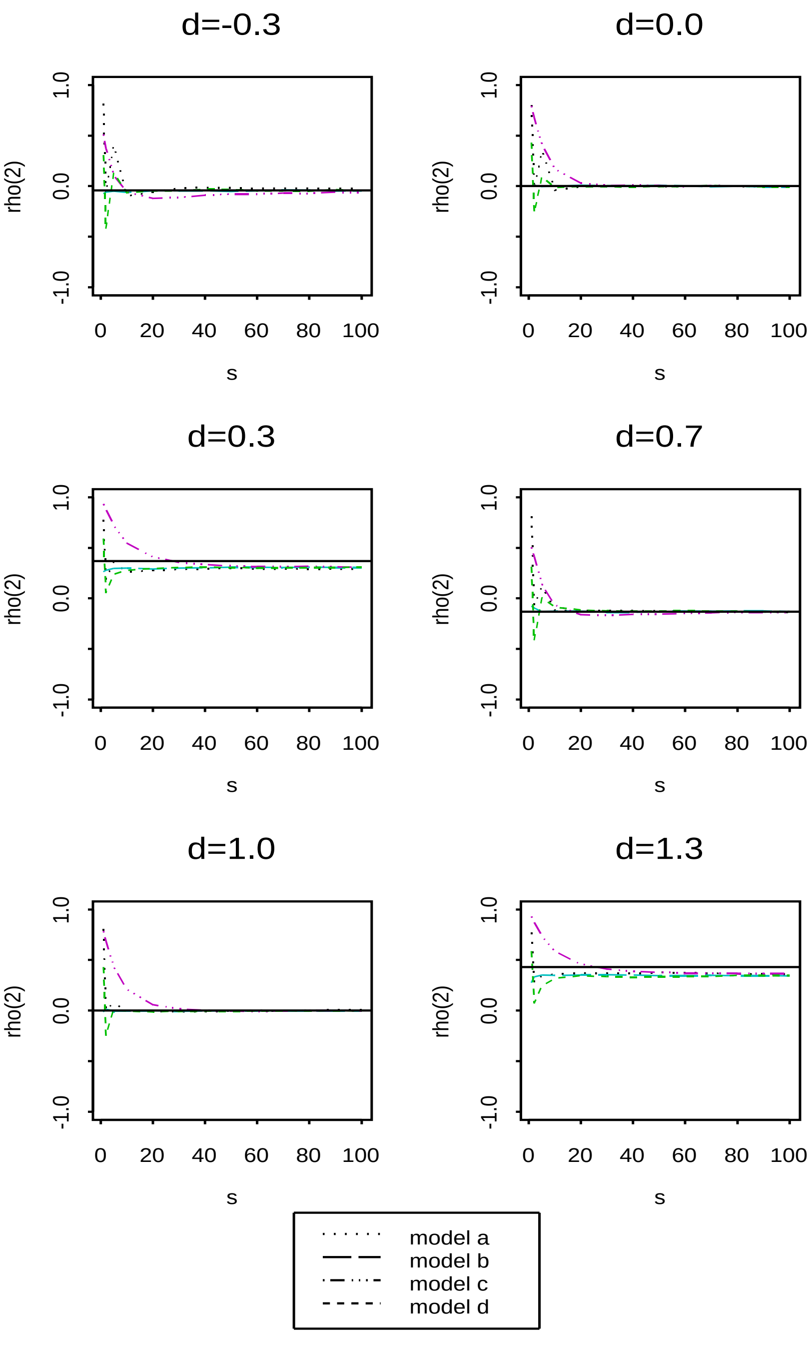

Figure 2: Lag-2 autocorrelations under temporal aggregation