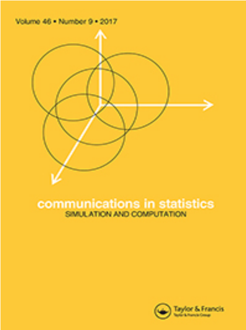

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Simulation and Computation

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: http://www.tandfonline.com/loi/lssp20

The effect of temporal aggregation on the estimation accuracy of time series models

Paulo Teles & Paulo S. A. Sousa

COMMUNICATIONS IN STATISTICS—SIMULATION AND COMPUTATION®
2017, VOL. 46, NO. 9, 6738–6759
http://dx.doi.org/10.1080/03610918.2016.1210169

Taylor & Francis
Taylor & Francis Group

The effect of temporal aggregation on the estimation accuracy of time series models

Paulo Teles[a,b] and Paulo S. A. Sousa[a]

[a]School of Economics, Porto University, Portugal; [b]LIAAD-INESC Porto LA, Porto University, Portugal

ABSTRACT
In time series analysis, Autoregressive Moving Average (ARMA) models play a central role. Because of the importance of parameter estimation in ARMA modeling and since it is based on aggregate time series so often, we analyze the effect of temporal aggregation on estimation accuracy. We derive the relationships between the aggregate and the basic parameters and compute the actual values of the former from those of the latter in order to measure and compare their estimation accuracy. We run a simulation experiment that shows that aggregation seriously worsens estimation accuracy and that the impact increases with the order of aggregation.

ARTICLE HISTORY
Received 19 February 2016
Accepted 4 July 2016

KEYWORDS
Aggregate time series; ARMA models; Estimation accuracy; Parameter values; Temporal aggregation

MATHEMATICS SUBJECT CLASSIFICATION
37M10; 62M10

1. Introduction

In time series analysis, Autoregressive Moving Average (ARMA) processes play a central role because they can describe a wide variety of time series in practice. Let us assume that the time series X_t follows a stationary and invertible Autoregressive Moving Average model of order (p, q) or ARMA(p, q):

$$\phi(B)X_t = \theta(B)a_t, \tag{1}$$

where $\phi(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$ and $\theta(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$ are the autoregressive and the moving average operators, respectively, B is the backshift operator such that $B^j X_t = X_{t-j}$ and a_t is an independent white noise process, that is, a sequence of iid random variables with zero mean and constant variance σ_a^2. We assume that the roots of $\phi(B)$ and of $\theta(B)$ are all outside the unit circle, that these polynomials have no roots in common and that $\phi(B) = \prod_{i=1}^{p}(1 - \delta_i B)$, where δ_i^{-1} $(i = 1, \ldots, p)$ denote the roots of $\phi(B)$. If $q = 0$, (1) becomes an ARMA$(p, 0)$ or AR(p) model, called an Autoregressive model:

$$\phi(B)X_t = a_t \Leftrightarrow \prod_{i=1}^{p}(1 - \delta_i B)X_t = a_t. \tag{2}$$

If $p = 0$, (1) becomes an ARMA$(0, q)$ or MA(q), called a moving average model:

$$X_t = \theta(B)a_t. \tag{3}$$

CONTACT Paulo Teles pteles@fep.up.pt School of Economics, Porto University, R. Dr. Roberto Frias, 4200-464 Porto, Portugal.

Many time series available in practice are obtained through temporal aggregation, that is, because of the process of data collection the only available observations are time series aggregates. For example, the series commonly used in the analysis are monthly, quarterly, or annual totals such as the Gross Domestic Product (GDP), investment, or rainfall amount. The effects of temporal aggregation on (univariate and multivariate) ARMA models and their use have been studied by several authors such as Amemiya and Wu (1972), Brewer (1973), Abraham (1982), Ahsanullah and Wei (1984), Weiss (1984), Stram and Wei (1986), Wei (1978, 2006), Drost (1994), Marcellino (1999), and Silvestrini and Veredas (2008). We first introduce the definition of aggregate time series and some results that will be used later concerning temporal aggregation. Suppose that the analyzed time series Y_T is the m-period nonoverlapping aggregates of X_t defined by

$$Y_T = \sum_{t=m(T-1)+1}^{mT} X_t = (1 + B + \cdots + B^{m-1})X_{mT} = \sum_{j=0}^{m-1} B^j X_{mT}, \tag{4}$$

where m is fixed and is called the order of aggregation and T is the aggregate time unit. For example, if X_t is a monthly time series and $m = 3$, then Y_T is a quarterly series. The time series X_t and Y_T will be called the basic and the aggregate time series, respectively (note that $m = 1$ is the situation of no aggregation, that is, the basic time series). Therefore, as expression (4) shows, we are discussing temporal aggregation of flow variables (such as the GDP or the rainfall amount), which is obtained through nonoverlapping sums of the basic series.

The derivation of the aggregate model is given in Stram and Wei (1986) and in Wei (2006, chap. 20). Assuming there are no hidden periodicities of order m in the AR operator $\phi(B)$ of (1), that is, assuming its roots are such that $\delta_i^m = \delta_j^m$ if and only if $\delta_i = \delta_j$ $(i, j = 1, \ldots, p;\ i \neq j)$ and multiplying by $\prod_{i=1}^{p} \frac{1-\delta_i^m B^m}{1-\delta_i B}(1 + B + \cdots + B^{m-1})$ in (1), we get

$$\begin{aligned}
&\prod_{i=1}^{p}(1 - \delta_i^m B^m)(1 + B + \cdots + B^{m-1})X_t = \prod_{i=1}^{p} \frac{1-\delta_i^m B^m}{1-\delta_i B}(1 + B + \cdots + B^{m-1})\theta(B)a_t \\
\Leftrightarrow &\prod_{i=1}^{p}(1 - \delta_i^m B^m)(1 + B + \cdots + B^{m-1})X_t = \prod_{i=1}^{p}\sum_{j=0}^{m-1}(\delta_i B)^j (1 + B + \cdots + B^{m-1})\theta(B)a_t.
\end{aligned} \tag{5}$$

Letting $W_t = \prod_{i=1}^{p}(1 - \delta_i^m B^m)(1 + B + \cdots + B^{m-1})X_t$, and since the a_t are iid variables with zero mean, it is easily seen that $E(W_{mT}) = 0$ and

$$\mathrm{Cov}(W_{mT}, W_{mT+mK}) = E(W_{mT}W_{mT+mK}) = 0 \tag{6}$$

for $K > [p + 1 + \frac{q-p-1}{m}]$, where $[z]$ denotes the integer part of z. Therefore, from (5) and (6), the aggregate series Y_T defined in (4) follows a stationary and invertible ARMA(p, Q) model:

$$\Phi(\mathcal{B})Y_T = \Theta(\mathcal{B})\varepsilon_T, \tag{7}$$

where $\mathcal{B}$ is the backshift operator on the aggregate time unit T such that $\mathcal{B}^j Y_T = Y_{T-j}$, $\Phi(\mathcal{B}) = (1 - \Phi_1\mathcal{B} - \cdots - \Phi_p\mathcal{B}^p) = \prod_{i=1}^{p}(1 - \delta_i^m\mathcal{B})$ and $\Theta(\mathcal{B}) = (1 - \Theta_1\mathcal{B} - \cdots - \Theta_Q\mathcal{B}^Q)$ are the aggregate autoregressive and moving average operators, respectively, whose roots are all outside the unit circle, $Q \leq [p + 1 + \frac{q-p-1}{m}]$ and ε_T is an independent white noise process with zero mean and variance σ_ε^2. We note that the autoregressive order remains unchanged by aggregation and that the roots of $\Phi(\mathcal{B})$ are the mth powers of the roots of $\phi(B)$.

However, as mentioned above, these results were based on the assumption that there are no hidden periodicities in the AR operator of the basic model (1) but sometimes that assumption may not be true, that is, $\delta_i \neq \delta_j$ and $\delta_i^m = \delta_j^m$ for some i, j, and m. In that case, both the AR and the MA order of the aggregate model are reduced (details can be found in Stram and Wei, 1986, and in Wei, 2006, chap. 20). Nevertheless, the assumption of no hidden periodicity is enough for our purposes. The parameters $\Phi_1, \ldots, \Phi_p$ are functions of $\phi_1, \ldots, \phi_p$ and the parameters $\Theta_1, \ldots, \Theta_Q$ and σ_ε^2 are functions of $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$ and σ_a^2 but these functions are generally very complicated.

Estimation of the parameters in (1) and (7) is extremely important because it is required for the use of ARMA models in practice and is based on the basic or on the aggregate time series, respectively. Assuming normal distribution of the white noise process, the maximum likelihood method is commonly used (Wei, 2006, chap. 7) and, denoting the vector of all the parameters of a given model by η, the information loss in estimation caused by aggregation is measured by $\tau(m) = 1 - \det(\mathbf{I}_a(\eta))/\det(\mathbf{I}_b(\eta))$, where $\mathbf{I}_b(\eta)$ and $\mathbf{I}_a(\eta)$ are the information matrices for the vector of all the parameters in the basic and in the aggregate models, respectively, and det denotes the determinant of a matrix (Wei, 2006, chap. 20). However, since the relationships between the parameters of models (1) and (7) are extremely complicated, as mentioned above, the derivation of $\tau(m)$ is generally very difficult.

Because of the central role of estimation in ARMA modeling and since it is based on aggregate data so often, assessing the effect of temporal aggregation on ARMA estimates is very important. Therefore, expressing the aggregate ARMA parameters as functions of those in the basic model for the most commonly used models in practice enables us to compute the actual values of the latter parameters from those of the former. With the parameter actual values, we next compare the estimation accuracy for the basic and the aggregate time series based on a simulation experiment, which shows how temporal aggregation can affect ARMA modeling in applied analysis through its impact on parameter estimation. Being based on the aggregate parameter actual values, such analysis has not been previously done (see the references mentioned above). First, we derive the relationships between the parameters in the aggregate and in the basic models in Section 2. Based on these relationships, we compute the actual values of the aggregate parameters for several basic autoregressive models and conduct a simulation experiment to measure the impact of aggregation on estimation accuracy in Section 3. Finally, concluding remarks are given and a reference to related issues is made in Section 4. The proofs of the results shown are left to the Appendix.

2. Relationship between aggregate and basic parameters

We first derive the relationships between the parameters in the aggregate and in the basic models, starting with autoregressive parameters.

2.1. Autoregressive parameters

As mentioned above, the aggregate AR parameters depend exclusively on the basic AR parameters for any ARMA model. That relationship, based on the roots of the AR polynomial, is given in Proposition 2.1.

Proposition 2.1. *Let X_t be a basic time series that follows the stationary and invertible Autoregressive Moving Average model of order (p, q) or ARMA(p, q) given in (1), where $\phi(B) = \prod_{i=1}^{p}(1 - \delta_i B)$ and δ_i^{-1} $(i = 1, \ldots, p)$ are the roots of $\phi(B)$. The m-order aggregate time series*

Y_T, given in (4), follows the ARMA(p, Q) model (7), where $\Phi(\mathcal{B}) = 1 - \Phi_1\mathcal{B} - \cdots - \Phi_p\mathcal{B}^p = \prod_{i=1}^{p}(1 - \delta_i^m\mathcal{B})$. Then, the autoregressive parameters of the ARMA model of Y_T, $\Phi_1, \ldots, \Phi_p$, are the following functions of $\delta_1, \ldots, \delta_p$:

$$\Phi_i = (-1)^{i-1} \sum_{j_1=1}^{p-(i-1)} \sum_{j_2=j_1+1}^{p-(i-2)} \cdots \sum_{j_i=j_{i-1}+1}^{p} \prod_{h=1}^{i} \delta_{j_h}^m \qquad i = 1, \ldots, p. \tag{8}$$

2.2. Moving average parameters

Since the above proposition shows the relationship between the aggregate autoregressive parameters and the parameters of the basic ARMA model, it is also necessary to find the corresponding relationship for the moving average parameters, which depends on the class of the basic model. In this article, we consider basic autoregressive models only, leaving moving average and autoregressive moving average models for future work.

Let the basic time series X_t follow the AR(p) model (2). Then, expression (5) becomes

$$W_t = \prod_{i=1}^{p} \sum_{j=0}^{m-1} (\delta_i B)^j (1 + B + \cdots + B^{m-1}) a_t \tag{9}$$

and the aggregate model is an ARMA(p, Q_1) where $Q_1 \le [p + 1 - \frac{p+1}{m}]$. In order to find the expressions of the parameters of the aggregate model as functions of those in the basic model, W_t given in (9) is written in the powers of B as

$$W_t = \left[\sum_{i=0}^{m-1} \left(1 + \sum_{j=1}^{i} \sum_{h_1=1}^{p} \sum_{h_2=h_1}^{p} \cdots \sum_{h_j=h_{j-1}}^{p} \prod_{u=1}^{j} \delta_{h_u} \right) B^i \right. \tag{10a}$$

$$+ \sum_{i=m}^{(m-1)p} \left(\sum_{j=i-(m-1)}^{i} \sum_{h_1=1}^{p} \sum_{h_2=h_1}^{p} \cdots \sum_{h_j=h_{j-1}}^{p} \prod_{u=1}^{j} \delta_{h_u} \right) B^i \tag{10b}$$

$$\left. + \sum_{i=(m-1)p+1}^{(m-1)(p+1)} \left(\sum_{j=i-(m-1)}^{(m-1)p} \sum_{h_1=1}^{p} \sum_{h_2=h_1}^{p} \cdots \sum_{h_j=h_{j-1}}^{p} \prod_{u=1}^{j} \delta_{h_u} \right) B^i \right] a_t, \tag{10c}$$

where, for $i = m, \ldots, (m-1)(p+1)$ and $j = m, \ldots, (m-1)p$, the powers $\delta_{h_u}^r$ are such that $r = \min(j, m-1)$, that is, the largest possible power is $\delta_{h_u}^{m-1}$, since the right-hand side of Eq. (9) involves only the powers $\delta_i, \ldots, \delta_i^{m-1}$ $(i = 1, \ldots, p)$. Let also Z_t denote

$$Z_j = \sum_{h_1=1}^{p} \sum_{h_2=h_1}^{p} \cdots \sum_{h_j=h_{j-1}}^{p} \prod_{u=1}^{j} \delta_{h_u}. \tag{11}$$

We note that, if the upper limit of a sum is less than its lower limit, that sum does not exist. Therefore, when $p = 1$, expression (10b) does not exist and W_t is simply the sum of (10a) and (10c). The determination of the moving average parameters of the aggregate model requires $\mathrm{Cov}(W_{mT}, W_{mT+mK})$ which is given in the next theorem.

Theorem 2.1. *Let X_t be a basic time series that follows the Autoregressive model of order p or AR(p) given in (2), W_t be defined in expressions (10a), (10b), and (10c) and Z_j be given in (11). Then, $\mathrm{Cov}(W_{mT}, W_{mT+mK}) = E(W_{mT} W_{mT+mK})$ for $K = 1, 2, \ldots$ is given in the following expressions.*

If $m = \frac{p}{p-K} \Leftrightarrow K = p - \frac{p}{m} \quad (p > K)$, *then*

$$Cov(W_{mT}, W_{mT+mK}) = \sum_{i=0}^{(m-1)p-mK} \left[\left(1 + \sum_{j=1}^{i} Z_j \right) \left(\sum_{j=i+1}^{i+m} Z_j \right) \right] \sigma_a^2 + \sum_{i=(m-1)p-mK+1}^{(m-1)p-m(K-1)-1} \left[\left(1 + \sum_{j=1}^{i} Z_j \right) \left(\sum_{j=i+m(K-1)+1}^{(m-1)p} Z_j \right) \right] \sigma_a^2. \quad (12)$$

If $\frac{p}{p-K} < m < \frac{p-1}{p-(K+1)} \Leftrightarrow p - 1 - \frac{p-1}{m} < K < p - \frac{p}{m} \quad (p > K+1)$, *then*

$$Cov(W_{mT}, W_{mT+mK}) = \sum_{i=0}^{(m-1)p-mK} \left[\left(1 + \sum_{j=1}^{i} Z_j \right) \left(\sum_{j=i+1}^{i+m} Z_j \right) \right] \sigma_a^2 + \sum_{i=(m-1)p-mK+1}^{m-1} \left[\left(1 + \sum_{j=1}^{i} Z_j \right) \left(\sum_{j=i+m(K-1)+1}^{(m-1)p} Z_j \right) \right] \sigma_a^2. \quad (13)$$

If $m = \frac{p-1}{p-(K+1)} \Leftrightarrow K = p - 1 - \frac{p-1}{m} \quad (p > K+1)$, *then*

$$Cov(W_{mT}, W_{mT+mK}) = \sum_{i=0}^{(m-1)p-mK} \left[\left(1 + \sum_{j=1}^{i} Z_j \right) \left(\sum_{j=i+1}^{i+m} Z_j \right) \right] \sigma_a^2 + \sum_{i=(m-1)p-mK+1}^{(m-1)p-m(K-1)-1} \left[\left(\sum_{j=i-(m-1)}^{i} Z_j \right) \left(\sum_{j=i+m(K-1)+1}^{(m-1)p} Z_j \right) \right] \sigma_a^2. \quad (14)$$

If $m > \frac{p-1}{p-(K+1)} \Leftrightarrow K < p - 1 - \frac{p-1}{m} \quad (p > K+1)$, *then*

$$Cov(W_{mT}, W_{mT+mK}) = \sum_{i=0}^{m-1} \left[\left(1 + \sum_{j=1}^{i} Z_j \right) \left(\sum_{j=i+1}^{i+m} Z_j \right) \right] \sigma_a^2 + \sum_{i=m}^{(m-1)p-mK} \left[\left(\sum_{j=i-(m-1)}^{i} Z_j \right) \left(\sum_{j=i+m(K-1)+1}^{i+mK} Z_j \right) \right] \sigma_a^2 + \sum_{i=(m-1)p-mK+1}^{(m-1)p-m(K-1)-1} \left[\left(\sum_{j=i-(m-1)}^{i} Z_j \right) \left(\sum_{j=i+m(K-1)+1}^{(m-1)p} Z_j \right) \right] \sigma_a^2. \quad (15)$$

If $m = \frac{p+1}{p-(K-1)} \Leftrightarrow K = p + 1 - \frac{p+1}{m} \quad (p > K-1)$, *then*

$$Cov(W_{mT}, W_{mT+mK}) = \sum_{i=0}^{(m-1)p-m(K-1)-1} \left[\left(1 + \sum_{j=1}^{i} Z_j \right) \left(\sum_{j=i+(m-1)(p-1)+1}^{(m-1)p} Z_j \right) \right] \sigma_a^2 + \sum_{i=m}^{(m-1)p-m(K-1)-1} \left[\left(\sum_{j=i-(m-1)}^{i} Z_j \right) \left(\sum_{j=i+(m-1)(p-2)}^{(m-1)p} Z_j \right) \right] \sigma_a^2. \quad (16)$$

If $\frac{p+1}{p-(K-1)} < m < \frac{p-1}{p-(K+1)} \Leftrightarrow p-1-\frac{p-1}{m} < K < p+1-\frac{p+1}{m} \quad (p > K+1)$, *then*

$$Cov(W_{mT}, W_{mT+mK}) = \sum_{i=m}^{(m-1)p-m(K-1)-1}\left[\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+(m-1)(p-2)}^{(m-1)p} Z_j\right)\right]\sigma_a^2. \quad (17)$$

If $\frac{p+1}{p-(K-1)} < m < \frac{p}{p-K} \Leftrightarrow p-\frac{p}{m} < K < p+1-\frac{p+1}{m} \quad (p > K)$, *then*

$$Cov(W_{mT}, W_{mT+mK}) = \sum_{i=0}^{(m-1)p-m(K-1)-1}\left[\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+(m-1)(p-1)+1}^{(m-1)p} Z_j\right)\right]\sigma_a^2. \quad (18)$$

$Cov\,(W_{mT}, W_{mT+mK}) = E(W_{mT}W_{mT+mK}) = 0$ *for other* $K > 0$.

Consequently, since the covariance is zero for $K > p+1-\frac{p+1}{m}$, W_{mT} is an MA(Q) process with $Q = [p+1-\frac{p+1}{m}]$, confirming that the aggregate time series follows the ARMA(p, Q) model (7). Based on the above results, it is now possible to find the parameter values of the aggregate model from the values of the parameters of the basic model. However, since it is not possible to determine a general expression for the moving average parameters as expression (8) for the autoregressive parameters, we focus on basic AR(p) models with $p = 1, 2, 3$, the most common in practice. The relationship between the aggregate and the basic white noise variances is also derived.

2.2.1. AR(1) models

When $p = 1$ in (2), the basic model is $(1-\phi_1 B)X_t = a_t$ with $\delta_1 = \phi_1$ and (5) is

$$\left(1-\delta_1^m B^m\right)(1+B+\cdots+B^{m-1})X_t = \sum_{j=0}^{m-1}(\delta_1 B)^j(1+B+\cdots+B^{m-1})a_t. \quad (19)$$

Consequently, from (10a), (10b), and (10c), W_t becomes

$$W_t = \left[\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}\delta_1^j\right)B^i + \sum_{i=m}^{2(m-1)}\left(\sum_{j=i-(m-1)}^{m-1}\delta_1^j\right)B^i\right]a_t.$$

Therefore, since a_t is an independent white noise process, W_{mT} follows an MA(1) process and, from (7), the aggregate time series Y_T follows the ARMA$(1, 1)$ model

$$(1-\Phi_1 B)Y_T = (1-\Theta_1\mathcal{B})\varepsilon_T. \quad (20)$$

Ahsanullah and Wei (1984) provide the expressions of the parameters of the aggregate model as functions of those in the basic model. We summarize those results next because they will be useful for higher-order models. Since $\delta_1 = \phi_1$ and from Proposition 2.1, $\Phi_1 = \delta_1^m = \phi_1^m$. Concerning the moving average parameters, since a_t is an independent white noise process with zero mean and variance $\mathrm{Var}(a_t) = \sigma_a^2$, $E(W_{2T}) = 0$ and we have

$$\mathrm{Cov}(W_{mT}, W_{mT+m}) = E(W_{mT}W_{mT+m}) = \left[\sum_{i=0}^{m-2}\left(1+\sum_{j=1}^{i}\delta_1^j\right)\left(\sum_{j=i+1}^{m-1}\delta_1^j\right)\right]\sigma_a^2 = \beta_2\sigma_a^2$$

and $\mathrm{Cov}(W_{mT}, W_{mT+mK}) = 0$ for $K \geq 2$ which shows that W_{mT} is an MA(1) process, that is, $W_{mT} = \varepsilon_T - \Theta_1\varepsilon_{T-1}$. Consequently, $\mathrm{Cov}(W_{mT}, W_{mT+m}) = -\Theta_1\sigma_\varepsilon^2 = \beta_2\sigma_a^2$ and, since ε_T and a_t are both independent white noise processes, $\mathrm{Var}(\varepsilon_T) = \sigma_\varepsilon^2$ is such that

$$\mathrm{Var}(W_{mT}) = \left(1+\Theta_1^2\right)\sigma_\varepsilon^2 = \left[\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}\delta_1^j\right)^2 + \sum_{i=m}^{2(m-1)}\left(\sum_{j=i-(m-1)}^{m-1}\delta_1^j\right)^2\right]\sigma_a^2 = \beta_1\sigma_a^2.$$

Thus, $\mathrm{Var}(W_{mT})/\mathrm{Cov}(W_{mT}, W_{mT+m}) = (1+\Theta_1^2)\sigma_\varepsilon^2/(-\Theta_1\sigma_\varepsilon^2) = \beta_1\sigma_a^2/(\beta_2\sigma_a^2) = \beta_3$, say. This expression leads to the quadratic equation

$$\Theta_1^2 + \beta_3\Theta_1 + 1 = 0 \tag{21}$$

and Θ_1 is the root such that $|\Theta_1| < 1$. Furthermore, $\sigma_\varepsilon^2 = \beta_1/(1+\Theta_1^2)\sigma_a^2$.

2.2.2. AR(2) models

When $p = 2$ in (2), the basic model is $(1-\phi_1 B - \phi_2 B^2)X_t = a_t \Leftrightarrow \prod_{i=1}^{2}(1-\delta_i B)X_t = a_t$ and (5) is

$$\prod_{i=1}^{2}\left(1-\delta_i^m B^m\right)(1+B+\cdots+B^{m-1})X_t = \sum_{j=0}^{m-1}(\delta_1 B)^j\sum_{j=0}^{m-1}(\delta_2 B)^j(1+B+\cdots+B^{m-1})a_t. \tag{22}$$

Therefore, since W_t is given by (10a), (10b), and (10c) with $p = 2$, W_{mT} follows an MA(Q) process with $Q = [3 - \frac{3}{m}]$. Consequently, from (7), the aggregate time series Y_T follows the ARMA$(2, Q)$ model where $\Phi(\mathcal{B}) = 1 - \Phi_1\mathcal{B} - \Phi_2\mathcal{B}^2 = \prod_{i=1}^{2}(1-\delta_i^m\mathcal{B})$ with, from (8), $\Phi_1 = \delta_1^m + \delta_2^m$ and $\Phi_2 = -\delta_1^m\delta_2^m$. Since the order Q of the moving average operator depends on the value of m, its parameters will also depend on this value and it is necessary to consider $m = 2$ separately to find these parameters using Theorem 2.1.

- $m = 2$

  When $m = 2$, $Q = 1$ and Θ_1, the moving average parameter in (7) is given in the next proposition.

  Proposition 2.2. *Let X_t be a basic time series that follows the stationary Autoregressive model of order 2 or AR(2) given in (2) with $p = 2$, where $\phi(B) = (1-\phi_1 B - \phi_2 B^2) = \prod_{i=1}^{2}(1-\delta_i B)$ and δ_i^{-1} $(i = 1, 2)$ are the roots of $\phi(B)$ with $|\delta_i^{-1}| > 1$. The aggregate time series with aggregation order $m = 2$ is $Y_T = \sum_{j=0}^{1} B^j X_{2T} = \sum_{t=2(T-1)+1}^{2T} X_t$ that follows the ARMA$(2, 1)$ model (7) with $P = 2$ and $Q = 1$, that is, $(1-\Phi_1\mathcal{B} - \Phi_2\mathcal{B}^2)Y_T = (1-\Theta_1\mathcal{B})\varepsilon_T$ where $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2) = \prod_{i=1}^{2}(1-\delta_i^2\mathcal{B})$. Then, the moving average parameter Θ_1 is the root of the quadratic equation $\Theta_1^2 + \beta_3\Theta_1 + 1 = 0$ such that $|\Theta_1| < 1$, where $\beta_3 = \beta_1/\beta_2$ with $\beta_1 = [1 + (1+\delta_1+\delta_2)^2 + (\delta_1+\delta_2+\delta_1\delta_2)^2 + (\delta_1\delta_2)^2]$ and $\beta_2 = [\delta_1+\delta_2+\delta_1\delta_2 + (1+\delta_1+\delta_2)\delta_1\delta_2]$. Furthermore, the white noise variance is $Var(\varepsilon_T) = \sigma_\varepsilon^2 = \beta_1\sigma_a^2/(1+\Theta_1^2)$.*

- $m \geq 3$

  When $m \geq 3$, $Q = 2$ and Θ_1 and Θ_2, the moving average parameters in (7), are given in the next proposition.

  Proposition 2.3. *Let X_t be a basic time series that follows the stationary Autoregressive model of order 2 or AR(2) given in (2) with $p = 2$, where $\phi(B) = (1-\phi_1 B - \phi_2 B^2) = \prod_{i=1}^{2}(1-\delta_i B)$ $\delta_i^{-1}(i = 1, 2)$ are the roots of $\phi(B)$ with $|\delta_i^{-1}| > 1$. The aggregate time series Y_T given in (4) with aggregation order $m \geq 3$ follows the ARMA$(2, 2)$ model (7) with $P = 2$ and $Q = 2$, that is, $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2)Y_T = (1-\Theta_1\mathcal{B}-\Theta_2\mathcal{B}^2)\varepsilon_T$ where $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2) = \prod_{i=1}^{2}(1-\delta_i^2\mathcal{B})$. Then, the first moving average parameter Θ_1 is a real root of the quartic equations $\beta_4^2\Theta_1^4 + 2\beta_4\Theta_1^3 + (1+2\beta_4^2+\beta_4\beta_5)\Theta_1^2 + (2\beta_4+\beta_5)\Theta_1 + 1 = 0$ or $\beta_4^2\Theta_1^4 + 2\beta_4\Theta_1^3 + (1+2\beta_4^2+\beta_4^2\beta_6)\Theta_1^2 + (\beta_6+2)\beta_4\Theta_1 + 1 = 0$ and the second moving average parameter is $\Theta_2 = \beta_4\Theta_1/(\beta_4\Theta_1+1)$ where the solutions retained* *have to be such that* $|\Theta_2| < 1$, $\Theta_2 - \Theta_1 < 1$ *and* $\Theta_2 + \Theta_1 < 1$ *with* $\beta_4 = \beta_3/\beta_2$, $\beta_5 = \beta_1/\beta_2$, $\beta_6 = \beta_1/\beta_3$,

$$\beta_1 = \sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i} Z_j\right)^2 + \sum_{i=m}^{2(m-1)}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)^2 + \sum_{i=2(m-1)+1}^{3(m-1)}\left(\sum_{j=i-(m-1)}^{2(m-1)} Z_j\right)^2,$$

$$\beta_2 = \sum_{i=0}^{m-2}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+1}^{i+m} Z_j\right) + \left(1+\sum_{j=1}^{m-1} Z_j\right)\left(\sum_{j=m}^{2(m-1)} Z_j\right)$$

$$+ \sum_{i=m}^{2(m-1)-1}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+1}^{2(m-1)} Z_j\right),$$

$$\beta_3 = \sum_{i=0}^{m-3}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+m+1}^{2(m-1)} Z_j\right)$$

and Z_j *is given in (11). Furthermore, the white noise variance is* $Var(\varepsilon_T) = \sigma_\varepsilon^2 = \beta_1\sigma_a^2/(1+\Theta_1^2+\Theta_2^2)$.

2.2.3. AR(3) models

When $p=3$ in (2), the basic model is $(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)X_t = a_t \Leftrightarrow \prod_{i=1}^{3}(1-\delta_i B)X_t = a_t$ and (5) is $\prod_{i=1}^{3}(1-\delta_i^m B^m)(1+B+\cdots+B^{m-1})X_t = \prod_{i=1}^{3}\sum_{j=0}^{m-1}(\delta_i B)^j(1+B+\cdots+B^{m-1})a_t$. Therefore, since W_t is given by (10a), (10b), and (10c) with $p=3$, W_{mT} follows an MA(Q) process with $Q = [4-\frac{4}{m}]$. Consequently, from (7), the aggregate time series Y_T follows the ARMA$(3, Q)$ model where $\Phi(\mathcal{B}) = 1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2-\Phi_3\mathcal{B}^3 = \prod_{i=1}^{3}(1-\delta_i^m\mathcal{B})$ with, from (8), $\Phi_1 = \delta_1^m+\delta_2^m+\delta_3^m$, $\Phi_2 = -(\delta_1^m\delta_2^m+\delta_1^m\delta_3^m+\delta_2^m\delta_3^m)$, and $\Phi_3 = \delta_1^m\delta_2^m\delta_3^m$. Since the order Q of the moving average operator depends on the value of m, its parameters will also depend on this value and it is necessary to discuss the cases $m=2$ and $m=3$ separately to find these parameters using Theorem 2.1.

- $m = 2$

  When $m=2$, $Q=2$ and Θ_1 and Θ_2, the moving average parameters in (7), are given in the next proposition.

Proposition 2.4. *Let* X_t *be a basic time series that follows the stationary Autoregressive model of order 3 or AR(3) given in (2) with* $p=3$, *where* $\phi(B) = (1-\phi_1 B-\phi_2 B^2-\phi_3 B^3) = \prod_{i=1}^{3}(1-\delta_i B)$ *and* δ_i^{-1} $(i=1,2,3)$ *are the roots of* $\phi(B)$ *with* $|\delta_i^{-1}| > 1$. *The aggregate time series* Y_T *given in (4) with aggregation order* $m=2$ *follows the ARMA(3, 2) model (7) with* $P=3$ *and* $Q=2$, *that is,* $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2-\Phi_3\mathcal{B}^3)Y_T = (1-\Theta_1\mathcal{B}-\Theta_2\mathcal{B}^2)\varepsilon_T$ *where* $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2-\Phi_3\mathcal{B}^3) = \prod_{i=1}^{3}(1-\delta_i^3\mathcal{B})$. *Then, the moving average parameters* Θ_1 *and* Θ_2 *and the white noise variance* $Var(\varepsilon_T) = \sigma_\varepsilon^2$ *are determined as in Proposition 2.3 with*

$$\beta_1 = \sum_{i=0}^{1}\left(1+\sum_{j=1}^{i} Z_j\right)^2 + \sum_{i=2}^{3}\left(\sum_{j=i-1}^{i} Z_j\right)^2 + Z_3^2,$$

$$\beta_2 = \sum_{j=1}^{2} Z_j + (1+Z_1)\left(\sum_{j=2}^{3} Z_j\right) + \left(\sum_{j=1}^{2} Z_j\right) Z_3,$$

$$\beta_3 = Z_3$$

where Z_j is given in (11).

- $m = 3$

When $m = 3$, $Q = 2$ and Θ_1 and Θ_2, the moving average parameters in (7), are given in the next proposition.

Proposition 2.5. *Let X_t be a basic time series that follows the stationary Autoregressive model of order 3 or $AR(3)$ given in (2) with $p=3$, where $\phi(B)=(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)=\prod_{i=1}^{3}(1-\delta_i B)$ and δ_i^{-1} $(i=1,2,3)$ are the roots of $\phi(B)$ with $|\delta_i^{-1}|>1$. The aggregate time series Y_T given in (4) with aggregation order $m=3$ follows the $ARMA(3,2)$ model (7) with $P=3$ and $Q=2$, that is, $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2-\Phi_3\mathcal{B}^3)Y_T=(1-\Theta_1\mathcal{B}-\Theta_2\mathcal{B}^2)\varepsilon_T$ where $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2-\Phi_3\mathcal{B}^3)=\prod_{i=1}^{3}(1-\delta_i^3\mathcal{B})$. Then, the moving average parameters Θ_1 and Θ_2 and the white noise variance $Var(\varepsilon_T)=\sigma_\varepsilon^2$ are determined as in Proposition 2.3 with*

$$\beta_1=\sum_{i=0}^{2}\left(1+\sum_{j=1}^{i}Z_j\right)^2+\sum_{i=3}^{6}\left(\sum_{j=i-2}^{i}Z_j\right)^2+\sum_{i=7}^{8}\left(\sum_{j=i-2}^{6}Z_j\right)^2,$$

$$\beta_2=\sum_{i=0}^{2}\left(1+\sum_{j=1}^{i}Z_j\right)\left(\sum_{j=i+1}^{i+3}Z_j\right)+\left(\sum_{j=1}^{3}Z_j\right)\left(\sum_{j=4}^{6}Z_j\right)+\sum_{i=4}^{5}\left(\sum_{j=i-2}^{i}Z_j\right)\left(\sum_{j=i+1}^{6}Z_j\right),$$

$$\beta_3=\sum_{j=4}^{6}Z_j+\sum_{i=1}^{2}\left(1+\sum_{j=1}^{i}Z_j\right)\sum_{j=i+4}^{6}Z_j$$

where Z_j is given in (11).

- $m \geq 4$

When $m \geq 4$, $Q = 3$ and Θ_1, Θ_2 and Θ_3, the moving average parameters in (7), are given in the next proposition.

Proposition 2.6. *Let X_t be a basic time series that follows the stationary Autoregressive model of order 3 or $AR(3)$ given in (2) with $p=3$, where $\phi(B)=(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)=\prod_{i=1}^{3}(1-\delta_i B)$ and δ_i^{-1} $(i=1,2,3)$ are the roots of $\phi(B)$ with $|\delta_i^{-1}|>1$. The aggregate time series Y_T given in (4) with aggregation order $m\geq 4$ follows the $ARMA(3,3)$ model (7) with $P=3$ and $Q=3$, that is, $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2-\Phi_3\mathcal{B}^3)Y_T=(1-\Theta_1\mathcal{B}-\Theta_2\mathcal{B}^2-\Theta_3\mathcal{B}^3)\varepsilon_T$ where $(1-\Phi_1\mathcal{B}-\Phi_2\mathcal{B}^2-\Phi_3\mathcal{B}^3)=\prod_{i=1}^{3}(1-\delta_i^3\mathcal{B})$. Then, the first moving average parameter Θ_1 is a real root of the equations $\Theta_1^2+\Theta_2^2+\Theta_3^2-\beta_8\Theta_1(\Theta_2-1)-\beta_8\Theta_2\Theta_3+1=0$ or $\Theta_1^2+\Theta_2^2+\Theta_3^2-\beta_9(\Theta_1\Theta_3-\Theta_2)+1=0$ or $\Theta_1^2+\Theta_2^2+\Theta_3^2+\beta_{10}\Theta_3+1=0$, where the second moving average parameter Θ_2 is a root of the equations $\beta_6\Theta_2^2+(\beta_6\Theta_1^2+\beta_5\Theta_1+1)\Theta_2-(\beta_6\Theta_1^2+\beta_5\Theta_1)=0$ or $\beta_5\beta_7\Theta_2^2+(\beta_5\beta_7\Theta_1^2+\beta_5\Theta_1+1)\Theta_2-(\beta_7\Theta_1+1)\beta_5\Theta_1=0$ or $\beta_6\beta_7\Theta_2^2+(\beta_6\beta_7\Theta_1^2+\beta_6\Theta_1+\beta_7)\Theta_2-(\beta_7\Theta_1+1)\beta_6\Theta_1=0$ and the third moving average parameter Θ_3 is $\Theta_3=[\Theta_2-(1-\Theta_2)\beta_5\Theta_1]/(\Theta_1-\beta_5\Theta_2)$ or $\Theta_3=[\beta_6\Theta_1(1-\Theta_2)]/(\beta_6\Theta_2+1)$ or $\Theta_3=\beta_7\Theta_2/(\beta_7\Theta_1+1)$ with $\beta_5=\beta_3/\beta_2$, $\beta_6=\beta_4/\beta_2$, $\beta_7=\beta_3/\beta_2$, $\beta_8=\beta_1/\beta_2$, $\beta_9=\beta_1/\beta_3$, $\beta_{10}=\beta_1/\beta_4$,*

$$\beta_1 = \sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i} Z_j\right)^2 + \sum_{i=m}^{3(m-1)}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)^2 + \sum_{i=(m-1)p+1}^{4(m-1)}\left(\sum_{j=i-(m-1)}^{3(m-1)} Z_j\right)^2,$$

$$\beta_2 = \sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+1}^{i+m} Z_j\right) + \sum_{i=m}^{2(m-1)-1}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+1}^{i+m} Z_j\right)$$
$$+ \sum_{i=2(m-1)}^{3(m-1)-1}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+1}^{3(m-1)} Z_j\right),$$

$$\beta_3 = \sum_{i=0}^{m-3}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+m+1}^{i+2m} Z_j\right) + \sum_{i=m-2}^{m-1}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+m+1}^{3(m-1)} Z_j\right)$$
$$+ \sum_{i=m}^{2(m-2)}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+m+1}^{3(m-1)} Z_j\right),$$

$$\beta_4 = \sum_{i=0}^{m-4}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+2m+1}^{3(m-1)} Z_j\right)$$

and Z_j is given in (11). The solutions retained have to be such that the roots of the moving average polynomial $(1 - \Theta_1\mathcal{B} - \Theta_2\mathcal{B}^2 - \Theta_3\mathcal{B}^3)$ are outside the unit circle. Furthermore, the white noise variance is $Var(\varepsilon_T) = \sigma_\varepsilon^2 = \beta_1\sigma_a^2/(1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2)$.

3. Simulation experiment

In order to analyze the effects of temporal aggregation on parameter estimation, we conducted a simulation experiment where 10,000 (basic) time series of 6,000 observations each were generated from several stationary Autoregressive models with a zero mean, normal white noise process. The simulated time series were subsequently aggregated with orders $m = 2, 3, 4, 6, 8, 12$ and the basic models considered were the following: AR(1) with $\delta_1 = -0.99, -0.95, -0.9, -0.8, -0.6, -0.5, -0.2, 0.2, 0.5, 0.6, 0.8, 0.9, 0.95, 0.99$ (recall that $\delta_1 = \phi_1$); AR(2) with $(\delta_1, \delta_2) = (-0.8, -0.7), (-0.2, -0.5), (-0.5, -0.1), (-0.7, 0.3), (-0.5, 0.6), (-0.2, 0.4), (-0.2, 0.7), (0.4, -0.7), (0.4, -0.9), (0.4, -0.95), (0.2, 0.3), (0.3, 0.95), (0.5, 0.7), (0.7, 0.8)$ or $(\phi_1, \phi_2) = (-1.5, -0.56), (-0.7, -0.1), (-0.6, -0.05), (-0.4, 0.21), (0.1, 0.3), (0.2, 0.08), (0.5, 0.14), (-0.3, 0.28), (-0.5, 0.36), (-0.55, 0.38), (0.5, -0.06), (1.25, -0.285), (1.2, -0.35), (1.5, -0.56)$, respectively; AR(3) with $(\delta_1, \delta_2, \delta_3) = (-0.8, -0.65, 0.4), (-0.85, 0.3, 0.6), (0.4, -0.6, 0.8), (0.2, 0.5, 0.7), (0.3, 0.6, 0.5)$ or $(\phi_1, \phi_2, \phi_3) = (-1.05, 0.06, 0.208), (0.05, 0.585, -0.153), (0.6, 0.4, -0.192), (1.4, -0.59, 0.07), (1.4, -0.63, 0.09)$, respectively.

In every model, we considered $\sigma_a^2 = 0.5, 1,$ and 4 for the basic white noise variance. The parameters of the above models were estimated from the simulated time series for the different orders of aggregation and the estimation Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) were computed as $\text{MAE} = \sum_{\upsilon=1}^{10,000} |\alpha - \hat{\alpha}_\upsilon|/10,000$ and $\text{MAPE} = (1/10,000)\sum_{\upsilon=1}^{10,000} |(\alpha - \hat{\alpha}_\upsilon)/\alpha| \times 100$ for any parameter α and its estimates $\hat{\alpha}_\upsilon$ (AR and MA parameters and the white noise variance). The aggregate parameter values were computed from Proposition 2.1 for the AR parameters and each value of m and, for the MA

Table 1. Estimation accuracy for AR(1) models.

δ_1			m						
0.50			1	2	3	4	6	8	12
	Φ_1	value	0.5000	0.2500	0.1250	0.0625	0.0156	0.0039	0.0002
		MAE	0.0088	0.0372	0.0623	0.0944	0.1860	0.2856	0.4101
		MAPE	1.8	14.9	49.9	151.0	1190.3	7304.6	170871
	Θ_1	value		−0.1459	−0.1667	−0.1579	−0.1246	−0.0962	−0.0625
		MAE		0.2926	0.3345	0.3193	0.2943	0.3354	0.4286
		MAPE		200.5	200.7	202.2	236.2	348.7	686.0
	σ_ε	value	1.0000	1.8512	2.5981	3.2386	4.2882	5.1403	6.5193
		MAE	0.0073	0.0190	0.0324	0.0470	0.0763	0.1058	0.1652
		MAPE	0.7	1.0	1.2	1.5	1.8	2.1	2.5

parameters, from (21) for the basic AR(1) model and each value of m, from Propositions 2.2 ($m = 2$) and 2.3 ($m \geq 3$) for the basic AR(2) model and from Propositions 2.4 ($m = 2$), 2.5 ($m = 3$), and 2.6 ($m \geq 4$) for the basic AR(3) model. This procedure enabled us to measure the actual effect of temporal aggregation on estimation accuracy, which has not been previously done. Since different equations were derived to find the MA parameter values for the basic AR(2) model with $m \geq 3$ and AR(3) with $m \geq 4$ in Propositions 2.3 and 2.6, respectively, the solutions obtained from those equations were computed and compared, being all equal according to the results in the propositions.

Tables 1, 2, and 3 display the simulation results and the parameter actual values for the basic models AR(1) with $\delta_1 = 0.5$, AR(2) with $(\delta_1, \delta_2) = (0.5, 0.7)$, and AR(3) with $(\delta_1, \delta_2, \delta_3) = (0.3, 0.6, 0.5)$, respectively, and for $\sigma_a^2 = 1$ (recall that $m = 1$ denotes the basic time series). To save space, we only show these results as examples and the complete tables are in http://www.fep.up.pt/docentes/paulus/Supp_Material_CiSSC2016.pdf. We conclude the following.

For the basic time series, estimation of the AR parameters and the white noise standard deviation is usually accurate or even very accurate as shown by the estimation errors.

For aggregate series and concerning the ARMA parameters, the error is very large most of the times and is often extremely large, showing a very strong negative impact of aggregation

Table 2. Estimation accuracy for AR(2) models.

δ_1	δ_2			m						
0.5	0.7			1	2	3	4	6	8	12
		Φ_1	value	1.2000	0.7400	0.4680	0.3026	0.1333	0.0616	0.0141
			MAE	0.0097	0.0405	0.3093	0.2562	0.3130	0.3495	0.3881
			MAPE	0.8	5.5	66.1	84.7	234.9	567.8	2754.4
		Φ_2	value	−0.3500	−0.1225	−0.0429	−0.0150	−0.0018	−0.0002	-3.4×10^{-6}
			MAE	0.0097	0.0335	0.1370	0.0862	0.0816	0.1225	0.2739
			MAPE	2.8	27.4	319.4	574.4	4433.7	53261	8.1×10^6
		Θ_1	value		−0.3028	−0.3480	−0.3404	−0.2920	−0.2422	−0.1690
			MAE		0.6035	0.8022	0.7161	0.5707	0.5002	0.4620
			MAPE		199.3	230.5	210.4	195.5	206.5	273.4
		Θ_2	value			−0.0051	−0.0057	−0.0036	−0.0017	−0.0003
			MAE			0.1214	0.1004	0.1088	0.1383	0.2796
			MAPE			2384.3	1758.5	3021.1	7946.0	84733
		σ_ε	value	1.0000	2.7681	4.9058	7.0566	10.948	14.229	19.472
			MAE	0.0073	0.0288	0.0626	0.1034	0.1966	0.2963	0.4973
			MAPE	0.7	1.0	1.3	1.5	1.8	2.1	2.6

Table 3. Estimation accuracy for AR(3) models.

δ_1	δ_2	δ_3						m			
0.3	0.6	0.5			1	2	3	4	6	8	12
			Φ_1	value	1.4000	0.7000	0.3680	0.2002	0.0630	0.0208	0.0024
				MAE	0.0101	0.4318	0.4018	0.3397	0.4031	0.4345	0.4365
				MAPE	0.7	61.7	109.2	169.7	639.8	2091.8	18035.5
			Φ_2	value	−0.6300	−0.1449	−0.0362	−0.0097	−0.0008	-6.7×10^{-5}	-5.3×10^{-7}
				MAE	0.0164	0.3045	0.1724	0.4613	0.5049	0.5248	0.4978
				MAPE	2.6	210.1	476.0	4775.8	65575.3	7.8×10^{5}	9.4×10^{7}
			Φ_3	value	0.0900	0.0081	0.0007	6.6×10^{-5}	5.3×10^{-7}	4.3×10^{-9}	2.8×10^{-13}
				MAE	0.0103	0.0661	0.0517	0.1030	0.098	0.1587	0.3474
				MAPE	11.5	815.7	7079.5	156076	1.9×10^{7}	3.7×10^{9}	12.4×10^{13}
			Θ_1	value		−0.3961	−0.4197	−0.3832	−0.2941	−0.2245	−0.1418
				MAE		0.9049	0.8087	0.7617	0.6169	0.5559	0.4843
				MAPE		228.5	192.7	198.8	209.8	247.6	341.6
			Θ_2	value		−0.0091	−0.0158	−0.0124	−0.0050	−0.0018	−0.0002
				MAE		0.1780	0.1697	0.4425	0.4958	0.5239	0.5073
				MAPE		1949.9	1076.8	3576.8	9876.3	29934	267011
			Θ_3	value				-1.1×10^{-5}	-2.4×10^{-6}	-2.4×10^{-7}	-1.5×10^{-9}
				MAE				0.1752	0.1678	0.1955	0.3655
				MAPE				15.9×10^{5}	7.0×10^{6}	8.1×10^{7}	24.4×10^{9}
			σ_ε	value	1.0000	3.1401	5.7527	8.2850	12.6295	16.1347	21.615
				MAE	0.0072	0.0320	0.0729	0.1215	0.2299	0.3414	0.5768
				MAPE	0.7	1.0	1.3	1.5	1.8	2.1	2.7

even for low orders of aggregation. Even for $m = 2$, the lowest order, the MAE, and the MAPE are much larger than for the basic series. The error generally increases with the order of aggregation, often achieving extremely large values for the highest values of m, and is larger for the smaller (in absolute value) basic parameter values. For any given value of m and for the three values of the white noise standard deviation considered, the percentage errors are very close, which means that this standard deviation has only a very weak effect on the estimation accuracy of the ARMA parameters. The effect of aggregation is stronger for the AR(2) and AR(3) models, but it is not possible to clearly rank them because their accuracy does not dominate the other even though, concerning the moving average parameters, the latter appears to lead to larger errors. Therefore, the impact increases with the model order and consequently with the number of parameters, which can be generalized to higher-order processes. These results agree with the conclusion in Wei (2006, chap. 20) concerning the increasing information loss in estimation $\tau(m)$ with the order of aggregation but, as mentioned above, $\tau(m)$ is usually very difficult to compute.

Concerning the white noise standard deviation, the error is also larger for aggregate series and increases with the order of aggregation, but the impact is much weaker than for ARMA parameters, since the error is always low for any value of m and any model. Therefore, the estimation accuracy of this parameter is affected by the order of aggregation only and not by the model or the parameter values, nor even by the standard deviation itself.

Moreover, the values displayed in the tables show that the aggregate ARMA parameters decrease in absolute value as the order of aggregation increases, reflecting the fact that they tend to 0 as $m \to \infty$. This result is straightforward for AR parameters because, since $|\delta_i| < 1$ $(i = 1, \ldots, p)$ in (2), then $|\delta_i^m| \to 0$ as $m \to \infty$ and therefore $|\Phi_i| \to 0$ in (7). For a general stationary ARMA(p, q) model, Tiao (1972) shows that, as $m \to \infty$, the limiting aggregate model is a white noise process, which means that the ARMA parameters tend to 0. Consequently, fitting an ARMA model with aggregate time series can easily be misleading about its order, that is, about the correct p and Q values in (7), because a parameter value close to zero is likely to result in a nonsignificant estimate, leading to the wrong model identification. However, we did not take such problem into account and considered the appropriate orders in the aggregate models because our purpose was to show the impact of aggregation on parameter estimation only, without any further effects. We then conclude that, in applied analysis, the impact of aggregation can be extremely serious because it can lead to large errors in parameter estimates and even to model identification failure.

4. Concluding remarks

We analyzed how temporal aggregation affects estimation accuracy of ARMA models, particularly focusing on AR models. To this purpose, we started by deriving the aggregate autoregressive parameters for any basic ARMA model and the aggregate moving average parameters and white noise standard deviation for the most commonly used basic AR models. These expressions allowed the computation of the aggregate parameter actual values from those in the basic model, making it possible to measure and compare their estimation accuracy. With that purpose, we conducted a simulation experiment which showed that the effect of aggregation on the estimation accuracy of both AR and MA parameters is very strong and increases with the order of aggregation. On the contrary, that impact on the white noise standard deviation is much weaker, even though it is still relevant and also grows with the order of aggregation. Therefore, estimation accuracy can be low for aggregate time series, that is, aggregate parameter estimates are subject to large estimation errors (in absolute value), which can have serious negative consequences on the use of ARMA models, namely, on model identification, inference, and forecasting.

We considered (basic) autoregressive models only, but this problem also arises for moving average and mixed autoregressive moving average models. The impact of aggregation on forecasting accuracy is also particularly important (Hotta and Neto, 1993; Koreisha and Fang, 2004; Lütkepohl, 2009, e.g.) and should also be analyzed for the most commonly used models in practice. This analysis is currently under way but it will be reported later for reasons of space. Furthermore, the basic AR models considered are all stationary, but many time series are nonstationary, that is, the autoregressive polynomial has at least a unit root. Nevertheless, our conclusions will generally be valid for nonstationary data since, as shown by Wei (2006, chap. 20), the integration order of the basic model, that is, the number of differences required to remove nonstationarity (or the number of those unit roots), remains unchanged by aggregation. The parameter values of the aggregate model also depend on that order and their expressions as functions of the basic parameters will be more complicated but that will not change the consequences of aggregation on estimation accuracy and will worsen its impact. Therefore, we may conclude that, both for stationary and nonstationary time series, temporal aggregation negatively affects the estimation accuracy of ARMA parameters with a very strong impact.

Appendix: Proofs

Proof of Proposition 2.1. The autoregressive polynomial in the aggregate model (7) can be expanded in the powers of $\mathcal{B}$ as

$$\Phi(\mathcal{B}) = 1 - \Phi_1\mathcal{B} - \cdots - \Phi_p\mathcal{B}^p = \prod_{i=1}^{p}\left(1 - \delta_i^m\mathcal{B}\right) = 1 - \sum_{j_1=1}^{p}\delta_{j_1}^m\mathcal{B} + \sum_{j_1=1}^{p-1}\sum_{j_2=j_1+1}^{p}\delta_{j_1}^m\delta_{j_2}^m\mathcal{B}^2$$

$$-\sum_{j_1=1}^{p-2}\sum_{j_2=j_1+1}^{p-1}\sum_{j_3=j_2+1}^{p}\delta_{j_1}^m\delta_{j_2}^m\delta_{j_3}^m\mathcal{B}^3+\sum_{j_1=1}^{p-3}\sum_{j_2=j_1+1}^{p-2}\sum_{j_3=j_2+1}^{p-1}\sum_{j_4=j_3+1}^{p}\delta_{j_1}^m\delta_{j_2}^m\delta_{j_3}^m\delta_{j_4}^m\mathcal{B}^4-\cdots$$

$$+(-1)^{p-1}\sum_{j_1=1}^{2}\sum_{j_2=j_1+1}^{3}\cdots\sum_{j_{p-1}=j_{p-2}+1}^{p}\delta_{j_1}^m\ldots\delta_{j_{p-1}}^m\mathcal{B}^{p-1}+(-1)^p\delta_1^m\ldots\delta_p^m\mathcal{B}^p$$

$$=1+\sum_{i=1}^{p}(-1)^i\sum_{j_1=1}^{p-(i-1)}\sum_{j_2=j_1+1}^{p-(i-2)}\cdots\sum_{j_i=j_{i-1}+1}^{p}\prod_{h=1}^{i}\delta_{j_h}^m\mathcal{B}^i.$$

Thus, result (8) is obtained by equating the powers of $\mathcal{B}$ on both sides of this expression. □

Proof of Theorem 2.1. Since a_t is an independent white noise process with zero mean and variance $\mathrm{Var}(a_t)=E(a_t^2)=\sigma_a^2$, using (10a), (10b), and (10c), we obtain, for $K=0,1,2,\ldots,$

$$E\left[\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}Z_j\right)B^ia_{mT}\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}Z_j\right)B^ia_{mT+mK}\right]$$

$$=\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}Z_j\right)E\left(B^ia_{mT}\right)\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}Z_j\right)E\left(B^ia_{mT+mK}\right)=0. \tag{A.1}$$

$$E\left[\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}Z_j\right)B^ia_{mT}\sum_{i=m}^{(m-1)p}\left(\sum_{j=i-(m-1)}^{i}Z_j\right)B^ia_{mT+mK}\right]$$

$$=\begin{cases}\displaystyle\sum_{i=0}^{(m-1)p-mK}\left[\left(1+\sum_{j=1}^{i}Z_j\right)\left(\sum_{j=i+1}^{i+m}Z_j\right)\right]\sigma_a^2 & \text{if}\\ \quad -m+1\le -mp+p+mK\le 0\Leftrightarrow \dfrac{p}{p-K}\le m\le\dfrac{p-1}{p-(K+1)}\ \text{with } p>K+1 & \text{(A.2)}\\ \displaystyle\sum_{i=0}^{m-1}\left[\left(1+\sum_{j=1}^{i}Z_j\right)\left(\sum_{j=i+1}^{i+m}Z_j\right)\right]\sigma_a^2 & \text{if}\\ \quad -m+1> -mp+p+mK\Leftrightarrow m>\dfrac{p-1}{p-(K+1)}\ \text{with } p>K+1 & \text{(A.3)}\\ 0 \qquad \text{if } -mp+p+mK>0. & \text{(A.4)}\end{cases}$$

Note that if $-m+1\le -mp+p+mK\le 0\Leftrightarrow \frac{p-1}{p-(K+1)}\le m\le\frac{p}{p-K}$ with $p<K$, then $m\le 0$, or if $-m+1>-mp+p+mK\Leftrightarrow m<\frac{p-1}{p-(K+1)}$ with $p<K+1$, then $m\le 0$, which is impossible. Therefore, these conditions are not valid to determine the covariance.

$$E\left[\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}Z_j\right)B^i a_{mT}\sum_{i=(m-1)p+1}^{(m-1)(p+1)}\left(\sum_{j=i-(m-1)}^{(m-1)p}Z_j\right)B^i a_{mT+mK}\right]$$

$$=\begin{cases}\displaystyle\sum_{i=(m-1)p-mK+1}^{(m-1)p-m(K-1)-1}\left[\left(1+\sum_{j=1}^{i}Z_j\right)\left(\sum_{j=i+m(K-1)+1}^{(m-1)p}Z_j\right)\right]\sigma_a^2 & \text{if}\\ \quad -m+1<-mp+p+mK\le 0 \text{ and } -m+1\le -mp+p+m(K-1)+1 & \\ \Leftrightarrow \dfrac{p}{p-K}\le m<\dfrac{p-1}{p-(K+1)} \text{ with } p>K+1 \text{ and } m\le\dfrac{p}{p-K} \text{ with } p>K & \\ \Rightarrow m=\dfrac{p}{p-K} \text{ with } p>K & \text{(A.5)}\\ \displaystyle\sum_{i=(m-1)p-mK+1}^{m-1}\left[\left(1+\sum_{j=1}^{i}Z_j\right)\left(\sum_{j=i+m(K-1)+1}^{(m-1)p}Z_j\right)\right]\sigma_a^2 & \text{if}\\ \quad -m+1<-mp+p+mK\le 0 \text{ and } -m+1>-mp+p+m(K-1)+1 & \\ \Leftrightarrow \dfrac{p}{p-K}\le m<\dfrac{p-1}{p-(K+1)} \text{ with } p>K+1 \text{ and } m>\dfrac{p}{p-K} \text{ with } p>K & \\ \Rightarrow \dfrac{p}{p-K}<m<\dfrac{p-1}{p-(K+1)} \text{ with } p>K+1 & \text{(A.6)}\\ \displaystyle\sum_{i=0}^{(m-1)p-(K-1)m-1}\left[\left(1+\sum_{j=1}^{i}Z_j\right)\left(\sum_{j=i+(m-1)(p-1)+1}^{(m-1)p}Z_j\right)\right]\sigma_a^2 & \text{if}\\ \quad -mp+p+mK>0 \text{ and } -m+1<-mp+p+m(K-1)+1\le 0 & \\ \Leftrightarrow m<\dfrac{p}{p-K} \text{ with } p>K \text{ and } \dfrac{p+1}{p-(K-1)}\le m<\dfrac{p}{p-K} \text{ with } p>K & \\ \Rightarrow \dfrac{p+1}{p-(K-1)}\le m<\dfrac{p}{p-K} \text{ with } p>K & \text{(A.7)}\\ 0 \qquad\qquad \text{if } -mp+p+m(K-1)m+1>0. & \text{(A.8)}\end{cases}$$

Note that $-mp+p+mK>0 \Leftrightarrow m<\frac{p}{p-K}$ and $-m+1\ge -mp+p+m(K-1)+1 \Leftrightarrow m\ge\frac{p}{p-K}$ with $p>K$, which is impossible. Therefore, these conditions are not valid to determine the covariance.

$$E\left[\sum_{i=m}^{(m-1)p}\left(\sum_{j=i-(m-1)}^{i}Z_j\right)B^i a_{mT}\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i}Z_j\right)B^i a_{mT+mK}\right]=0. \quad \text{(A.9)}$$

$$E\left[\sum_{i=m}^{(m-1)p}\left(\sum_{j=i-(m-1)}^{i}Z_j\right)B^i a_{mT}\sum_{i=m}^{(m-1)p}\left(\sum_{j=i-(m-1)}^{i}Z_j\right)B^i a_{mT+mK}\right]$$

$$= \begin{cases} \sum_{i=m}^{(m-1)p-mK} \left[\left(\sum_{j=i-(m-1)}^{i} Z_j \right) \left(\sum_{j=i+m(K-1)+1}^{i+mK} Z_j \right) \right] \sigma_a^2 \quad \text{if} \\ -m+1 > -mp+p+mK \Leftrightarrow m > \dfrac{p-1}{p-(K+1)} \quad \text{with } p > K+1 & \text{(A.10)} \\ 0 \text{ if } -m+1 \leq -mp+p+mK \\ \Leftrightarrow m \leq \dfrac{p-1}{p-(K+1)} \text{ with } p > K+1 \text{ or } m \geq \dfrac{p-1}{p-(K+1)} \quad \text{with } p < K+1. & \text{(A.11)} \end{cases}$$

Note that if $-m+1 > -mp+p+mK \Leftrightarrow m < \frac{p-1}{p-(K+1)}$ with $p < K+1$, then $m \leq 0$ which is impossible. Therefore, this condition is not valid to determine the covariance.

$$E\left[\sum_{i=m}^{(m-1)p} \left(\sum_{j=i-(m-1)}^{i} Z_j \right) B^i a_{mT} \sum_{i=(m-1)p+1}^{(m-1)(p+1)} \left(\sum_{j=i-(m-1)}^{(m-1)p} Z_j \right) B^i a_{mT+mK} \right]$$

$$= \begin{cases} \sum_{i=(m-1)p-mK+1}^{(m-1)p-m(K-1)-1} \left[\left(\sum_{j=i-(m-1)}^{i} Z_j \right) \left(\sum_{j=i+m(K-1)+1}^{(m-1)p} Z_j \right) \right] \sigma_a^2 \quad \text{if} \\ -m+1 \geq -mp+p+mK \Leftrightarrow m \geq \dfrac{p-1}{p-(K+1)} \quad \text{with } p > K+1 & \text{(A.12)} \\ \sum_{i=m}^{(m-1)p-m(K-1)-1} \left[\left(\sum_{j=i-(m-1)}^{i} Z_j \right) 0 \left(\sum_{j=i+(m-1)(p-2)}^{(m-1)p} Z_j \right) \right] \sigma_a^2 \quad \text{if} \\ -m+1 < -mp+p+mK \text{ and } -mp+p+m(K-1)+1 \leq 0 \\ \Leftrightarrow \dfrac{p+1}{p-(K-1)} \leq m < \dfrac{p-1}{p-(K+1)} \quad \text{with } p > K+1 & \text{(A.13)} \\ 0 \text{ if } -mp+p+m(K-1)+1 > 0. & \text{(A.14)} \end{cases}$$

Note that if $-m+1 \geq -mp+p+mK \Leftrightarrow m \leq \frac{p-1}{p-(K+1)}$ with $p < K+1$, then $m \leq 0$, or if $-m+1 < -mp+p+mK$ with $p < K+1$ and $-mp+p+m(K-1)+1 \leq 0$ with $p < K-1$, then $m \leq 0$, which is impossible. Therefore, these conditions are not valid to determine the covariance.

$$E\left[\sum_{i=(m-1)p+1}^{(m-1)(p+1)} \left(\sum_{j=i-(m-1)}^{(m-1)p} Z_j \right) B^i a_{mT} \sum_{i=0}^{m-1} \left(1 + \sum_{j=1}^{i} Z_j \right) B^i a_{mT+mK} \right] = 0. \quad \text{(A.15)}$$

$$E\left[\sum_{i=(m-1)p+1}^{(m-1)(p+1)} \left(\sum_{j=i-(m-1)}^{(m-1)p} Z_j \right) B^i a_{mT} \sum_{i=m}^{(m-1)p} \left(\sum_{j=i-(m-1)}^{i} Z_j \right) B^i a_{mT+mK} \right] = 0. \quad \text{(A.16)}$$

$$E\left[\sum_{i=(m-1)p+1}^{(m-1)(p+1)} \left(\sum_{j=i-(m-1)}^{(m-1)p} Z_j \right) B^i a_{mT} \sum_{i=(m-1)p+1}^{(m-1)(p+1)} \left(\sum_{j=i-(m-1)}^{(m-1)p} Z_j \right) B^i a_{mT+mK} \right] = 0. \quad \text{(A.17)}$$

Then, joining expressions (A.2) and (A.5), we obtain expression (12); expressions (A.2) and (A.6) lead to (13); expressions (A.2) and (A.12) lead to (14); expressions (A.3), (A.10), and (A.12) lead to (15); expressions (A.7) and (A.13) lead to (16); expression (A.13) is (17); finally, expression (A.7) is (18). For all other values of $K > 0$, $\mathrm{Cov}(W_{mT}, W_{mT+mK}) = 0$. □

Proof of Proposition 2.2. Since $p = 2$ in (2) and $m = 2$, from (10a), (10b), and (10c) we obtain $W_{2T} = [1 + (1+\delta_1+\delta_2)B + (\delta_1 + \delta_2 + \delta_1\delta_2)B^2 + \delta_1\delta_2 B^3]a_{2T}$. Since a_t is an independent white noise process with zero mean and variance $\mathrm{Var}(a_t) = \sigma_a^2$, $E(W_{2T}) = 0$ and we can easily determine the autocovariance function of W_{2T}, which is

$$\mathrm{Cov}(W_{2T}, W_{2T+2K}) = E(W_{2T}W_{2T+2K})$$
$$= \begin{cases} [1 + (1 + \delta_1 + \delta_2)^2 + (\delta_1 + \delta_2 + \delta_1\delta_2)^2 + (\delta_1\delta_2)^2]\sigma_a^2 = \beta_1\sigma_a^2 & K = 0 \\ [\delta_1 + \delta_2 + \delta_1\delta_2 + (1 + \delta_1 + \delta_2)\delta_1\delta_2]\,\sigma_a^2 = \beta_2\sigma_a^2 & K = 1 \\ 0 & K \geq 2. \end{cases} \quad (A.18)$$

Thus, W_{2T} is an MA(1) process, $W_{2T} = \varepsilon_T - \Theta_1\varepsilon_{T-1}$. Consequently, $\mathrm{Cov}(W_{2T}, W_{2T+2}) = -\Theta_1\sigma_\varepsilon^2$ and $\mathrm{Var}(W_{2T}) = (1 + \Theta_1^2)\sigma_\varepsilon^2$. Therefore, from expression (A.18), we have $\mathrm{Var}(W_{2T})/\mathrm{Cov}(W_{2T}, W_{2T+2}) = (1 + \Theta_1^2)\sigma_\varepsilon^2/(-\Theta_1\sigma_\varepsilon^2) = \beta_1\sigma_a^2/(\beta_2\sigma_a^2) = \beta_3$, say. Then, Θ_1 is the root of the quadratic equation $\Theta_1^2 + \beta_3\Theta_1 + 1 = 0$ such that $|\Theta_1| < 1$. Furthermore, $\mathrm{Var}(W_{2T}) = (1 + \Theta_1^2)\sigma_\varepsilon^2 = \beta_1\sigma_a^2 \Leftrightarrow \sigma_\varepsilon^2 = \beta_1/(1 + \Theta_1^2)\sigma_a^2$. □

Proof of Proposition 2.3. We have $p = 2$ in (2), (10a), (10b), and (10c) and, since a_t is an independent white noise process with zero mean and variance $\mathrm{Var}(a_t) = \sigma_a^2$, $E(W_{mT}) = 0$ and the autocovariance function of W_{mT} (with $m \geq 3$) is

$$\mathrm{Cov}(W_{mT}, W_{mT+mK}) = E(W_{mT}W_{mT+mK})$$
$$= \begin{cases} \left[\sum_{i=0}^{m-1}\left(1 + \sum_{j=1}^{i} Z_j\right)^2 + \sum_{i=m}^{2(m-1)}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)^2 \right. \\ \left. + \sum_{i=2(m-1)+1}^{3(m-1)}\left(\sum_{j=i-(m-1)}^{2(m-1)} Z_j\right)^2\right]\sigma_a^2 = \beta_1\sigma_a^2 & K = 0 \\ \left[\sum_{i=0}^{m-2}\left(1 + \sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+1}^{i+m} Z_j\right) + \left(1 + \sum_{j=1}^{m-1} Z_j\right)\left(\sum_{j=m}^{2(m-1)} Z_j\right) \right. \\ \left. + \sum_{i=m}^{2(m-1)-1}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+1}^{2(m-1)} Z_j\right)\right]\sigma_a^2 = \beta_2\sigma_a^2 & K = 1 \\ \sum_{i=0}^{m-3}\left(1 + \sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+m+1}^{2(m-1)} Z_j\right)\sigma_a^2 = \beta_3\sigma_a^2 & K = 2 \\ 0 & K \geq 3. \end{cases}$$

(A.19)

Thus, W_{mT} is an MA(2) process, that is, $W_{mT} = \varepsilon_T - \Theta_1\varepsilon_{T-1} - \Theta_2\varepsilon_{T-2}$ and consequently

$$\mathrm{Cov}(W_{mT}, W_{mT+mK}) = \begin{cases} \left(1+\Theta_1^2+\Theta_2^2\right)\sigma_\varepsilon^2 & K=0 \\ (-\Theta_1+\Theta_1\Theta_2)\,\sigma_\varepsilon^2 & K=1 \\ -\Theta_2\sigma_\varepsilon^2 & K=2 \\ 0 & K\geq 3. \end{cases} \tag{A.20}$$

Therefore, from expression (A.19), we have $\mathrm{Cov}(W_{mT}, W_{mT+2m})/\mathrm{Cov}(W_{mT}, W_{mT+m}) =-\Theta_2\sigma_\varepsilon^2/[(-\Theta_1+\Theta_1\Theta_2)\sigma_\varepsilon^2] = \beta_3\sigma_a^2/(\beta_2\sigma_a^2) = \beta_4$, say. Consequently,

$$\frac{-\Theta_2}{-\Theta_1+\Theta_1\Theta_2} = \beta_4 \Leftrightarrow \Theta_2 = \frac{\beta_4\Theta_1}{\beta_4\Theta_1+1}. \tag{A.21}$$

Furthermore, $\mathrm{Var}(W_{mT})/\mathrm{Cov}(W_{mT}, W_{mT+m}) = (1+\Theta_1^2+\Theta_2^2)\sigma_\varepsilon^2/[(-\Theta_1+\Theta_1\Theta_2)\sigma_\varepsilon^2] = \beta_1\sigma_a^2/(\beta_2\sigma_a^2) = \beta_5$, say. Thus, $(1+\Theta_1^2+\Theta_2^2)/(-\Theta_1+\Theta_1\Theta_2) = \beta_5 \Leftrightarrow \beta_4^2\Theta_1^4 + 2\beta_4\Theta_1^3 + (1+2\beta_4^2+\beta_4\beta_5)\Theta_1^2 + (2\beta_4+\beta_5)\Theta_1 + 1 = 0$, using (A.21). Moreover, we also have $\mathrm{Var}(W_{mT})/\mathrm{Cov}(W_{mT}, W_{mT+2m}) = (1+\Theta_1^2+\Theta_2^2)\sigma_\varepsilon^2/(-\Theta_2\sigma_\varepsilon^2) = \beta_1\sigma_a^2/(\beta_3\sigma_a^2) = \beta_6$, say. Consequently, $(1+\Theta_1^2+\Theta_2^2)/(-\Theta_2) = \beta_6 \Leftrightarrow \beta_4^2\Theta_1^4 + 2\beta_4\Theta_1^3 + (1+2\beta_4^2+\beta_4^2\beta_6)\Theta_1^2 + (\beta_6+2)\beta_4\Theta_1 + 1 = 0$ using (A.21). We then conclude that there may be more than one solution for Θ_1 (and consequently for Θ_2) but the solutions retained have to be such that $|\Theta_2| < 1$, $\Theta_2 - \Theta_1 < 1$ and $\Theta_2 + \Theta_1 < 1$ to ensure invertibility of the aggregate model. Furthermore, $\mathrm{Var}(W_{mT}) = (1+\Theta_1^2+\Theta_2^2)\sigma_\varepsilon^2 = \beta_1\sigma_a^2 \Leftrightarrow \sigma_\varepsilon^2 = \beta_1\sigma_a^2/(1+\Theta_1^2+\Theta_2^2)$. □

Proof of Proposition 2.4. We have $p = 3$ in (2) and, from (10a), (10b), and (10c) with $m = 2$, since a_t is an independent white noise process with zero mean and variance $\mathrm{Var}(a_t) = \sigma_a^2$, $E(W_{2T}) = 0$ and the autocovariance function of W_{2T} is

$$\mathrm{Cov}\left(W_{2T}, W_{2T+2K}\right) = E\left(W_{2T}W_{2T+2K}\right)$$

$$= \begin{cases} \left[\sum_{i=0}^{1}\left(1+\sum_{j=1}^{i}Z_j\right)^2 + \sum_{i=2}^{3}\left(\sum_{j=i-1}^{i}Z_j\right)^2 + Z_3^2\right]\sigma_a^2 = \beta_1\sigma_a^2 & K=0 \\ \left[\sum_{j=1}^{2}Z_j + (1+Z_1)\left(\sum_{j=2}^{3}Z_j\right) + \left(\sum_{j=1}^{2}Z_j\right)Z_3\right]\sigma_a^2 = \beta_2\sigma_a^2 & K=1 \\ Z_3\sigma_a^2 = \beta_3\sigma_a^2 & K=2 \\ 0 & K\geq 3. \end{cases} \tag{A.22}$$

Thus, W_{2T} is an MA(2) process, that is, $W_{2T} = \varepsilon_T - \Theta_1\varepsilon_{T-1} - \Theta_2\varepsilon_{T-2}$ and consequently $\mathrm{Cov}\left(W_{2T}, W_{2T+2K}\right)$ is given by (A.20) in Proposition 2.3. Therefore, Θ_1, Θ_2, and σ_ε^2 are determined as in that proposition with β_1, β_2, and β_3 given by (A.22). □

Proof of Proposition 2.5. We have $p = 3$ in (2) and, from (10a), (10b), and (10c) with $m = 3$, since a_t is an independent white noise process with zero mean and variance $\mathrm{Var}(a_t) = \sigma_a^2$, $E\left(W_{3T}\right) = 0$ and the autocovariance function of W_{3T} is

$$\mathrm{Cov}\,(W_{3T}, W_{3T+3K}) = E\,(W_{3T} W_{3T+3K})$$
$$= \begin{cases} \left[\sum_{i=0}^{2}\left(1+\sum_{j=1}^{i} Z_j\right)^2 + \sum_{i=3}^{6}\left(\sum_{j=i-2}^{i} Z_j\right)^2 + \sum_{i=7}^{8}\left(\sum_{j=i-2}^{6} Z_j\right)^2\right]\sigma_a^2 = \beta_1\sigma_a^2 & K=0 \\ \left[\sum_{i=0}^{2}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+1}^{i+3} Z_j\right) + \left(\sum_{j=1}^{3} Z_j\right)\left(\sum_{j=4}^{6} Z_j\right) \right. & \\ \left. + \sum_{i=4}^{5}\left(\sum_{j=i-2}^{i} Z_j\right)\left(\sum_{j=i+1}^{6} Z_j\right)\right]\sigma_a^2 \beta_2\sigma_a^2 & K=1 \\ \left[\sum_{j=4}^{6} Z_j + \sum_{i=1}^{2}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+4}^{6} Z_j\right)\right]\sigma_a^2\beta_3\sigma_a^2 & K=2 \\ 0 & K\geq 3. \end{cases} \tag{A.23}$$

Thus, W_{3T} is an MA(2) process, that is, $W_{3T} = \varepsilon_T - \Theta_1\varepsilon_{T-1} - \Theta_2\varepsilon_{T-2}$ and consequently $\mathrm{Cov}\,(W_{3T}, W_{3T+3K})$ is given by (A.20) in Proposition 2.3. Therefore, Θ_1, Θ_2, and σ_ε^2 are determined as in that proposition with β_1, β_2, and β_3 given by (A.23). □

Proof of Proposition 2.6. We have $p = 3$ in (2), (10a), (10b), and (10c) and, since a_t is an independent white noise process with zero mean and variance $\mathrm{Var}(a_t) = \sigma_a^2$, $E(W_{mT}) = 0$ and the autocovariance function of W_{mT} (with $m \geq 4$) is

$$\mathrm{Cov}(W_{mT}, W_{mT+mK}) = E(W_{mT} W_{mT+mK})$$
$$= \begin{cases} \left[\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i} Z_j\right)^2 + \sum_{i=m}^{3(m-1)}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)^2 \right. & \\ \left. + \sum_{i=(m-1)p+1}^{4(m-1)}\left(\sum_{j=i-(m-1)}^{3(m-1)} Z_j\right)^2\right]\sigma_a^2 = \beta_1\sigma_a^2 & K=0 \\ \left[\sum_{i=0}^{m-1}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+1}^{i+m} Z_j\right) + \sum_{i=m}^{2(m-1)-1}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+1}^{i+m} Z_j\right) \right. & \\ \left. + \sum_{i=2(m-1)}^{3(m-1)-1}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+1}^{3(m-1)} Z_j\right)\right]\sigma_a^2 = \beta_2\sigma_a^2 & K=1 \\ \left[\sum_{i=0}^{m-3}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+m+1}^{i+2m} Z_j\right) + \sum_{i=m-2}^{m-1}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+m+1}^{3(m-1)} Z_j\right) \right. & \\ \left. + \sum_{i=m}^{2(m-2)}\left(\sum_{j=i-(m-1)}^{i} Z_j\right)\left(\sum_{j=i+m+1}^{3(m-1)} Z_j\right)\right]\sigma_a^2 = \beta_3\sigma_a^2 & K=2 \\ \left[\sum_{i=0}^{m-4}\left(1+\sum_{j=1}^{i} Z_j\right)\left(\sum_{j=i+2m+1}^{3(m-1)} Z_j\right)\right]\sigma_a^2 = \beta_4\sigma_a^2 & K=3 \\ 0 & K\geq 4. \end{cases} \tag{A.24}$$

Thus, W_{mT} is an MA(3) process, that is, $W_{mT} = \varepsilon_T - \Theta_1\varepsilon_{T-1} - \Theta_2\varepsilon_{T-2} - \Theta_3\varepsilon_{T-3}$ and consequently

$$\mathrm{Cov}(W_{mT}, W_{mT+mK}) = \begin{cases} \left(1+\Theta_1^2+\Theta_2^2+\Theta_3^2\right)\sigma_\varepsilon^2 & K=0 \\ (-\Theta_1+\Theta_1\Theta_2+\Theta_2\Theta_3)\,\sigma_\varepsilon^2 & K=1 \\ (-\Theta_2+\Theta_1\Theta_3)\,\sigma_\varepsilon^2 & K=2 \\ -\Theta_3\sigma_\varepsilon^2 & K=3 \\ 0 & K\geq 4. \end{cases} \tag{A.25}$$

Therefore, from (A.24), we have $\mathrm{Cov}(W_{mT}, W_{mT+2m})/\mathrm{Cov}(W_{mT}, W_{mT+m}) = (-\Theta_2 + \Theta_1\Theta_3)\sigma_\varepsilon^2/[(-\Theta_1 + \Theta_1\Theta_2 + \Theta_2\Theta_3)\sigma_\varepsilon^2] = \beta_3\sigma_a^2/(\beta_2\sigma_a^2) = \beta_5$, say. Consequently,

$$\frac{-\Theta_2+\Theta_1\Theta_3}{-\Theta_1+\Theta_1\Theta_2+\Theta_2\Theta_3} = \beta_5 \Leftrightarrow \Theta_3 = \frac{\Theta_2-(1-\Theta_2)\beta_5\Theta_1}{\Theta_1-\beta_5\Theta_2} \tag{A.26}$$

which expresses Θ_3 as a function of Θ_1 and Θ_2. Moreover, from (A.24), we also have $\mathrm{Cov}(W_{mT}, W_{mT+3m})/\mathrm{Cov}(W_{mT}, W_{mT+m}) = -\Theta_3\sigma_\varepsilon^2/[(-\Theta_1 + \Theta_1\Theta_2 + \Theta_2\Theta_3)\sigma_\varepsilon^2] = \beta_4\sigma_a^2/(\beta_2\sigma_a^2) = \beta_6$, say. Thus,

$$\frac{-\Theta_3}{-\Theta_1+\Theta_1\Theta_2+\Theta_2\Theta_3} = \beta_6 \Leftrightarrow \Theta_3 = \frac{\beta_6\Theta_1(1-\Theta_2)}{\beta_6\Theta_2+1}, \tag{A.27}$$

which provides an alternative expression for Θ_3. Furthermore, from (A.24), we also have $\mathrm{Cov}(W_{mT}, W_{mT+3m})/\mathrm{Cov}(W_{mT}, W_{mT+2m}) = -\Theta_3\sigma_\varepsilon^2/[(-\Theta_2 + \Theta_1\Theta_3)\sigma_\varepsilon^2] = \beta_4\sigma_a^2/(\beta_3\sigma_a^2) = \beta_7$, say. Consequently,

$$\frac{-\Theta_3}{-\Theta_2+\Theta_1\Theta_3} = \beta_7 \Leftrightarrow \Theta_3 = \frac{\beta_7\Theta_2}{\beta_7\Theta_1+1}, \tag{A.28}$$

which provides another alternative expression for Θ_3. Therefore, from (A.26) and (A.27), we have

$$\frac{\Theta_2-(1-\Theta_2)\beta_5\Theta_1}{\Theta_1-\beta_5\Theta_2} = \frac{\beta_6\Theta_1(1-\Theta_2)}{\beta_6\Theta_2+1} \Leftrightarrow \beta_6\Theta_2^2 + (\beta_6\Theta_1^2+\beta_5\Theta_1+1)\Theta_2 - (\beta_6\Theta_1^2+\beta_5\Theta_1) = 0$$

$$\Rightarrow \Theta_2 = \frac{-(\beta_6\Theta_1^2+\beta_5\Theta_1+1) \pm \sqrt{(\beta_6\Theta_1^2+\beta_5\Theta_1+1)^2 + 4\beta_6(\beta_6\Theta_1^2+\beta_5\Theta_1)}}{2\beta_6}. \tag{A.29}$$

Using this solution in (A.26) or (A.27), Θ_3 may be written as a function of Θ_1 only. Alternatively, from (A.26) and (A.28), we have

$$\frac{\Theta_2-(1-\Theta_2)\beta_5\Theta_1}{\Theta_1-\beta_5\Theta_2} = \frac{\beta_7\Theta_2}{\beta_7\Theta_1+1} \Leftrightarrow \beta_5\beta_7\Theta_2^2 + (\beta_5\beta_7\Theta_1^2+\beta_5\Theta_1+1)\Theta_2 - (\beta_7\Theta_1+1)\beta_5\Theta_1 = 0$$

$$\Rightarrow \Theta_2 = \frac{-(\beta_5\beta_7\Theta_1^2+\beta_5\Theta_1+1) \pm \sqrt{(\beta_5\beta_7\Theta_1^2+\beta_5\Theta_1+1)^2 + 4\beta_5\beta_7(\beta_7\Theta_1+1)\beta_5\Theta_1}}{2\beta_5\beta_7}. \tag{A.30}$$

Using this solution in (A.26) or (A.28), Θ_3 may again be written as a function of Θ_1 only. Alternatively, from (A.27) and (A.28), we have

$$\frac{\beta_6\Theta_1(1-\Theta_2)}{\beta_6\Theta_2+1} = \frac{\beta_7\Theta_2}{\beta_7\Theta_1+1} \Leftrightarrow \beta_6\beta_7\Theta_2^2 + (\beta_6\beta_7\Theta_1^2+\beta_6\Theta_1+\beta_7)\Theta_2 - (\beta_7\Theta_1+1)\beta_6\Theta_1 = 0$$

$$\Rightarrow \Theta_2 = \frac{-(\beta_6\beta_7\Theta_1^2+\beta_6\Theta_1+\beta_7) \pm \sqrt{(\beta_6\beta_7\Theta_1^2+\beta_6\Theta_1+\beta_7)^2 + 4\beta_6\beta_7(\beta_7\Theta_1+1)\beta_6\Theta_1}}{2\beta_6\beta_7}. \tag{A.31}$$

Using this solution in (A.27) or (A.28), Θ_3 may again be written as a function of Θ_1 only. Furthermore, $\mathrm{Var}(W_{mT})/\mathrm{Cov}(W_{mT}, W_{mT+m}) = (1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2)\sigma_\varepsilon^2/[(-\Theta_1 + \Theta_1\Theta_2 + \Theta_2\Theta_3)\sigma_\varepsilon^2] = \beta_1\sigma_a^2/(\beta_2\sigma_a^2) = \beta_8$, say. Thus,

$$\frac{1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2}{-\Theta_1 + \Theta_1\Theta_2 + \Theta_2\Theta_3} = \beta_8 \Leftrightarrow \Theta_1^2 + \Theta_2^2 + \Theta_3^2 - \beta_8\Theta_1(\Theta_2 - 1) - \beta_8\Theta_2\Theta_3 + 1 = 0, \tag{A.32}$$

which will lead to possible solutions for Θ_1, writing Θ_3 and Θ_2 as functions of Θ_1 from (A.26), (A.27) or (A.28) and (A.29), (A.30) or (A.31), respectively. Then, Θ_3 and Θ_2 are computed from these expressions. Moreover, we also have $\mathrm{Var}(W_{mT}) / \mathrm{Cov}(W_{mT}, W_{mT+2m}) = (1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2)\sigma_\varepsilon^2/[(-\Theta_2 + \Theta_1\Theta_3)\sigma_\varepsilon^2)] = \beta_1\sigma_a^2/(\beta_3\sigma_a^2) = \beta_9$, say. Consequently,

$$\frac{1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2}{-\Theta_2 + \Theta_1\Theta_3} = \beta_9 \Leftrightarrow \Theta_1^2 + \Theta_2^2 + \Theta_3^2 - \beta_9(\Theta_1\Theta_3 - \Theta_2) + 1 = 0, \tag{A.33}$$

which provides an alternative equation to find Θ_1, writing again Θ_3 and Θ_2 as functions of Θ_1 from (A.26), (A.27), or (A.28), and (A.29), (A.30), or (A.31), respectively. Furthermore, we also have $\mathrm{Var}(W_{mT})/\mathrm{Cov}(W_{mT}, W_{mT+3m}) = (1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2)\sigma_\varepsilon^2/(-\Theta_3\sigma_\varepsilon^2) = \beta_1\sigma_a^2/(\beta_4\sigma_a^2) = \beta_{10}$, say. Therefore,

$$\frac{1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2}{-\Theta_3} = \beta_{10} \Leftrightarrow \Theta_1^2 + \Theta_2^2 + \Theta_3^2 + \beta_{10}\Theta_3 + 1 = 0, \tag{A.34}$$

which provides another alternative equation to find Θ_1 similarly to (A.32) and (A.33). To ensure invertibility of the aggregate model, the solutions retained have to be such that the roots of the moving average polynomial $(1 - \Theta_1\mathcal{B} - \Theta_2\mathcal{B}^2 - \Theta_3\mathcal{B}^3)$ are outside the unit circle, that is, $|\mathcal{B}_i| > 1$ $(i = 1, 2, 3)$. Finally, $\mathrm{Var}(W_{mT}) = (1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2)\sigma_\varepsilon^2 = \beta_1\sigma_a^2 \Leftrightarrow \sigma_\varepsilon^2 = \beta_1\sigma_a^2/(1 + \Theta_1^2 + \Theta_2^2 + \Theta_3^2)$. We note that it is not possible to find analytical solutions for Θ_1 in the above nonlinear equations (A.32), (A.33), or (A.34) and consequently they have to be solved numerically. □

Acknowledgments

The authors gratefully acknowledge the helpful comments and suggestions of two anonymous referees that significantly contributed to improve this article.

ORCID

Paulo S. A. Sousa http://orcid.org/0000-0002-0578-1593

References

Abraham, B. (1982). Temporal aggregation and time series. *International Statistical Review* 50:285–291.

Ahsanullah, M., Wei, W. W. S. (1984). Effects of temporal aggregation on forecasts in an ARMA (1, 1) process. *American Statistical Association Proceedings of the Business and Economics Statistics Section*. pp. 297–302.

Amemiya, T., Wu, R. Y. (1972). The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association* 67:628–632.

Brewer, K. R. W. (1973). Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models. *Journal of Econometrics* 1:133–154.

Drost, F. C. (1994). Temporal aggregation of time-series. In: Kaehler, J., Kugler, P., eds. *Econometric Analysis of Financial Markets*. Heidelberg: Physica, pp. 11–21.

Hotta, L. K., Neto, J. C. (1993). The effect of aggregation on prediction in autoregressive integrated moving-average models. *Journal of Time Series Analysis* 14:261–269.

Koreisha, S. G., Fang, Y. (2004). Updating ARMA predictions for temporal aggregates. *Journal of Forecasting* 23:275–296.

Lütkepohl, H. (2009). *Forecasting Aggregated Time Series Variables: A Survey*. Working Paper (ECO 20019/17). Florence: European University Institute, Department of Economics.

Marcellino, M. (1999). Some consequences of temporal aggregation in empirical analysis. *Journal of Business and Economic Statistics* 17:129–136.

Silvestrini, A., Veredas, D. (2008). Temporal aggregation of univariate and multivariate time series models: A survey. *Journal of Economic Surveys* 22:458–497.

Stram, D. O., Wei, W. W. S. (1986). Temporal aggregation in the ARIMA process. *Journal of Time Series Analysis* 7:279–292.

Tiao, G. C. (1972). Asymptotic behavior of temporal aggregates of time series. *Biometrika* 59:525–531.

Wei, W. W. S. (1978). Some consequences of temporal aggregation in seasonal time-series models. In: Zellner, A., ed. *Seasonal Analysis of Economic Time Series*. Washington DC: U.S. Department of Commerce, Bureau of the Census, pp. 433–444.

Wei, W. W. S. (2006). *Time Series Analysis: Univariate and Multivariate Methods*. 2nd ed. New York: Pearson.

Weiss, A. A. (1984). Systematic sampling and temporal aggregation in time series models. *Journal of Econometrics* 26:271–281.