Journal of Forecasting
J. Forecast. **23**, 275–296 (2004)
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/for.913

Updating ARMA Predictions for Temporal Aggregates

SERGIO G. KOREISHA[1]* AND YUE FANG[1,2]
[1] *University of Oregon, USA*
[2] *China Europe International Business School, China*

ABSTRACT

This article develops and extends previous investigations on the temporal aggregation of ARMA predications. Given a basic ARMA model for disaggregated data, two sets of predictors may be constructed for future temporal aggregates: predictions based on models utilizing aggregated data or on models constructed from disaggregated data for which forecasts are updated as soon as the new information becomes available. We show that considerable gains in efficiency based on mean-square-error-type criteria can be obtained for short-term predications when using models based on updated disaggregated data. However, as the prediction horizon increases, the gain in using updated disaggregated data diminishes substantially. In addition to theoretical results associated with forecast efficiency of ARMA models, we also illustrate our findings with two well-known time series. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS ARMA processes; seasonal ARMA processes; nonstationary processes; temporal aggregation; prediction

INTRODUCTION

The effects of temporal aggregation of time series on predictions have been studied extensively over the past two or three decades. Amemiya and Wu (1972) investigated the effect of temporal aggregation on prediction when the disaggregated series followed an autoregressive (AR) process and the data were aggregated over m nonoverlapping periods. They presented numerical computations for the efficiency loss on prediction due to the temporal aggregation for AR(1) and AR(2) processes. They also showed that the ratio of the mean square error of the predictors based on disaggregated and aggregated series, ρ, tended to unity when m increased. Tiao (1972) studied integrated moving average (IMA(d,q)) processes, and obtained similar results to those of Amemiya and Wu. In addition, Tiao showed that the ratio ρ was close to unity for any forecast horizon when the process was stationary ($d = 0$). However, when the process was nonstationary ($d > 0$), substantial gains in forecasting efficiency could be obtained by using disaggregated data, especially for short-term forecasts. Wei (1978) extended these studies to include seasonal autoregressive integrated moving average (SARIMA) models. He showed that the loss in forecasting efficiency through aggregation could be

* Correspondence to: Sergio Koreisha, Lundquist College of Business, University of Oregon, Eugene, OR 97403, USA. E-mail: seigiok@oregon.uoregon.edu

substantial if the nonseasonal component of the disaggregated series was nonstationary. This loss in efficiency was less serious for long-term forecasts. Furthermore, he showed that there was no loss in efficiency due to aggregation if the disaggregated series followed a purely seasonal process. Lütkepohl (1984, 1987) and Wei (1990) extended these results to multivariate cases.

Although theory predicts that the disaggregated series will always provide more accurate forecasts than those based on aggregated data, several empirical studies have shown that there are circumstances in which predictions from aggregated data may yield more accurate forecasts than those derived from disaggregated data. For example, Butter (1976), using monthly and quarterly observations of the difference between the yield on mortgages and the yield on government loans from 1961 to 1974, generated quarterly forecasts based on quarterly models and on the aggregation of the three monthly predictions. His study indicated that the quarterly models in most cases were superior to the monthly models if the forecast horizon was more than one quarter. Similar findings have been reported in Nijman and Palm (1990).

In this article we examine the forecast accuracy of models over different time horizons based on aggregated data, and on disaggregated data which are immediately updated as the new information becomes available. Assuming that the disaggegated data are generated by an ARMA model, we will show how new monthly observations available during the current quarter can be used to significantly improve the performance of quarterly forecasts for short-term predictions. We will also show that there is not much to be gained in using (updated) monthly data instead of quarterly data for long-term forecasts.

The article is organized as follows. In the next section we provide a brief overview of models for aggregated and disaggregated data, and derive the minimum mean square prediction error (MMSE) of temporally aggregated prediction horizons for general ARMA processes based on aggregated series which are updated regularly, and on disaggregated series which are updated within the time span of the forecast interval as soon as new information becomes available. In the following section we compare the relative impact of aggregation on MMSE predictions for several ARMA parametrizations (with and without seasonal components) and forecast horizons. Then we extend our findings to some nonstationary models. Finally, we conclude with some general comments and offer some directions for future work. Proofs of theorems are outlined in the Appendix.

THE AGGREGATED AND DISAGGREGATED SERIES

Models and predictors

Suppose the disaggregated time series x_t follows a stationary and invertible ARMA process, $\Phi(B)x_t = \theta(B)a_t$, where $\Phi(B)$ and $\theta(B)$ are finite polynomials in the back-shift operator B such that $B^j w_t = w_{t-j}$ and $\{a_t\}$ is a white noise process with variance σ_a^2. Such a stochastic process can also be expressed as autoregressive or moving average processes respectively, i.e., $\Pi(B)x_t = a_t$ or $x_t = \Psi(B)a_t$, where $\Pi(B) = \theta^{-1}(B)\Phi(B)$ and $\Psi(B) = \Phi^{-1}(B)\theta(B)$. Now consider the time series X_T, which is the m-period nonoverlapping aggregate of x_t, namely

$$X_T = \sum_{t=m(T-1)+1}^{mT} x_t = (1 + B + B^2 + \ldots + B^{m-1})x_{mT} \tag{1}$$

This aggregated time series X_T follows and ARMA(p, $[p + 1 + (q - p - 1)/m]$) (Brewer, 1973; Harvey, 1981), where $[k]$ denotes the integer part of k. The orders of the aggregate series given by p and

$[p + 1 + (q - p - 1)/m]$ are only maximum since some cancellations may occur (Stram and Wei, 1986). We will refer to X_T as the aggregated series generated from x_t.

Suppose that $\{x_t\}$ is available and hence so is $\{X_T\}$. Two different predictions of X_{T+L} (L is a positive integer) can be obtained: one is the sum of m predictions based on the disaggregated x_t and the other one generated directly from models based on the aggregated series X_T.

The MMSE l-step-ahead prediction of x_t, based on a weighted average of previous observations and the forecasts made at previous lead times from the same origin, is defined in terms of the conditional expectation, $\hat{x}_t(l) = E(x_{t+l}|x_t, x_{t-1}, \ldots)$. If we restrict ourselves to the class of linear predictors, then $\hat{x}_t(l) = E(x_{t+l}|x_t, x_{t-1}, \ldots) = \Sigma_{i=-\infty}^{t}\pi_{i,x}^{l}x_i$, where $\pi_{i,x}^{l}$s are the weights estimated from the data which will yield the MMSE for $\hat{x}_t(l)$. A natural L-step-ahead prediction of X_{T+L} based on x_t, denoted as $\hat{\hat{x}}_{T,0}(L)$, can be derived by summing m forecasts $\hat{x}_t(l)$s:

$$\hat{\hat{x}}_{T,0}(L) \equiv \sum_{i=1}^{m} \hat{x}_{mT}(m(L-1)+i) \tag{2}$$

When new disaggregated observations become available, the predictor in (2) can be improved, i.e.

$$\hat{\hat{x}}_{T,k}(L) \equiv \begin{cases} \sum_{j=1}^{k} x_{mT+j} + \sum_{i=k+1}^{m} \hat{x}_{mT}(m(L-1)+i), & \text{if } L=1 \\ \sum_{i=I}^{m} \hat{x}_{mT+k}(m(L-1)+i-k), & \text{if } L>1 \end{cases} \tag{3}$$

where $k(1 \le k < m)$ is the number of disaggregated observations available after the last aggregated observation. When $k = 0$, $\hat{\hat{x}}_{T,k}(L)$ becomes $\hat{\hat{x}}_{T,0}(L)$ in (2).

The predictor $\hat{\hat{x}}_{T,0}(L)$ yields the linear unbiased MMSE forecast of X_{T+L} given $\{x_{mT}, x_{mT-1}, x_{mT-2}, \ldots\}$ (Box and Jenkins, 1976, p. 128; Pino *et al*., 1987). Following similar arguments we have:

Theorem 1 *For* $0 \le k < m$, $\hat{\hat{x}}_{T,k}(L) = E(X_{T+L}|x_{mT+k}, x_{mT+k-1}, \ldots)$. *That is,* $\hat{\hat{x}}_{T,k}(L)$ *is the linear unbiased MMSE forecast of* X_{T+L} *given* $\{x_{mT+k}, x_{mT+k-1}, \ldots\}$.

The linear unbiased MMSE predictor of X_{T+L}, which is based on X_T, is

$$\hat{X}_T(L) = E(X_{T+L}|X_T, X_{T-1}, \ldots) = \sum_{i=-\infty}^{T} \pi_{i,X}^{L} X_i \tag{4}$$

where $\pi_{i,X}^{L}$s are the weights estimated from the data which will yield the MMSE for $\hat{X}_T(L)$.

Minimum mean square errors

To derive MMSE for general ARMA processes, we first rewrite the disaggregated series $\{x_t\}$ as a moving average process of infinite order (MA(∞)):

$$x_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \tag{5}$$

where $\Sigma_{j=0}^{\infty}\psi_j^2 < \infty$ (i.e., the sequence $\{\psi_j\}$ is square summable). Then, we have:

Theorem 2 *For* $0 \le k < m$,

$$\text{MMSE}(\hat{\hat{x}}_{T,k}(L)) = \begin{cases} \sigma_\varepsilon^2 \sum_{i=0}^{m-k-1} \left(\sum_{j=0}^{i} \psi_j\right)^2, & \text{if } L = 1 \\ \sigma_\varepsilon^2 \left[\sum_{i=0}^{m-1} \left(\sum_{j=0}^{i} \psi_j\right)^2 + \sum_{i=1}^{m(L-1)-k} \left(\sum_{j=0}^{m-1} \psi_{i+j}\right)^2\right], & \text{if } L > 1 \end{cases} \tag{6}$$

From (6) it is evident that MMSE($\hat{\hat{x}}_{T,k}(L)$) depends not only on the magnitude of the ψ weights, but also on their signs. Since MMSE($\hat{\hat{x}}_{T,k}(L)$) consists essentially of summations of several non-negative terms which are expressed as squares of ($\Sigma_{j=0}^{v}\psi_j$), if all the ψ's share the same sign, then there will be no cancellations when the summations are performed. However, if some of the ψ_j's alternate in sign, then cancellations will surely occur, thus impacting the magnitude of MMSE($\hat{\hat{x}}_{T,k}(L)$) as well as that of MMSE($\hat{X}_T(L)$), the MMSE based on aggregated series. [See the next section for further discussion on how the magnitude and sign of the ψ weights affect the measurement of the relative impact of aggregation on MMSE prediction (defined later in this section).]

From Theorem 2, it is easy to obtain the following results:

Corollary 1 *For* $0 \le k_1 \le k_2 < m$,

$$\text{MMSE}\left(\hat{\hat{x}}_{T,k_1}(L)\right) \ge \text{MMSE}\left(\hat{\hat{x}}_{T,k_2}(L)\right)$$

Corollary 2 *For* $0 < k < m$,

$$\frac{\text{MMSE}\left(\hat{\hat{x}}_{T,k}(L)\right)}{\text{MMSE}\left(\hat{\hat{x}}_{T,0}(L)\right)} \to 1$$

as $L \to \infty$.

The MMSE of $\hat{X}_T(L)$ can be obtained in the traditional way for ARMA processes although the series is sampled in a less frequent time scale. The MMSE of $\hat{X}_T(L)$ for X_{T+L} is simply MMSE($\hat{X}_T(L)$) $\equiv E(\hat{X}_T(L) - X_{T+L})^2$. It can be evaluated from the relation

$$\text{MMSE}(\hat{X}_T(L)) = \sigma_{a*}^2 \sum_{i=-\infty}^{L-1} \Psi_i^2 \tag{7}$$

where σ_{a*}^2 is the variance of the white noise term of the X_T process and the Ψ_i's are the coefficients of the MA(∞) representation of X_T. Both σ_{a*}^2 and the Ψ_i's are determined by the parameters of the disaggregated series in a complicated way. General expressions for simple ARIMA models such as AR(1), MA(1) and ARIMA(0, 1, 1) $\times$ SARIMA$(0, 1, 1)_{12}$ can easily be derived (see later sections as well as Amemiya and Wu, 1972; Nijman and Palm, 1990).

It is also possible to evaluate MMSE($\hat{X}_T(L)$) if the autocovariance structures of the X_T process are known. Let $\mathbf{X}_T = (X_T, X_{T-1}, \ldots, X_1)'$ and $R_{T,X}^L = (\gamma_X(L), \gamma_X(1 + L), \ldots, \gamma_X(T - 1 + L))'$, where $\gamma_X(.)$ is the autocovariance function of X_T and $\Gamma_{T,X} = [\gamma_X(i - j)]_{i,j=1,2,\ldots T}$. Then, the MMSE of $\hat{X}_T(L)$ is

$$E(\hat{X}_T(L) - X_{T+L})^2 = \gamma_{X_T}(0) - \mathcal{R}_{T,X}^{L\prime}\Gamma_{T,X}^{-1}\mathcal{R}_{T,X}^L \tag{8}$$

The autocovariance function of aggregated series $\{X_T\}$ can be obtained in terms of the autocovariance function of the disaggregated series $\{x_t\}$. The following results from Stram and Wei (1986) are useful in building that connection. If $m \geq 2$ and $k \geq 1$, the kth autocovariance of X_T is

$$\gamma_{X_T}(k) = (1 + B + B^2 + \ldots + B^{m-1})^2 \gamma_{x_t}(mk + m - 1)$$

where $B\gamma_{xt}(i) = \gamma_{xt}(i - 1)$. Although (7) is based on the knowledge of X_T's going all the way back to the infinite past, (8) utilizes only the finite distant past.

In general, prediction from an aggregated series is less efficient than prediction from a disaggregated series. In fact, we have the following results:

Theorem 3 *For* $0 \leq k < m$,

$$\text{MMSE}(\hat{\hat{x}}_{T,k}(L)) \leq \text{MMSE}(\hat{X}_T(L))$$

Note that Theorem 3 holds under the assumption that the series $\{x_t\}$ (and hence the aggregated series $\{X_T\}$) is correctly identified. If $\{x_t\}$ is misspecified, then the inequality in Theorem 3 need not hold. We demonstrate this later, when we examine Butter's (1976) series.

Denote the difference between $\text{MMSE}(\hat{\hat{x}}_{T,k}(L))$ (where $0 \leq k < m$) and $\text{MMSE}(\hat{X}_T(L))$ as ΔMMSE, that is

$$\Delta\text{MMSE} = \text{MMSE}(\hat{X}_T(L)) - \text{MMSE}(\hat{\hat{x}}_{T,k}(L))$$

Consequently,

$$\nabla\text{MMSE} \equiv \frac{\Delta\text{MMSE}}{\text{MMSE}(\hat{X}_T(L))} = 1 - \frac{\text{MMSE}(\hat{\hat{x}}_{T,k}(L))}{\text{MMSE}(\hat{X}_T(L))} \tag{9}$$

is a measurement of the relative impact of aggregation on MMSE predictions. More specifically, ∇MMSE is the percentage improvement of MMSE using disaggregated data over the MMSE based on aggregated data.

The magnitude of the gain in efficiency for prediction from a disaggregated series, measured by either ΔMMSE or ∇MMSE, can be evaluated using (7) (or (8)) and (6). Since both $\text{MMSE}(X_T(L))$ and $\text{MMSE}(\hat{\hat{x}}_{T,k}(L))$ can be expressed explicitly in either $\{\psi_i\}$ or $\{\Psi_i\}$, we derive explicit expressions for ∇MMSE for AR(1) and MA(q) processes. As will be shown, there is not much insight to be gained from similar expressions for general ARMA(p,q) models.

From Theorem 3, we can also establish the bounds $0 \leq \nabla\text{MMSE} \leq 1$ for all m and L. When m becomes large, X_T tends to white noise, and in this case, there is no gain in using the disaggregated data when long-term predictions are made. Applying the result of Corollary 2, the limit holds also for $\hat{\hat{x}}_{T,k}(L)$ for all $k = 0, 1, 2, \ldots, m - 1$. That is,

Corollary 3 *For sufficiently large m,*

$$\nabla\text{MMSE} \to 0$$

as $L \to \infty$.

Corollary 3 generalizes the results in Tiao (1972) where he considered just the case $k = 0$, and showed that $\mathrm{MMSE}(\hat{\hat{x}}_{T,0}(L)) / \mathrm{MMSE}(\hat{X}_T(L)) \to 1$ as $L \to \infty$.

We would like to point out, however, that the results of Corollary 3 appear to hold even for finite m, as will be demonstrated in the following sections.

Note that when we derived the above results we assumed that the model parameters were known. In practice, it is necessary to replace these parameters with their estimates. The effect of estimation errors on the variance of prediction errors has been evaluated in Butter (1976) for the AR(1) model. The bias is small in comparison to the variance of the disturbances. It is very complicated to estimate the effect of estimation errors on the variance of prediction errors for other models such as ARMA(1,1). However, as pointed out by Box and Jenkins (1976), for model parameter estimates based on series of moderate length, the effect of such estimation errors is small and may be negligible. Hence, if all parameters are to be estimated consistently, those results will remain valid in a certain asymptotic sense.

PREDICTIONS BASED ON ARMA MODELS

∇MMSE's generated from ARMA parametrizations without seasonal components

In this section, we derive exact values for ∇MMSE associated with several ARMA models. We have chosen model structures and parametrizations of $\{x_t\}$ which not only conform to other previously published studies, such as Pukkila *et al*. (1990) and Koreisha and Fang (1999, 2001), but which also cover a wide spectrum of parameter values ranging from well-defined stationary and invertible processes to nearly noninvertible or nonstationary processes. Without loss of generality, we will assume that the variance of $\{a_t\}$, σ_a^2, is one.

Tables I–V contain ∇MMSE's for selected ARMA(p,q) structures for several prediction horizons, L, and various levels of m and k ($0 \le k < m$). If $m = 3$, for example, the disaggregated and aggregated series x_t and X_T can be viewed as monthly and quarterly data, respectively. For the sake of brevity we will only report results for $m = 3$. However, for illustrative purposes, for mixed

Table I. ∇MMSE of AR (1) process: $x_t - \phi_1 x_{t-1} = a_t$ with $\sigma_a^2 = 1$ ($m = 3$)

L	k	ϕ_1							
		–0.99	–0.90	–0.50	–0.30	0.30	0.50	0.90	0.99
1	0	0.326645	0.265426	0.060589	0.019561	0.022384	0.064814	0.193495	0.223919
	1	0.659956	0.596366	0.352130	0.308994	0.431040	0.518518	0.688978	0.720682
	2	0.659990	0.600362	0.481704	0.536238	0.788491	0.851851	0.932533	0.943687
2	0	0.158394	0.091793	0.000914	0.000014	0.000016	0.000933	0.054147	0.087162
	1	0.320021	0.206243	0.005310	0.000223	0.000308	0.007466	0.192804	0.280531
	2	0.484929	0.347539	0.022898	0.002546	0.003560	0.033600	0.363984	0.477826
5	0	0.057693	0.010307	0.0	0.0	0.0	0.0	0.005580	0.028396
	1	0.116564	0.023159	0.0	0.0	0.0	0.0	0.019870	0.091390
	2	0.176630	0.039025	0.0	0.0	0.0	0.0	0.037512	0.155664
10	1	0.024530	0.000418	0.0	0.0	0.0	0.0	0.000224	0.011667
	1	0.049561	0.000941	0.0	0.0	0.0	0.0	0.000799	0.037552
	2	0.075101	0.001586	0.0	0.0	0.0	0.0	0.001510	0.063961

Table II. ∇MMSE of MA (1) process: $x_t = a_t - \theta_1 a_{t-1}$ with $\sigma_a^2 = 1$ ($m = 3$)

L	k	θ_1							
		–0.99	–0.90	–0.50	–0.30	0.30	0.50	0.90	0.99
1	0	0.089802	0.080474	0.036134	0.015680	0.022496	0.058421	0.0555556	0.008171
	1	0.4938829	0.484304	0.430442	0.395474	0.264403	0.215351	0.064814	0.00827
	2	0.897962	0.888135	0.824751	0.775269	0.506311	0.372281	0.074074	0.008369
2, 5, 10	0–2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

ARMA(p,q) structures ($p \geq 1$ and $q \geq 1$), we will also report results for $m = \{4,12\}$. For a more complete set of results including $m = \{4,12\}$ please see http://lcb1.uoregon.edu/sergiok/updating.pdf.

Several general conclusions emerge from the results in Tables I–V. First, the impact of temporal aggregation on one or two-step-ahead forecasts is significant for all models examined. The updating of disaggregated series can also dramatically improve prediction irrespective of the underlying model parametrizations for the disaggregated series and the level of aggregation. From Table I which contains the ∇MMSE's for AR(1) processes we can see, for example, when the level of aggregation $m = 3$ and the forecast horizon $L = 1$, the improvement in MMSE prediction using disaggregated data over aggregated data ranges from 2% to 33% when there is no updating. As new disaggregated data become available the improvement in MMSE is even more noticeable: for this same model structure, as well as level of aggregation and forecast horizon, for instance, when the number of updates $k = 2$, the improvement in MMSE prediction is well above 48% in all cases.

Second, the impact of temporal aggregation on prediction becomes less significant as the prediction horizon L increases. As can be seen, for the majority of ARMA parametrizations, the ∇MMSE's for $L = 5$ and 10 are generally less than 1%. For pure MA(q) processes (Tables II and III) there appears to be no difference in predictive ability of models using updated disaggregated over aggregated data for $L > q$ (see next subsection for a more detailed discussion).

Third, the impact on predictive MMSE appears to depend on the model parametrization of the disaggregated process. Consider, for example, the ARMA(1,1) parametrizations (0.8, –0.7) and (–0.8, –0.7) of Table IV. For $m = 3$, $L = 1$ and $k = 0$ the improvement in MMSE is 32.9% for the (0.8, –0.7) parametrization and just about 1% for the other parametrization. For $m = 4$, $L = 2$ and $k = 1$ the corresponding improvements are 9.3% and practically 0%, respectively. For these parametrizations when $m = 12$ noticeable changes are most apparent only when $L = 1$.

Fourth, the improvement in predictive MMSE in most cases seems to decrease as the level of aggregation increases, particularly for mixed ARMA processes. As can be seen, for example, from Table V for the ARMA(1,2) parametrization with $\Phi = 0.6$, $\theta_1 = -0.5$ and $\theta_2 = -0.9$ the one-step-ahead forecast improvement in MMSE when there is no updating is 42.1% for $m = 3$, 36.0% for $m = 4$ and 12.6% for $m = 12$. For this same parametrization when $L = 2$ and $k = 2$ the corresponding improvements are 30.4%, 8.3% and almost 0%, respectively.

Theoretical insights and an application

To gain a better understanding as to why there is so much variation in ∇MMSE due to model parametrization and level of aggregation we will examine in more detail the AR(1) and MA(q) cases since the formulae for MMSE($\hat{\hat{x}}_{T,k}(L)$) in these cases are relatively simple to derive.

Table III. ∇MMSE of AR (2) MA (2) processes ($m = 3$)

L	k	AR(2): (ϕ_1,ϕ_2)*				MA(2): (θ_1,θ_2)**			
		(1.42,–0.73)	(1.08,–0.90)	(0.50,0.30)	(–0.30,0.50)	(1.42,–0.73)	(1.08,–0.90)	(0.50,0.30)	(–0.30,0.50)
1	0	0.434233	0.570075	0.092367	0.174428	0.249624	0.110974	0.046513	0.022887
	1	0.811636	0.892212	0.604185	0.609998	0.306293	0.116362	0.076078	0.210680
	2	0.972527	0.987806	0.878210	0.738254	0.410313	0.461196	0.260862	0.706572
2	0	0.009929	0.217100	0.025395	0.051170	0.0	0.0	0.0	0.0
	1	0.021031	0.442142	0.121375	0.063282	0.0	0.0	0.0	0.0
	2	0.152802	0.671172	0.261654	0.175729	0.233574	0.247706	0.044554	0.069061
5	0	0.003678	0.062889	0.001155	0.004378	0.0	0.0	0.0	0.0
	1	0.011111	0.120576	0.005641	0.007135	0.0	0.0	0.0	0.0
	2	0.021858	0.170006	0.011819	0.010595	0.0	0.0	0.0	0.0
10	0	0.000064	0.000261	0.000009	0.000073	0.0	0.0	0.0	0.0
	1	0.000134	0.000380	0.000046	0.000119	0.0	0.0	0.0	0.0
	2	0.000181	0.002419	0.000095	0.000179	0.0	0.0	0.0	0.0

* The AR(2) process: $x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} = a_t$ with $\sigma_a^2 = 1$.
** The MA(2) process: $x_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$ with $\sigma_a^2 = 1$.

Table IV. ∇MMSE of ARMA(1,1) process: $x_t - \phi_1 x_{t-1} = a_t - \theta_1 a_{t-1}$ with $\sigma_a^2 = 1$

Panel A: $m = 3$

L	k	(ϕ_1,θ_1)							
		(0.8,–0.7)	(–0.8,0.7)	(0.8,0.7)	(–0.8,–0.7)	(0.9,–0.9)	(–0.9,0.9)	(0.3,0.5)	(–0.3,–0.5)
1	0	0.329152	0.157299	0.002261	0.005788	0.403205	0.062724	0.011518	0.011518
	1	0.767734	0.394611	0.387907	0.350446	0.814082	0.335265	0.258955	0.258955
	2	0.967963	0.515689	0.723035	0.641130	0.978968	0.594674	0.548143	0.548143
2	0	0.052718	0.016763	0.000574	0.001515	0.104771	0.009126	0.000007	0.000007
	1	0.160116	0.144794	0.007118	0.002552	0.265312	0.204475	0.000283	0.000283
	2	0.327925	0.344842	0.017342	0.004171	0.463509	0.445646	0.003344	0.003344
5	0	0.000836	0.000250	0.000010	0.000027	0.010535	0.000805	0.0	0.0
	1	0.002539	0.002160	0.000126	0.000046	0.026678	0.018052	0.0	0.0
	2	0.005201	0.005146	0.000308	0.000075	0.046609	0.039344	0.0	0.0
10	0	0.000001	0.0	0.0	0.0	0.000422	0.000032	0.0	0.0
	1	0.000003	0.000002	0.0	0.0	0.001071	0.000715	0.0	0.0
	2	0.000006	0.000006	0.0	0.0	0.001871	0.001559	0.0	0.0

Panel B: $m = 4$

L	k	(ϕ_1,θ_1)							
		(0.8,–0.7)	(–0.8,0.7)	(0.8,0.7)	(–0.8,–0.7)	(0.9,–0.9)	(–0.9,0.9)	(0.3,0.5)	(–0.3,–0.5)
1	0	0.317069	0.092993	0.004124	0.000811	0.396214	0.019671	0.013439	0.004560
	1	0.664743	0.126913	0.303381	0.233083	0.727729	0.166406	0.203288	0.267296
	2	0.883925	0.372782	0.572636	0.498946	0.915180	0.408798	0.402721	0.521928
	3	0.983990	0.498226	0.806624	0.723175	0.990405	0.639511	0.635805	0.804069
2	0	0.036390	0.012965	0.000668	1.000136	0.090894	0.007673	0.000001	0.0
	1	0.092684	0.039424	0.004934	0.000214	0.195024	0.026101	0.000021	0.000004
	2	0.180643	0.080765	0.011600	0.000336	0.323579	0.048852	0.000254	0.000041
	3	0.318079	0.145362	0.022015	0.000527	0.482289	0.076939	0.002838	0.000451
5	0	0.000162	0.000059	0.000003	0.000001	0.005472	0.000574	0.0	0.0
	1	0.000412	0.000180	0.000023	0.000001	0.011742	0.001954	0.0	0.0
	2	0.000802	0.000369	0.000054	0.000001	0.019482	0.003658	0.0	0.0
	3	0.001413	0.000664	0.000103	0.000002	0.029037	0.005761	0.0	0.0
10	0	0.0	0.0	0.0	0.0	0.000079	0.000008	0.0	0.0
	1	0.0	0.0	0.0	0.0	0.000170	0.000029	0.0	0.0
	2	0.0	0.0	0.0	0.0	0.000282	0.000054	0.0	0.0
	3	0.0	0.0	0.0	0.0	0.000420	0.000085	0.0	0.0

Panel C: $m = 12$

L	k	(ϕ_1,θ_1)							
		(0.8,–0.7)	(–0.8,0.7)	(0.8,0.7)	(–0.8,–0.7)	(0.9,–0.9)	(–0.9,0.9)	(0.3,0.5)	(–0.3,–0.5)
1	0	0.187624	0.185211	0.012950	0.000681	0.307495	0.052572	0.009984	0.001602
	1	0.304535	0.188285	0.114150	0.082637	0.431803	0.065409	0.084317	0.086108
	5	0.713330	0.243560	0.500112	0.413158	0.809133	0.197923	0.381667	0.424139
	11	0.998105	0.660070	0.952331	0.907182	0.999302	0.785544	0.854307	0.936525
2	0	0.000829	0.000633	0.000060	0.000002	0.019154	0.002929	0.0	0.0
	1	0.001467	0.001726	0.000135	0.000004	0.028101	0.009746	0.0	0.0
	5	0.010258	0.016791	0.001170	0.000029	0.090404	0.057217	0.0	0.0
	11	0.153424	0.262113	0.018017	0.000440	0.368337	0.268990	0.001172	0.000148
5	0	0.0	0.0	0.0	0.0	0.000009	0.000001	0.0	0.0
	1	0.0	0.0	0.0	0.0	0.000013	0.000004	0.0	0.0
	5	0.0	0.0	0.0	0.0	0.000044	0.000028	0.0	0.0
	11	0.0	0.0	0.0	0.0	0.000183	0.000133	0.0	0.0
10	0–11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table V. ∇MMSE of ARMA(1,2) and ARMA(2,1) processes

Panel A: $m = 3$

L	k	ARMA(1,2): $(\phi_1,\theta_1,\theta_2)$				ARMA(2,1): (ϕ_1,ϕ_2,θ_1)			
		(−0.8,1.4,−0.6)	(0.6,−0.5,0.9)	(0.3,−0.5,0.3)	(−0.3,0.3,−0.5)	(1.4,−0.6,−0.8)	(−0.5,−0.9,0.6)	(0.3,−0.5,0.3)	(−0.3,0.3,−0.5)
1	0	0.126251	0.421094	0.063589	0.123626	0.578762	0.321948	0.094035	0.026696
	1	0.436826	0.833460	0.453687	0.563018	0.891162	0.435190	0.194697	0.473843
	2	0.769191	0.969216	0.871152	0.623292	0.990316	0.440782	0.597348	0.784362
2	0	0.026649	0.025311	0.000005	0.000645	0.056436	0.109312	0.028499	0.001113
	1	0.211879	0.099109	0.000085	0.010416	0.186253	0.109470	0.040753	0.001403
	2	0.501300	0.304103	0.000977	0.118986	0.427078	0.276998	0.106928	0.020623
5	0	0.000370	0.000002	0.0	0.0	0.000021	0.063242	0.000045	0.000001
	1	0.002942	0.000009	0.0	0.0	0.000230	0.087497	0.000047	0.000003
	2	0.006961	0.000030	0.0	0.0	0.001548	0.119597	0.000241	0.000006
10	0	0.0	0.0	0.0	0.0	0.0	0.006492	0.0	0.0
	1	0.000003	0.0	0.0	0.0	0.0	0.006668	0.0	0.0
	2	0.000008	0.0	0.0	0.0	0.0	0.015401	0.0	0.0

Panel B: $m = 4$

L	k	ARMA(1,2): $(\phi_1,\theta_1,\theta_2)$				ARMA(2,1): (ϕ_1,ϕ_2,θ_1)			
		(−0.8,1.4,−0.6)	(0.6,−0.5,0.9)	(0.3,−0.5,0.3)	(−0.3,0.3,−0.5)	(1.4,−0.6,−0.8)	(−0.5,−0.9,0.6)	(0.3,−0.5,0.3)	(−0.3,0.3,−0.5)
1	0	0.073742	0.360402	0.046384	0.088285	0.532597	0.340093	0.082456	0.027189
	1	0.187493	0.698763	0.322723	0.314425	0.806060	0.535809	0.033146	0.329921
	2	0.476301	0.913340	0.604869	0.658156	0.949891	0.613334	0.226517	0.637763
	3	0.785369	0.983982	0.906809	0.705307	0.995542	0.617162	0.613258	0.851542
2	0	0.012336	0.008029	0.0	0.000048	0.003370	0.065377	0.003092	0.000051
	1	0.055691	0.027945	0.000006	0.000691	0.028020	0.065412	0.016778	0.000624
	2	0.0123435	0.083270	0.000066	0.007830	0.118320	0.230313	0.022968	0.001292
	3	0.229284	0.236949	0.000738	0.087157	0.311478	0.278287	0.057850	0.012451
5	0	0.000055	0.0	0.0	0.0	0.000110	0.012515	0.000001	0.0
	1	0.000247	0.0	0.0	0.0	0.000258	0.012676	0.000003	0.0
	2	0.000548	0.0	0.0	0.0	0.000452	0.044745	0.000008	0.0
	3	0.001017	0.000001	0.0	0.0	0.000602	0.052037	0.000009	0.0
10	0	0.0	0.0	0.0	0.0	0.0	0.003654	0.0	0.0
	1	0.0	0.0	0.0	0.0	0.0	0.005581	0.0	0.0
	2	0.0	0.0	0.0	0.0	0.0	0.006205	0.0	0.0
	3	0.0	0.0	0.0	0.0	0.0	0.010129	0.0	0.0

Panel C: $m = 12$

L	k	ARMA(1,2): $(\phi_1,\theta_1,\theta_2)$				ARMA(2,1): (ϕ_1,ϕ_2,θ_1)			
		(–0.8,1.4,–0.6)	(0.6,–0.5,0.9)	(0.3,–0.5,0.3)	(–0.3,0.3,–0.5)	(1.4,–0.6,–0.8)	(–0.5,–0.9,0.6)	(0.3,–0.5,0.3)	(–0.3,0.3,–0.5)
1	0	0.118128	0.126443	0.014915	0.029947	0.047236	0.271568	0.059049	0.012239
	1	0.118264	0.229015	0.100680	0.113180	0.084344	0.283497	0.120819	0.102705
	5	0.164280	0.629048	0.443748	0.446058	0.240443	0.497164	0.357717	0.464939
	11	0.845281	0.997128	0.970816	0.902317	0.953396	0.770980	0.819356	0.959678
2	0	0.000557	0.0	0.0	0.0	0.000304	0.036117	0.000020	0.0
	1	0.002207	0.000002	0.0	0.0	0.000780	0.036772	0.000020	0.0
	5	0.024939	0.000145	0.0	0.0	0.002781	0.206936	0.000635	0.000005
	11	0.395100	0.067086	0.000235	0.029369	0.104039	0.506176	0.045794	0.003978
5	0	0.0	0.0	0.0	0.0	0.0	0.850732	0.0	0.0
	1	0.0	0.0	0.0	0.0	0.0	0.000908	0.0	0.0
	5	0.0	0.0	0.0	0.0	0.0	0.004012	0.0	0.0
	11	0.0	0.0	0.0	0.0	0.0	0.009838	0.0	0.0
10	0	0.0	0.0	0.0	0.0	0.0	0.000003	0.0	0.0
	1	0.0	0.0	0.0	0.0	0.0	0.000003	0.0	0.0
	5	0.0	0.0	0.0	0.0	0.0	0.000008	0.0	0.0
	11	0.0	0.0	0.0	0.0	0.0	0.000020	0.0	0.0

(1) For the ARMA(1,2) case, $x_t - \phi_1 x_{t-1} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$ with $\sigma_a^2 = 1$.
(2) For the ARMA(2,1) case, $x_t - \phi_1 x_{t-1} - \theta_2 x_{t-2} = a_t - \theta_1 a_{t-1}$ with $\sigma_a^2 = 1$.

AR(1) process

Suppose that x_t follows an AR(1) process: $x_t = \phi x_{t-1} + a_t$ with $|\phi| < 1$. Since $\Psi(B) = (1 - \phi B)^{-1} = 1 + \phi B + \phi^2 B^2 + \ldots$, we have $\psi_i = \phi^i$ for $i = 0, 1, 2, \ldots$. Consequently, for $L = 1$

$$\mathrm{MMSE}(\hat{\hat{x}}_{T,k}(1)) = \sigma_a^2 (1 + \phi + \phi^2 + \ldots + \phi^{m-k-1})^2 = \sigma_a^2 \left(\sum_{i=0}^{m-k-1} \phi^i \right)^2 \tag{10}$$

If $L > 1$

$$\begin{aligned} \mathrm{MMSE}(\hat{\hat{x}}_{T,k}(L)) &= \sigma_a^2 \left[1 + (1+\phi)^2 + \ldots + (1 + \phi + \ldots + \phi^{m-1})^2 \right] \\ &\quad + \sigma_a^2 (1 + \phi + \phi^2 + \ldots + \phi^{m-1})^2 \sum_{i=1}^{m(L-1)-k} \phi^{2i} \\ &= \sigma_a^2 \left[\sum_{i=0}^{m-1} \left(\sum_{j=0}^{i} \phi^j \right)^2 + \left(\sum_{i=0}^{m-1} \phi^i \right)^2 \sum_{i=1}^{m(L-1)-k} \phi^{2i} \right] \end{aligned} \tag{11}$$

Therefore, the efficiency gain from updating disaggregated data depends not only on σ_a^2 but also ϕ. Thus, the magnitude and the sign of ϕ affect ∇MMSE.

In addition, note that $\mathrm{MMSE}(\hat{x}_t((L-1)m - k + 1)) = \sigma_a^2 \Sigma_{i=0}^{m(L-1)-k} \phi^{2i}$. Therefore, (11) can be expressed as

$$\begin{aligned} \mathrm{MMSE}(\hat{\hat{x}}_{T,k}(L)) &= \sigma_a^2 \left[1 + (1+\phi)^2 + \ldots + (1 + \phi + \ldots + \phi^{m-2})^2 \right] \\ &\quad + (1 + \phi + \phi^2 + \ldots + \phi^{m-1})^2 (\mathrm{MMSE}(\hat{x}_t((L-1)m - k + 1)) - \sigma_a^2) \end{aligned}$$

As $L \to \infty$, for any given k and m with $0 \le k < m$

$$\mathrm{MMSE}(\hat{\hat{x}}_{T,k}(L)) \to \sigma_a^2 \left[1 + (1+\phi)^2 + \ldots + (1 + \phi + \ldots + \phi^{m-2})^2 + \frac{\phi^2 (1 + \phi + \phi^2 + \ldots + \phi^{m-1})^2}{1 - \phi^2} \right]$$

Example 1 *Differences between yield on mortgages (RH) and the yield on government loans (RO)*

The data span the period from January 1961 to March 1974 with 159 monthly or 53 quarterly observations. In studying the difference between RH and RO, Butter (1976) identified the process governing the behaviour of the series as an AR(1) model if monthly data were used. He estimated the following (de-meaned) model:

$$x_t = \underset{(0.00182)}{0.841}\, x_{t-1} + a_t, \quad \mathrm{var}(a_t) = 0.0246$$

If quarterly data are used, the quarterly aggregation of x_t, namely X_T, follows the ARMA(1, 1) process

$$X_T = \Phi X_{T-1} + a_T^* - \Theta a_{T-1}^* \tag{13}$$

where $m = 3$, $\Phi = \phi^m$, $\mathrm{var}(a_T^*) = (3 + 4\phi + 5\phi^2 + 4\phi^3 + 3\phi^4)\mathrm{var}(a_t)/(1 + \Theta^2)$ and Θ is the root of $\Theta^2 + \Theta/\rho + 1 = 0$ with $\rho = (\phi + 2\phi^2 + \phi^3)/(3 + 4\phi + 5\phi^2 + 4\phi^3 + 3\phi^4)$. There exists a unique solution for Θ with $|\Theta| < 1$, since if Θ is a solution, then so is $1/\Theta$. Hence, knowledge of ϕ and $\text{var}(a_t)$ yields $\Phi = 0.595$, $\Theta = -0.217$ and $\text{var}(a_T^*) = 0.324$.

These parameters can also be estimated directly from quarterly data as: $\Phi = 0.707$, $\Theta = -0.109$ and $\text{var}(a_T^*) = 0.284$. As expected, they are close to those derived from the disaggregated monthly model.

To compare the predictive performance of $\hat{\hat{x}}_{T,k}(L)$ against $\hat{X}_T(L)$, we calculated ∇MMSE for $L = 1, 2, 3, 4$ (i.e., predicting up to one year ahead). In this simple AR(1) case, from (10), (11) and (13), we have

$$
\nabla\text{MMSE} = \frac{\text{MMSE}(\hat{X}_T(L)) - \text{MMSE}(\hat{\hat{x}}_{T,k}(L))}{\text{MMSE}(\hat{X}_T(L))}
$$

$$
= \begin{cases} \dfrac{\text{var}(a_T^*) - \sigma_a^2\left(\sum_{i=0}^{m-k-1}\phi^i\right)^2}{\text{var}(a_T^*)}, & L = 1 \\[2ex] \dfrac{\begin{array}{c}\text{var}(a_T^*)\left(1+\sum_{j=1}^{L-1}[\Phi^{j-1}(\Phi-\Theta)]^2\right) \\ -\sigma_a^2\left[\sum_{i=0}^{m-1}\left(\sum_{j=0}^{i}\phi^i\right)^2 + \left(\sum_{i=0}^{m-1}\phi^i\right)^2\sum_{i=1}^{m(L-1)-k}\phi^{2i}\right]\end{array}}{\text{var}(a_T^*)\left(1+\sum_{j=1}^{L-1}[\Phi^{j-1}(\Phi-\Theta)]^2\right)}, & L > 1 \end{cases}
$$

$$
= \begin{cases} \dfrac{(3+4\phi+5\phi^2+4\phi^3+3\phi^4)(1+\Theta^2)^{-1} - \left(\sum_{i=0}^{m-k-1}\phi^i\right)^2}{(3+4\phi+5\phi^2+4\phi^3+3\phi^4)(1+\Theta^2)^{-1}}, & L = 1 \\[2ex] \dfrac{\begin{array}{c}(3+4\phi+5\phi^2+4\phi^3+3\phi^4)(1+\Theta^2)^{-1}\left(1+\sum_{j=1}^{L-1}[\phi^{m(j-1)}(\phi^m-\Theta)]^2\right) \\ -\left[\sum_{i=0}^{m-1}\left(\sum_{j=0}^{1}\phi^3\right)^2 + \left(\sum_{i=0}^{m-1}\phi^i\right)^2\sum_{i=1}^{m(L-1)-k}\phi^{2i}\right]\end{array}}{(3+4\phi+5\phi^2+4\phi^3+3\phi^4)(1+\Theta^2)^{-1}\left(1+\sum_{j=1}^{L-1}[\phi^{m(j-1)}(\phi^m-\Theta)]^2\right)}, & L > 1 \end{cases}
$$

where Θ is the root of $\Theta^2 + \Theta/\rho + 1 = 0$ with $\rho = (\phi + 2\phi^2 + \phi^3)/(3 + 4\phi + 5\phi^2 + 4\phi^3 + 3\phi^4)$.

In Table VI two ∇MMSE's are presented: $\nabla\text{MMSE}_\text{I}$ which is evaluated where $\hat{X}_T(L)$ is based on ARMA(1, 1): $X_T = \Phi X_{T-1} + a_T^* - \Theta a_{T-1}^*$ with $\Phi = 0.595$, $\Theta = -0.217$ and $\text{var}(a_T^*) = 0.324$; and $\nabla\text{MMSE}_\text{II}$ based on the same model with $\Phi = 0.707$, $\Theta = -0.109$ and $\text{var}(a_T^*) = 0.284$, estimated directly from quarterly data.

As shown in Table VI, for short time horizons the predictions based on the AR(1) monthly model (with and without updating) appear to be more accurate than those derived from the aggregated quarterly process. There are significant gains from updating the forecasts based on the disaggregated model when new monthly data are available. When $L = 1$ both $\nabla\text{MMSE}_\text{I}$ and $\nabla\text{MMSE}_\text{II}$ are above 60% and 90% for $k = 1$ and 2, respectively. As expected, as the forecast horizon increases the differences between predictions based on monthly and quarterly data become much smaller, except for the two-quarter-ahead predictions ($L = 2$) with monthly data updating (i.e., $k = 1$ or 2).

Note that there are some differences between $\Delta\text{MMSE}_\text{I}$ and $\nabla\text{MMSE}_\text{II}$. These differences, although small in magnitude, become more noticeable as the forecast horizon increases. $\nabla\text{MMSE}_\text{I}$ is positive

Table VI. ∇MMSE's of Butter's data

L	k	$\nabla\text{MMSE}_{\text{I}}$	$\nabla\text{MMSE}_{\text{II}}$
1	0	0.173697	0.058642
	1	0.666740	0.620336
	2	0.924074	0.913502
2	0	0.035906	–0.092396
	1	0.141276	0.026997
	2	0.290254	0.195802
3	0	0.010804	–0.066012
	1	0.043480	–0.030798
	2	0.089680	0.018989
4	0	0.003738	–0.036784
	1	0.014812	–0.024135
	2	0.030466	–0.007861

for all prediction periods with or without monthly information updating (Theorem 2). On the other hand, $\nabla\text{MMSE}_{\text{II}}$ can become negative, hence validating Butter's original observations, when more than one quarter ahead is predicted because the model's parameters were allowed to be freely estimated. Note, however, that when there is updating $\nabla\text{MMSE}_{\text{II}}$ becomes positive (or nearly so) for several predictive horizons greater than 1.

The negative sign is some $\nabla\text{MMSE}_{\text{II}}$'s is due to model misspecification (see the Remark on Theorem 3) and estimation error.

MA(q) processes

Suppose x_t follows an MA(q) process. For sufficiently large L, ∇MMSE ≡ 0 for all k. In fact, both MMSE($\hat{\hat{x}}_{T,k}(L)$) and MMSE($\hat{X}_T(L)$) converge to a constant as L increases. For example, suppose x_t follows an MA(1) process: $x_t = a_t - \theta a_{t-1}$. From Theorem 2, if $m = 3$ and $L > 1$, MMSE($\hat{\hat{x}}_{T,k}(L)$) = $\sigma_a^2(1 + \theta^2 + 2(1 - \theta)^2)$. On the other hand, for all $m > 1$, X_T follows an MA(1) process: $X_T = a_T^* - \Theta a_T^*$ with $\sigma_{a*}^2 = \sigma_a^2(1 + \theta^2 + 2(1 - \theta)^2)$. Therefore, MMSE($\hat{X}_T(L)$) = MMSE($\hat{\hat{x}}_{T,k}(L)$) for $L > 1$. Hence, both ΔMMSE and $\nabla\text{MMSE}_{\text{I}}$ are zero for $L > 1$.

∇MMSE's generated from seasonal ARMA parametrizations

We also investigated the impact of temporal aggregation on the forecastibility of some typically used seasonal ARMA models. Table VII, for instance, contains ∇MMSE's for several parametrizations of the multiplicative structure ARMA(0, 1) × SARMA$(1, 0)_{12}$. The pattern of behaviour of ∇MMSE's observed her is not too dissimilar from those discussed earlier for nonseasonal processes.

In order to compare the results of Table VII with those for nonseasonal models we chose some parametrizations used in Table IV for ARMA(1, 1) processes. In addition, we added some new parametrizations to investigate the cases in which the process is purely seasonal. As can be seen, most of the general observations made earlier about ∇MMSE appear to hold. The major difference in behaviour between the results from those tables lies in the fact that the ∇MMSE's of the seasonal models do not appear to be affected by the sign of the autoregressive parameter. In fact, neither MMSE($\hat{\hat{x}}_{T,k}(L)$) nor MMSE($\hat{X}_T(L)$) depend on ϕ for $m \leq 12$ when $L = 1$, as shown below. From Theorem 2, MMSE($\hat{\hat{x}}_{T,k}(L)$) = $\sigma_\varepsilon^2 \Sigma_{i=0}^{m-k-1} (\Sigma_{j=0}^{i} \psi_j)^2$, where

Table VII. ∇MMSE of ARMA(0, 1) × SARMA(1, 0)$_{12}$ process: $x_t - \phi_{1,12}x_{t-12} = a_t - \theta_1 a_{t-1}$ with $\sigma_a^2 = 1$

Panel A: $m = 3$

L	k	($\phi_{1,12}$, θ_1)							
		(0.8, –0.7)	(–0.8, 0.7)	(0.8, 0.7)	(–0.8, –0.7)	(0.9, –0.9)	(–0.9, 0.9)	(0.9, 0.0)	(–0.5, 0.0)
1	0	0.058589	0.085427	0.085427	0.058589	0.080474	0.055555	0.0	0.0
	1	0.459869	0.155182	0.155182	0.459869	0.484304	0.064814	0.333333	0.333333
	2	0.861148	0.224938	0.224938	0.861148	0.888135	0.074074	0.666666	0.666666
2	0–2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0	0.022733	0.028264	0.028264	0.022733	0.035812	0.017968	0.0	0.0
	1	0.178433	0.051343	0.051343	0.178433	0.215526	0.020962	0.149171	0.066666
	2	0.334133	0.074423	0.074423	0.334133	0.395239	0.023957	0.298342	0.133333
10	0–2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Panel B: $m = 4$

L	k	($\phi_{1,12}$, θ_1)							
		(0.8, –0.7)	(–0.8, 0.7)	(0.8, 0.7)	(–0.8, –0.7)	(0.9, –0.9)	(–0.9, 0.9)	(0.9, 0.0)	(–0.5, 0.0)
1	0	0.043666	0.101388	0.101388	0.043666	0.059280	0.072730	0.0	0.0
	1	0.329478	0.165069	0.165069	0.329478	0.346352	0.081732	0.250000	0.250000
	2	0.615291	0.228751	0.228751	0.615291	0.633416	0.090735	0.500000	0.500000
	3	0.901103	0.292432	0.292432	0.901103	0.920480	0.099738	0.750000	0.750000
2	0–3	0.390243	0.0	0.0	0.390243	0.0	0.0	0.0	0.0
5	0–3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	0	0.004930	0.009443	0.009443	0.004930	0.010467	0.008370	0.0	0.0
	1	0.037200	0.015374	0.015374	0.037204	0.061148	0.009408	0.044323	0.002941
	2	0.069476	0.021305	0.021305	0.069476	0.111828	0.010445	0.088646	0.005882
	3	0.101749	0.027236	0.027236	0.101749	0.162509	0.011481	0.132969	0.008825

Panel C: m = 12

L	k	($\phi_{1,12}$, θ_1)							
		(0.8, –0.7)	(–0.8, 0.7)	(0.8, 0.7)	(–0.8, –0.7)	(0.9, –0.9)	(–0.9, 0.9)	(0.9, 0.0)	(–0.5, 0.0)
1	0	0.014287	0.120839	0.120839	0.014287	0.019047	0.142238	0.0	0.0
	1	0.101164	0.160600	0.160600	0.101164	0.106034	0.149966	0.083333	0.083333
	5	0.448674	0.319644	0.319644	0.448674	0.453982	0.180876	0.416666	0.416666
	11	0.969938	0.558210	0.558210	0.969938	0.975903	0.227241	0.916666	0.916666
2	0	0.005461	0.034673	0.062326	0.005690	0.008341	0.0	0.0	0.000771
	1	0.038674	0.046082	0.082835	0.040295	0.046438	0.034260	0.037292	0.016666
	5	0.171524	0.091718	0.164867	0.178713	0.198823	0.041321	0.186464	0.083333
	11	0.370800	0.160172	0.287915	0.386339	0.427401	0.051914	0.410220	0.183333
5	0	0.000930	0.005273	0.013023	0.000997	0.002312	0.007095	0.0	0.0
	1	0.006634	0.007008	0.017308	0.007063	0.012870	0.007480	0.010464	0.000244
	5	0.029422	0.013949	0.034448	0.031329	0.055104	0.009021	0.052322	0.001222
	11	0.063606	0.024360	0.060159	0.067727	0.118455	0.011334	0.115109	0.002688
10	0	0.000090	0.000499	0.001313	0.000096	0.000595	0.001723	0.0	0.0
	1	0.000641	0.000664	0.001745	0.000686	0.003315	0.001817	0.002705	0.0
	5	0.002847	0.001321	0.003475	0.003044	0.014195	0.002191	0.013527	0.0
	11	0.006155	0.002308	0.006068	0.006580	0.030516	0.002753	0.029759	0.000002

$$\psi_{12j+m} = \begin{cases} \phi^j, & m = 0 \\ -\theta\phi^j, & m = 1 \\ 0, & \text{otherwise} \end{cases}$$

for $j = 0, 1, 2, 3, \ldots$. When $m \leq 12$, only ψ_0 and ψ_1, which do not depend on ϕ, are required to calculate MMSE($\hat{\hat{x}}_{T,k}(L)$). When $L > 1$ and $m = 3$ or 4, MMSE($\hat{\hat{x}}_{T,k}(L)$) depends on the magnitude of ϕ but not on its sign. When $L > 1$, MMSE($\hat{\hat{x}}_{T,k}(L)$) = $\sigma_\varepsilon^2[\Sigma_{i=0}^{m-1}(\Sigma_{j=0}^{i}\psi_j)^2 + \Sigma_{i=1}^{m(L-1)-k}(\Sigma_{j=0}^{m-1}\psi_{i+j})^2]$. In the cases that $m = 3$ and 4, MMSE($\hat{\hat{x}}_{T,k}(L)$) is essentially the summation of several $\sigma_\varepsilon^2(\phi^j - \theta\phi^j)^2$ factors, which do not depend on the sign of ϕ. However, if $m = 12$, MMSE($\hat{\hat{x}}_{T,k}(L)$) contains terms such as $\sigma_\varepsilon^2(\phi^j - \theta\phi^{j+1})^2$, which depend on both the magnitude and the sign of ϕ. Similar arguments can be made for MMSE($\hat{X}_T(L)$).

When the data are generated by purely seasonal AR processes the results in Table VII confirm findings of published studies such as Wei (1978) which showed that there is no gain in predictive ability for seasonal models by using models based on nonupdated disaggregated data ($k = 0$) over those constructed from aggregated data. However, as shown in Table VII, ∇MMSE is, in general, not zero if updating is allowed ($k > 0$).

We also studied the impact of temporal aggregation for other seasonal processes. For the various parametrizations of the ARMA(0, 1) × SARMA$(0, 1)_{12}$ structure (results not shown here), for example, we found that there is a substantial gain in using predictions based on disaggregated data for short horizons, but these gains decrease as L increases. When $L = 5$ or 10, there appears to be no loss in predictive efficiency by using aggregated data (see http://lcb1.uoregon.edu/sergiok/updating.pdf).

PREDICTIONS BASED ON NONSTATIONARY MODELS

Nonseasonal integrated models

A widely used class of models to represent business and economic series (Ermini, 1991; Ali and Zarowin, 1992; Brown, 1993; Koreisha and Pukkila, 1995) is the IMA(1, 1) structure: $(1 - B)x_t = (1 - \theta B)a_t$. For this model structure, the MMSE forecast is an exponentially weighted average of the observations with the smoothing parameter $(1 - \theta)$.

The temporally aggregated series, X_T, follows IMA(1, 1): $(1 - B)X_T = (1 - \Theta B)a_T^*$. The parameters Θ and σ_{a*}^2 can be obtained from the relations

$$\frac{-\Theta}{1+\Theta^2} = cp + \frac{1}{4}(1-c) \tag{14}$$

and

$$\sigma_{a*}^2(1+\Theta^2) = m\sigma_a^2(1+\theta^2)c^{-1} \tag{15}$$

where $c^{-1} = 1 + \frac{2}{3}(m^2 - 1)(1 + 2\rho)$ and $\rho = -\theta/(1 + \theta^2)$. Note that the first-order autocorrelation of X_T is $c\rho + \frac{1}{4}(1 - c)$, which approaches 0.25 as $m \to \infty$. The limit for Θ is -0.268 as $m \to \infty$.

Following the approach in Box and Jenkins (1976), we express both x_t and X_T as infinite moving averages: $x_t = \Sigma_{j=0}^{\infty}\psi_j a_{t-j}$ and $X_T = \Sigma_{j=0}^{\infty}\Psi_j a_T^* - j$. Consequently, MMSE($\hat{\hat{x}}_{T,k}(L)$) can be obtained using (6). Since X_T can be expressed as $X_T = a_T^* + \Sigma_{i=1}^{\infty}(1-\Theta)a_{T-i}^*$, MMSE($\hat{X}_T(L)$) = $[1 + (L-1)(1-\Theta)^2]\sigma_{a^*}^2$. In Table VIII we present ∇MMSE's for several IMA(1, 1) parametrizations.

As shown the behaviour of ∇MMSE's depends on L, m, k and the MA coefficient θ. ∇MMSE's show similar characteristics to those observed earlier for stationary processes. For any given m, k and θ, ∇MMSE can be quite large even for large L. The magnitude of the gain in predictive efficiency by using disaggregated data, however, decreases as L increases but at a much slower decay rate than the corresponding stationary MA(1) processes in Table II. This slower decay rate is likely due to the fact that for nonstationary processes the forecasting errors increase without bound as L increases. We also note that for fixed L, m and k, ∇MMSE decreases as θ increases.

It should be noted, however, that although ∇MMSE for nonstationary processes decreases at a slower rate than that of corresponding stationary processes, for large m, ∇MMSE converges to zero as $L \to \infty$. From (6) (which also holds for nonstationary processes), we see that

$$\begin{aligned}\text{MMSE}\left(\hat{\hat{x}}_{T,k}(L)\right) &= \sigma_a^2\left[\sum_{i=0}^{m-1}(1+i(1-\theta))^2 + \sum_{i=1}^{m(L-1)-k}(m(1-\theta))^2\right] \\ &= \sigma_a^2 m^3(1-\theta)^2 L + O(1) \qquad (16)\end{aligned}$$

for sufficiently large L. Similarly, from (15) we have

$$\begin{aligned}\text{MMSE}(\hat{X}_T(L)) &= \left[1+(L-1)(1-\Theta)^2\right]\sigma_{a*}^2 \\ &= m\sigma_a^2(1+\theta^2)\left[1+\frac{2}{3}(m^2-1)(1+2\rho)\right]\frac{(1-\Theta)^2}{1+\Theta^2}L + O(1) \\ &= \frac{2}{3}\sigma_a^2 m^3(1-\theta)^2\frac{(1-\Theta)^2}{1+\Theta^2}L + \sigma_a^2 m\left[(1+\theta^2) - \frac{2}{3}(1-\theta^2)\right]\frac{(1-\Theta)^2}{1+\Theta^2}L + O(1) \qquad (17)\end{aligned}$$

From (14), $(1-\Theta)^2/(1+\Theta^2) \approx 3/2$ since for large m, $c \approx 0$. Hence, for large m, (17) will be dominated by the first term, namely $\sigma_a^2 m^3(1-\theta)^2 L$, which from (16) equals MMSE($\hat{\hat{x}}_{T,k}(L)$). Therefore, for sufficiently large m, ∇MMSE converges to zero as $L \to \infty$.

A seasonally integrated model

Example 2 Monthly totals of international airline passengers (Box and Jenkins, 1976).

The monthly airline passenger data can be represented by the nonstationary model:

$$(1-B)(1-B^{12})x_t = (1-\theta_1 B)(1-\theta_{1,12}B^{12})a_t \qquad (18)$$

where $\theta_1 = 0.4$, $\theta_{1,12} = 0.6$ and var(a_t) = 0.00134. The model structure based on monthly data implies that the model based on quarterly data would follow an ARIMA(0, 1, 1) × SARIMA(0, 1, 1)$_4$ (see, for example, Wei, 1978), namely

$$(1-B)(1-B^4)X_T = (1-\Theta_1 B)(1-\Theta_{1,4}B^4)a_T^* \qquad (19)$$

where $\Theta_{1,4} = \theta_{1,2}$ and Θ_1 and var(a_T^*) are determined by

Table VIII. ∇MMSE of IMA(1, 1) process: $x_t = x_{t-1} + a_t - \theta_1 a_{t-1}$ with $\sigma_a^2 = 1$

Panel A: $m = 3$

L	k	θ_1								
		–0.99	–0.90	–0.50	–0.30	0.0	0.30	0.50	0.90	0.99
1	0	0.455066	0.441833	0.367346	0.318140	0.227238	0.123359	0.060589	0.000862	0.000001
	1	0.844081	0.838140	0.802721	0.777200	0.724013	0.646618	0.578885	0.395042	0.339945
	2	0.984314	0.982799	0.972789	0.964578	0.944802	0.909156	0.870426	0.726263	0.673256
2	0	0.169980	0.165074	0.138461	0.121600	0.091248	0.056024	0.032321	0.000803	0.000001
	1	0.378804	0.373870	0.346153	0.327526	0.290730	0.237971	0.187840	0.023744	0.000294
	2	0.587627	0.582665	0.553846	0.533452	0.490212	0.419917	0.343360	0.046684	0.000588
5	0	0.059032	0.057334	0.048257	0.042616	0.032642	0.021241	0.013468	0.000665	0.000001
	1	0.131555	0.129853	0.120643	0.114787	0.104005	0.090225	0.078277	0.019680	0.000294
	2	0.204078	0.202373	0.193029	0.186957	0.175367	0.159209	0.143086	0.038695	0.000587
10	0	0.028274	0.027461	0.023136	0.020463	0.015766	0.010439	0.006829	0.000518	0.000001
	1	0.063010	0.062196	0.057840	0.055118	0.050233	0.044341	0.039691	0.015313	0.000292
	2	0.097746	0.096931	0.092544	0.089773	0.084700	0.078244	0.072554	0.030108	0.000584

Panel B: $m = 4$

L	k	θ_1								
		–0.99	–0.90	–0.50	–0.30	0.0	0.30	0.50	0.90	0.99
1	0	0.458961	0.449206	0.393382	0.355058	0.278775	0.176018	0.098997	0.001960	0.000002
	1	0.774416	0.768900	0.736376	0.713011	0.663428	0.587153	0.516128	0.317819	0.257414
	2	0.935455	0.932984	0.917794	0.906225	0.879795	0.833577	0.783091	0.586953	0.509852
	4	0.993506	0.992878	0.988661	0.985091	0.975959	0.957217	0.933259	0.813101	0.757364
2	0	0.173472	0.169806	0.149418	0.135907	0.109813	0.075168	0.047874	0.001751	0.000002
	1	0.328980	0.325302	0.304460	0.290213	0.261334	0.218406	0.176976	0.028460	0.000390
	2	0.484488	0.480799	0.459503	0.444519	0.412855	0.361643	0.306078	0.055169	0.000777
	3	0.639996	0.636295	0.614546	0.598826	0.564377	0.504880	0.435179	0.08187	0.001165
5	0	0.060525	0.059249	0.052234	0.047658	0.038965	0.027647	0.018780	0.001326	0.000002
	1	0.114783	0.113506	0.106435	0.101769	0.092729	0.080330	0.069423	0.021552	0.000388
	2	0.169041	0.167762	0.160636	0.155880	0.146493	0.133013	0.120067	0.041778	0.000774
	3	0.223298	0.222018	0.214837	0.209990	0.200257	0.185696	0.170710	0.062005	0.001159
10	0	0.029026	0.028415	0.025064	0.022888	0.018775	0.013462	0.009329	0.000944	0.000002
	1	0.055048	0.054436	0.051072	0.048875	0.044682	0.039115	0.034489	0.015344	0.000385
	2	0.081068	0.080456	0.07708	0.074862	0.070589	0.064768	0.059649	0.029745	0.000768
	3	0.107089	0.106477	0.103088	0.100849	0.096496	0.090422	0.084809	0.044146	0.001150

Panel C: $m = 12$

L	k	θ_1								
		–0.99	–0.90	–0.50	–0.30	0.0	0.30	0.50	0.90	0.99
1	0	0.463329	0.460136	0.441445	0.427963	0.398305	0.346374	0.284962	0.028108	0.000064
	1	0.586740	0.584055	0.568295	0.556877	0.531604	0.486787	0.432690	0.169748	0.092203
	2	0.893784	0.892694	0.886197	0.881384	0.870404	0.849755	0.822551	0.611050	0.444406
	4	0.999764	0.999741	0.999585	0.999449	0.999074	0.998144	0.996503	0.967881	0.925213
2	0	0.177398	0.176178	0.169108	0.164082	0.153219	0.134736	0.113508	0.018076	0.000063
	1	0.228825	0.227605	0.220518	0.215465	0.204496	0.185654	0.163647	0.047818	0.001126
	2	0.434533	0.433310	0.426158	0.420997	0.409603	0.389323	0.364205	0.166789	0.005379
	3	0.743093	0.741868	0.734619	0.729296	0.717264	0.694827	0.665041	0.345246	0.011758
5	0	0.062215	0.061787	0.059320	0.057576	0.053837	0.047559	0.040466	0.008729	0.000061
	1	0.080251	0.079823	0.077354	0.075607	0.071854	0.065532	0.058341	0.023093	0.001085
	2	0.152394	0.151966	0.149489	0.147729	0.143924	0.137423	0.129840	0.080546	0.005181
	3	0.260609	0.260181	0.257692	0.255911	0.252028	0.256260	0.237089	0.166726	0.011325
10	0	0.029880	0.029675	0.028491	0.027656	0.025870	0.022883	0.0195	0.004688	0.000057
	1	0.038542	0.038337	0.037153	0.036317	0.034528	0.031530	0.028150	0.012403	0.001022
	2	0.073191	0.072985	0.071799	0.070961	0.06915	0.066120	0.062649	0.043262	0.004881
	3	0.125163	0.124958	0.123769	0.122926	0.121106	0.118006	0.114398	0.089551	0.010600

Table IX. ∇MMSE's of airline passenger data

L	k	$\nabla \text{MMSE}_{\text{I}}$	$\nabla \text{MMSE}_{\text{II}}$
1	0	0.090056	0.134153
	1	0.614357	0.633046
	2	0.891673	0.896923
2	0	0.043746	0.001462
	1	0.214732	0.180008
	2	0.385718	0.358555
5	0	0.028341	–0.02282
	1	0.154005	0.109457
	2	0.258169	0.219106
10	0	0.015022	0.004172
	1	0.074360	0.064163
	2	0.133698	0.124155
15	0	0.010357	0.034605
	1	0.051565	0.074803
	2	0.092772	0.115001
20	0	0.007923	0.057932
	1	0.039674	0.088082
	2	0.071424	0.118231

$$\text{var}(a_T^*)(1+\Theta_1^2) = \text{var}(a_t)(19\theta_1^2 - 32\theta_1 + 19)$$
$$-\text{var}(a_T^*)\Theta_1 = \text{var}(a_t)(4\theta_1^2 - 11\theta_1 1 + 4)$$

Hence, based on (18), $\Theta_1 = -0.026$ and $\text{var}(a_T^*) = 0.01237$.

Direct estimation of model (19) based on quarterly data yields $\Theta_1 = 0.067$, $\Theta_{1,4} = 0.524$ and $\text{var}(a_T^*) = 0.013$. The 95% confidence intervals for Θ_1 and $\theta_{1,4}$ are (–0.237, 0.370) and (0.271, 0.778), respectively.

As was the case for Butter's data, in Table IX we present two measures of predictive efficiency: $\nabla \text{MMSE}_{\text{I}}$ based on derived estimates for the quarterly parameters and $\nabla \text{MMSE}_{\text{II}}$ based on direct estimates of these parameters. As can be observed, the quarterly forecasts based on the model estimated from monthly data (with and without updating) are superior to those generated from the model based on quarterly (aggregated) data. Like the IMA(1, 1) case, the most striking difference between the results here and those of stationary processes is that the ∇MMSE's decrease at a much slower rate as L increases than those associated stationary processes SARMA(0, 1) × SARMA$(0, 1)_{12}$ with similar parameters.

CONCLUSIONS

In this article we have compared the performance of two predictors for temporally aggregated series: one based on ARMA models generated from aggregated data, and the other based on aggregation of predictions constructed from disaggregated data and which are also permitted to be updated as the new information becomes available. We have shown that for lower order ARMA models, the loss of information can be substantial if models based on aggregated data are used for short-term predictions. Updating models based on disaggregated data can improve forecasts dramatically. However, this gain in predictive efficiency decreases as the forecast horizon increases.

In the seasonal and nonstationary cases we examined, the gain in predictive efficiency as measured by MMSE by using models based on disaggregated data (with and without updating) appear to be more long-lasting than for stationary ARMA processes. This has important practical implications because most economic time series seem to exhibit some sort of trend and seasonality.

It should be noted, however, that in many applications such as financial analysis of stock prices and exchange rates, the more finely the data is sampled, e.g. daily, hourly, minute-by-minute prices, the more likely it is that the error component in the data (relating to the signal) will be larger than when the data are sampled less frequently, such as weekly, monthly or annually (Merville and Pieptea, 1989; Hasbrouck, 1993; Zhou, 1996). A very interesting as well as productive line of research might examine to what extent the results derived in this article for data without measurement noise carry over to observations measured with noise.

APPENDIX

Proof of Theorem 1: Using the fact that (3) is the conditional expectation of X_{T+L} give $\{x_{mT+k}, x_{mT+k-1}, \ldots\}$, the result in Theorem 1 follows.

Proof of Theorem 2: Note that for any L, the MMSE of $\hat{\hat{x}}_{T,k}(L)$ is the same as the MMSE of $\Sigma_{j=1}^{k} x_{mT+j} + \Sigma_{i=k+1}^{m} \hat{x}_{mT}(m(L-1)+i)$. Hence

$$\begin{aligned}\text{MMSE}(\hat{\hat{x}}_{T,k}(L)) &= E\left(\sum_{i=1}^{k} x_{mT+j} + \sum_{i=k+1}^{m} \hat{x}_{mT}(m(L-1)+i) - X_{T+L}\right)^2 \\ &= E\left(\sum_{i=k+1}^{m} \hat{x}_{mT}(m(L-1)+i) - \sum_{i=k+1}^{m} x_{mT+(m(L-1)+i}\right)^2 \\ &= E\left(\sum_{i=k+1}^{m} \sum_{j=0}^{m(L-1)+i-1} \psi_j \varepsilon_{mT+m(L-1)+i-j}\right)^2\end{aligned}$$

Then, (6) follows.

Proof of Theorem 3:

$$\begin{aligned}\text{MMSE}(\hat{\hat{x}}_{T,k}(L)) &= E(\hat{\hat{x}}_{T,k}(L) - X_{T+L})^2 = \min_{\tilde{x} \in \mathcal{H}_{d,k}} E(\tilde{x} - X_{T+L})^2 \\ &\le \min_{\tilde{X} \in \mathcal{H}_a} E(\tilde{X} - X_{T+L})^2 = E(\hat{X}_T(L) - X_{T+L})^2 = \text{MMSE}(\hat{X}_T(L))\end{aligned}$$

where $\mathcal{H}_{d,k}$ and $\mathcal{H}_a$ are the closed linear manifold spanned by $\{x_i\}_{i=1}^{t+k}$ and $\{X_i\}_{i=1}^{T}$ ($t = mT$ and $0 \le k < m$), respectively.

REFERENCES

Ali A, Zarowin P. 1992. Permanent versus transitory components of annual earnings and estimation error in earnings response coefficients. *Journal of Accounting and Economics* **15**: 249–264.

Amemiya T, Wu R. 1972. The effect of aggregation on prediction in the autoregressive model. *Journal of the American Statistical Association* **67**: 628–632.
Box GEP, Jenkins GM. 1976. *Time Series Analysis: Forecasting and Control*. Holden-Day: New York.
Brewer K. 1973. Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models. *Journal of Econometrics* **1**: 133–154.
Brown L. 1993. Earnings forecasting research: its implications for capital markets research. *International Journal of Forecasting* **9**: 295–320.
Butter F. 1976. The use of monthly and quarterly data in an ARMA model. *Journal of Econometrics* **4**: 311–324.
Ermini L. 1991. Reinterpreting a temporally aggregated consumption CAP model. *Journal of Business and Economic Statistics* **9**: 325–328.
Harvey A. 1981. *Time Series Models*. Helsted Press: New York.
Hasbrouck J. 1993. Assessing the quality of a security market: a new approach to transaction-cost measurement. *Review of Financial Studies* **6**: 191–212.
Koreisha S, Fang Y. 1999. The impact of measurement errors on ARMA prediction. *Journal of Forecasting* **18**: 95–110.
Koreisha S, Fang Y. 2001. Generalized least squares with misspecified serial correlation structures. *Journal of the Royal Statistical Society, Series B* **63**: 515–531.
Koreisha S, Pukkila T. 1995. A comparison between different order-determination criteria for identification of ARIMA models. *Journal of Business & Economic Statistics* **13**: 127–131.
Lütkepohl H. 1984. Forecasting contemporaneously aggregated vector ARMA processes. *Journal of Business and Economic Statistics* **2**: 201–214.
Lütkepohl H. 1987. *Forecasting Aggregated Vector ARMA Processes*. Springer-Verlag: Berlin.
Merville L, Pieptea D. 1989. Stock-price volatility, mean-reverting diffusion and noise. *Journal of Financial Economics* **24**: 193–214.
Nijman T, Palm F. 1990. Predictive accuracy gain from disaggregate sampling in ARIMA models. *Journal of Business & Economic Statistics* **8**: 405–415.
Pino F, Morettin P, Mentz R. 1987. Modelling and forecasting linear combinations of time series. *International Statistical Review* **55**: 295–313.
Pukkila T, Koreisha S, Kallinen A. 1990. The identification of ARMA models. *Biometrika* **77**: 537–548.
Stram D, Wei W. 1986. Temporal aggregation in the ARIMA process. *Journal of Time Series Analysis* **7**: 279–292.
Tiao G. 1972. Asymptotic behavior of temporal aggregates of time series. *Biometrika* **59**: 525–531.
Tiao G, Wei W. 1976. Effect of temporal aggregation on the dynamic relationship of two time series variables. *Biometrika* **63**: 513–523.
Wei W. 1978. Some consequences of temporal aggregation in seasonal time series models. In *seasonal Analysis of Economic Time Series Analysis*, Zellner A (eds). US Department of Commerce, Bureau of the Census: Washington, DC.
Wei W. 1990. *Time Series Analysis*. Addison-Wesley: New York.
Zhou B. 1996. High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics* **14**: 45–52.

Authors' biographies:
Sergio G. Koreisha is the Charles H. Lundquist Research Scholar and Professor of Decision Sciences at the Lundquist College of Business, University of Oregon. He specializes in the areas of forecasting, mathematical model building, applied econometrics, production analysis and manufacturing strategy. Some of his latest works in the field of forecasting include articles dealing with new fast estimation methods for large-scale, multiple time series models, new approaches for identifying the mathematical structures governing the behaviour of economic time series, and how causality among economic and business variables might be ascertained and used to formulate more accurate models. He has taught courses in Europe, South America and the Pacific.

Yue Fang is Associate Professor of Decision Sciences at the Lundquist College of Business, University of Oregon. He has a PhD in Statistics from MIT. His publications include articles on financial econometrics, time series forecasting and statistical quality controls. His recent work focuses on the quantitative analysis of financial markets using mathematical and statistical models.

Authors' addresses:
Sergio G. Koreisha, Lundquist College of Business, University of Oregon, Eugene, OR 97403, USA.

Yue Fang, Lundquist College of Business, University of Oregon, Eugene, OR 97403, USA.

