GENERALIZATIONS OF THE BOX-JENKINS AIRLINE MODEL WITH FREQUENCY-SPECIFIC SEASONAL COEFFICIENTS AND A GENERALIZATON OF AKAIKE'S MAIC

John Aston, David Findley, Kellie Wills, U.S Census Bureau, Washington, D.C.
Donald Martin, U.S Census Bureau and Howard University, Washington, D.C.

Key Words: Time series, seasonal adjustment, forecasting, Model Selection, AIC

Abstract. The Box-Jenkins "airline" model is the most widely used ARIMA model for seasonal time series. Findley, Martin and Wills (2002) examined a generalization of the airline model with a more restricted seasonal moving average factor that models only seasonal effects and with a second-order nonseasonal moving average factor. Here, we generalize the seasonal factor further by associating a subset of the frequencies 1, 2, 3, 4, 5 and 6 cycles per year (in the case of monthly data) with one coefficient and the complementary subset with a second coefficient. A generalization of Akaike's Minimum AIC criterion is presented for choosing among subsets of a given size or, more generally, among generalizations of the airline model having the same number of coefficients. Properties of model-based seasonal adjustment filters obtained from the new models are considered as well as forecasting performance relative to the airline model.

1. Introduction

Box and Jenkins (1976) developed a two-coefficient time series model, now known as the airline model, which is by far the most widely used ARIMA model for monthly and quarterly macroeconomic time series. (Fischer and Planas (2000) deem it adequate for 50% of 13,232 Eurostat time series.) The Box-Jenkins airline model for a seasonal time series Z_t observed $s \geq 2$ times per year has the form

$$(1-B)(1-B^s)Z_t = (1-\theta B)(1-\Theta B^s)\varepsilon_t \qquad (1)$$

When $\Theta > 0$, the airline model can be written

$$(1-B)^2(\sum_{j=0}^{s-1} B^j)Z_t = (1-\theta B)(1-\Theta^{1/s}B)(\sum_{j=0}^{s-1}\Theta^{j/s}B^j)\varepsilon_t \qquad (2)$$

Findley, Martin and Wills (2002) substituted a general MA(2) polynomial for $(1-\theta B)(1-\Theta^{1/s}B)$ in (2), yielding a *generalized airline model*,

$$(1-B)^2(\sum_{j=0}^{s-1} B^j)Z_t = (1-aB-bB^2)(\sum_{j=0}^{s-1} c^j B^j)\varepsilon_t \qquad (3)$$

In this model the seasonal sum polynomial has a third coefficient c distinct from the coefficients associated with the other factors in the model.

In the present paper, we investigate various generalizations of (2) and (3), which we call *frequency-specific* models. In these models, the final moving average factor, e.g. $\sum_{j=0}^{s-1} c^j B^j$ in (3), is decomposed into several factors with different coefficients. Restricting attention to monthly data, i.e. $s=12$, the model (3) can be generalized by factoring $\sum_{j=0}^{11} c^j B^j$ in terms of frequencies of 1, 2, 3, 4, 5 and 6 cycles per year to obtain a general frequency-specific model,

$$(1-B)^2(\sum_{j=0}^{11} B^j)Z_t = (1-aB-bB^2)\left[(1+c_6B)\prod_{j=1}^{5}\left(1-2c_j\cos(\tfrac{2\pi j}{12})B+c_j^2B^2\right)\right]\varepsilon_t \qquad (4)$$

If the six c_i's are distinct, this model has a different seasonal coefficient for each seasonal frequency, for a total of eight coefficients. Eight moving average coefficients cannot be estimated reliably from macroeconomic time series of typical lengths. We shall consider instead the most parsimonious such generalizations of (2) in which there are only two distinct c_i's. That is, the seasonal frequencies are divided into two groups, with all frequencies in a group having the same coefficient. This reduces the total number of coefficients requiring estimation in the model to four in general, and to three when we constrain the MA(2) in (4) to have a factor whose coefficient is one of the seasonal coefficients c_i, in analogy with (2).

We consider two types of four-coefficient models. The first, designated the 5-1(4) type, is the one in which five of the frequency factors in brackets in (4) have the same coefficient c_1 and the sixth has a different coefficient c_2. There are six such models, an example being

---

This paper is released to inform about ongoing research and to encourage discussion. Any views expressed are the authors' and not necessarily those of the U.S. Census Bureau. John Aston is now at Academia Sinica, Taipei. Kellie Wills is now at the Corporate Executive Board, Washington, D.C.

$$(1-B)^2(\sum_{j=0}^{11} B^j)Z_t = (1-aB-bB^2)$$
$$\times\left\{(1+c_2B)\prod_{j=1}^{5}\left(1-2c_1\cos(\tfrac{2\pi j}{12})B+c_1^2B^2\right)\right\}\varepsilon_t. \quad (5)$$

The second type, designated 4-2(4), is the one in which four of the frequency factors in brackets in (4) have the same coefficient c_1 and the remaining two have a different coefficient c_2. There are fifteen such models.

We also consider the corresponding types of three-coefficient frequency-specific generalizations of (2), denoted 5-1(3) and 4-2(3) models. An example of the former is

$$(1-B)^2(\sum_{j=0}^{11} B^j)Z_t = (1-aB)(1-c_1B)$$
$$\times\left\{(1+c_2B)\prod_{j=1}^{5}\left(1-2c_1\cos(\tfrac{2\pi j}{12})B+c_1^2B^2\right)\right\}\varepsilon_t. \quad (6)$$

There are six 5-1(3) models and fifteen 4-2(3) models. We did not consider the twenty models that arise by grouping the frequency factors in (4) into two groups of three factors because it would increase the number of models to be compared from 22 to 42.

These new models cannot be estimated with standard ARIMA modeling software. We performed the estimation in the object-oriented matrix programming environment Ox (Doornik 2001), using the state space functions in the SSFPack library (Koopman, Shephard and Doornik 1999). Some details are given in the Appendix.

Before presenting our evaluation of the new models relative to the airline model for seasonal adjustment and forecasting using Census Bureau, we present our model selection criterion for deciding when one of the new models should be considered for replacing the airline model. This criterion is a generalization of Akaike's Minimum AIC criterion (MAIC).

2 A Modification of Akaike's Minimum AIC Procedure for Multiple Fixed-Dimension Comparisons to a Nested Model

For the airline model, let $\hat{\theta}^A$, $\dim\theta^A$, and $L(\hat{\theta}^A)$ denote the estimated parameter vector, its dimension, and the associated maximum log-likelihood value respectively. Let $\hat{\theta}^F$, $\dim\theta^F$, and $L(\hat{\theta}^F)$ denote the corresponding quantities for a frequency-specific model. Consider the AIC difference

$$\Delta AIC^{A,F} \equiv AIC(\hat{\theta}^A) - AIC(\hat{\theta}^F) =$$
$$-2\{\ln L(\hat{\theta}^A) - \ln L(\hat{\theta}^F)\} - 2(\dim\theta^F - \dim\theta^A). \quad (7)$$

Because the airline model is a special case of each type of frequency-specific model, when the airline model is correct,

$$-2\{\ln L(\hat{\theta}^A) - \ln L(\hat{\theta}^F)\} \sim \chi^2_{\dim\theta^F - \dim\theta^A} \quad (8)$$

asymptotically under standard assumptions; see Taniguchi and Kakizawa (2000, p. 61). Under (8), the probability that the frequency-specific model will have a smaller AIC and thus be incorrectly preferred by Akaike's Minimum AIC criterion (MAIC) is

$$P\{AIC(\hat{\theta}^A) - AIC(\hat{\theta}^F) > 0\} =$$
$$P\{\chi^2_{\dim\theta^F - \dim\theta^A} > 2(\dim\theta^F - \dim\theta^A)\}. \quad (9)$$

Thus the probability of incorrectly rejecting the airline model in favor of a frequency-specific model is $p^{(4)} \equiv P\left\{\chi_2^2 > 4\right\} = 0.135$ for a four-coefficient model and $p^{(3)} \equiv P\left\{\chi_1^2 > 2\right\} = 0.157$ for a three-coefficient model.

We now describe a generalization of MAIC that is directed toward achieving the same type I error probabilities $p^{(i)}$ when the minimum AIC model from a set $\mathcal{F}(i)$ containing several i-coefficient models more general than (2) is compared to the estimated airline model, for $i = 3, 4$. Our criterion is to prefer this MAIC model over the estimated airline model for a time series of length N when, for a threshold $\Delta_N^{(i)} = \Delta_N^{(i)}(\mathcal{F}(i)) \geq 0$ with a certain property, the inequality

$$AIC(\hat{\theta}^A) - \min_{F\in\mathcal{F}(i)} AIC(\hat{\theta}^F) > \Delta_N^{(i)} \quad (10)$$

holds. The property desired of $\Delta_N^{(i)}$ is

$$P\{AIC(\hat{\theta}^A) - \min_{F\in\mathcal{F}(i)} AIC(\hat{\theta}^F) > \Delta_N^{(i)}\} \doteq p^{(i)} \quad (11)$$

when the airline model is correct. This is the property of $\Delta_N^{(i)} = 0$ when $\mathcal{F}(i)$ contains only one model. We call the resulting generalization of MAIC the GMAIC criterion.

Such $\Delta_N^{(i)}$'s can be obtained from the empirical distribution of $AIC(\hat{\theta}^A) - \min_{F\in\mathcal{F}(i)} AIC(\hat{\theta}^F)$ when the models in $\mathcal{F}(i)$ are fit to simulated Gaussian time series of length N generated by an airline model, for example, the airline model estimated from the time series of interest, or from a nearby model for a series of approximately the same length. In this paper, for illustrative purposes, we use the $\Delta_N^{(i)}$'s in Table 1 below. These values were obtained from simulated series from a single airline model with coefficients θ = 0.5 and Θ = 0.5. These are fairly typical values. The lengths $N = 120$ and $N = 150$ in Table 1 are close to the lengths of the Census Bureau series we model.

Table 1. Thresholds $\Delta_N^{(i)}$ for the Four Model Types.

Thresholds	*4-coefficient*		*3-coefficient*	
	5-1	**4-2**	**5-1**	**4-2**
$\Delta_N^{(i)}, N = 120$	1.9	2.0	1.8	2.3
$\Delta_N^{(i)}, N = 150$	1.6	2.1	2.3	1.5

Note that the use of GMAIC always requires the fitting of the airline model to the series being modeled**.**

Remark 1. While the $p^{(i)}$ values are large relative to empirically chosen significance levels of tests like .05, they are more fundamental quantities than such empirical choices because of $\Delta AIC^{A,F}$'s unbiasedness property as an estimator of the accuracy difference of the two models in the Kullback-Leibler sense; see Akaike (1973) and Findley (1999).

3. Preferred Models for 75 Census Bureau Series

We fit the airline model and each of the four model sets defined by the 5-1(4), 4-2(4), 5-1(3) and 4-2(3) models to the logarithms of 75 Census Bureau series consisting of the value of shipments series from the monthly Survey of Manufacturers' Shipments, Inventories and Orders beginning in January 1992 and ending in September, 2001 (length $N = 117$) and of the Foreign Trade series from January, 1989 through December, 2000 (length $N = 144$). These are the Shipments and Foreign Trade series for which an airline model had originally been chosen over other standard ARIMA models for the given time span.[1] Table 2 below gives the breakdowns for MAIC and GMAIC by model type of the frequency-specific models that are preferred over the airline model. For GMAIC, the $\Delta_{120}^{(i)}$ values of Table 1 were used for the Shipments series and the $\Delta_{150}^{(i)}$ values for the Foreign Trade series. Excluded from preference were models with an estimated c_1 or c_2 equal to one. Some problems with such noninvertible models are discussed in Section 6.

Table 2. Numbers of Invertible Models of Each Type Preferred over the Airline Model by MAIC and GMAIC.

Models	*4-coefficient*		*3-coefficient*	
	5-1	**4-2**	**5-1**	**4-2**
MAIC preferred	8	15	24	43
GMAIC Preferred	4	7	9	24

The first row of Table 2 applies to 47 series and the second to 27. (For some series, more than one frequency-specific model is preferred over the airline model.) Thus the use of GMAIC in place of MAIC reduces the percentage of the 75 series for which a frequency-specific model is preferred from 63 percent to 36 percent.

Among the 27 series, a 4-2(3) model has the minimum AIC for 16 series and a 5-1(3) model has the minimum AIC for 6 series. Among the remaining 5 series, the 4-2(4) model has the minimum AIC for 4 series, and a 5-1(4) model has the minimum AIC for one series. For 7, 2, 2, and 1 of these models respectively, the largest or smallest seasonal spectral peak of the (differenced, log-transformed) modeled series occurred at a frequency associated with c_2. Thus the spectrum provides an interpretation of the GMAIC choice for almost half of the series, but for slightly more than half, the spectrum does not unambiguously indicate the distinctive nature of the frequency or frequency pair associated with c_2 (see Section 4 for two illustrative spectra). The three-coefficient models are by far the most favored of the frequency-specific models by GMAIC, being the preferred model for 22 of the 27 series, and therefore for 29 percent of the 75 series, a substantial percentage given that airline models were initially selected over other standard seasonal ARIMA models for these series.

As we shall discuss in Section 6, the four-coefficient models have an unexpectedly strong tendency for the estimate of c_1 or c_2 to be equal to one. A consequence is that only the three-coefficient models seem promising as frequency-specific generalizations of (2) for the purpose of the ARIMA model-based seasonal adjustment procedure of Hillmer and Tiao (1982) and Burman (1980) that we shall refer to as the AMB procedure.

For data $Z_t, 1 \le t \le N$ regarded as having an additive seasonal decomposition, most simply $Z_t = S_t + A_t$ with seasonal component S_t and a nonseasonal component A_t, the AMB procedure is able to decompose most seasonal ARIMA models for Z_t into the sum of a noninvertible ARIMA ("canonical") model for S_t and an ARIMA model for A_t. With these models, Gaussian conditional mean calculations can be used to obtain optimal linear estimates $\hat{A}_t$ of A_t that form the seasonally adjusted series

$$\hat{A}_t = \sum_{j=t-N}^{t-1} a_{t,j} Z_{t-j}, 1 \le t \le N,$$

(or $\exp(\hat{A}_t), 1 \le t \le N$ when the Z_t are the logs of the data).

Because of our interest in seasonal adjustment, for the remainder of the paper, we shall concentrate on the properties of the three-coefficient models.

4. Seasonal Adjustment Properties of Two GMAIC Three-Coefficient Models

Fig. 1 shows the Airline and GMAIC 5-1(3) models' AMB seasonal adjustments of the series U34EVS, *Shipments of Defense Communications Equipment* (January, 1992 through September, 2001) from the Census Bureau's monthly Manufacturers' Shipments, Inventories and Orders Survey. In the 5-1(3) model, the quarterly-effect frequency, 4 cycles/year, is associated with c_2. For this series $a = 0.6604$ and c_1 = 0.9867, a value slightly larger than the twelfth root of the

[1] These are the two major categories of Census Bureau series for which an interesting number of series had a lower AIC for model (3) than for model (1); see Findley, Martin and Wills (2002). For other major categories (Retail Trade, Construction), airline models usually had $\Theta^{1/12}$ very close to 1.

estimated airline model seasonal coefficient Θ ($\sqrt[12]{0.7798} = 0.9795$). By contrast, $c_2 = 0.8925$ ($= \sqrt[12]{0.2554}$). c_2's frequency has by far the largest seasonal peak in the spectrum of the modeled series, see Fig. 2. By contrast, the peaks at c_1's frequencies are small to non-existent. Thus there is compelling evidence for treating quarterly components differently from the other seasonal components, as well as evidence supporting the treatment of the remaining seasonal components in a uniform way. This is what the GMAIC 5-1(3) model does instead of treating all seasonal component uniformly the way the airline model does.

Basic features of each model's seasonal adjustment can be seen in the squared gain functions of the adjustment filters,

$$\left|\sum_{j=t-N}^{t-1} a_{t,j} \exp(i2\pi j\lambda)\right|^2, 1 \le t \le N \quad .$$

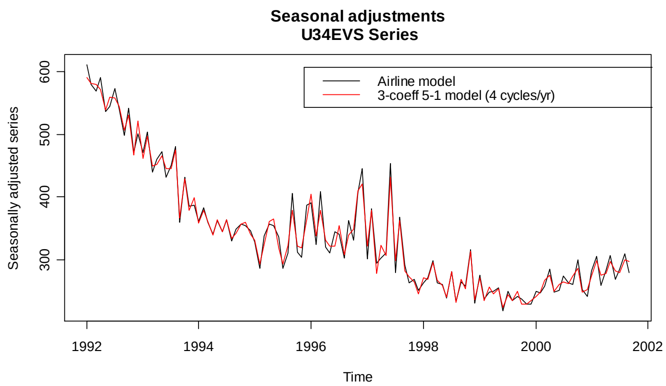


Figure 1. Airline and 5-1(3) model-based seasonal adjustment of Shipments (U34EVS). The GMAIC 5-1(3) model has c_2 assigned to the frequency 4 cycles/year.

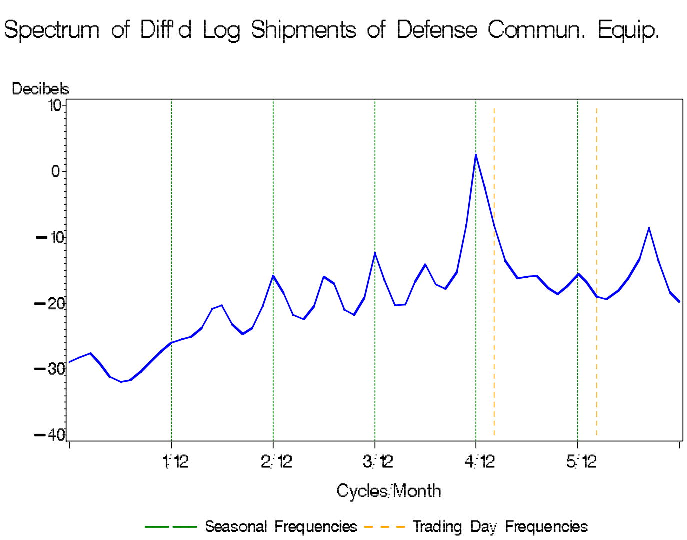


Figure 2. Spectrum of first-differenced logs of Shipments (U34EVS) showing a dominant peak at 4 cycles/year.

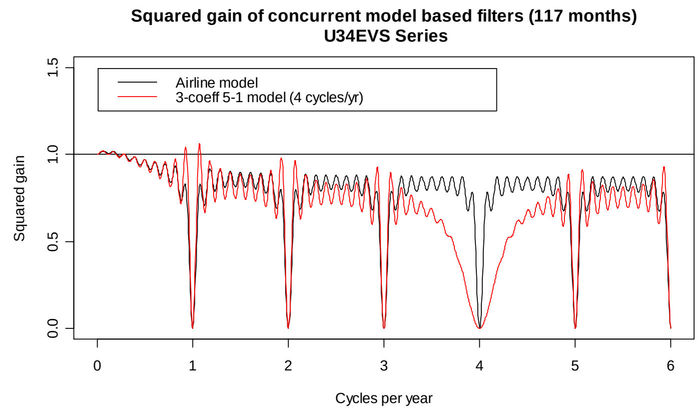


Figure 3. Squared gain of the finite concurrent model-based filters for logs of Shipments (U34EVS).

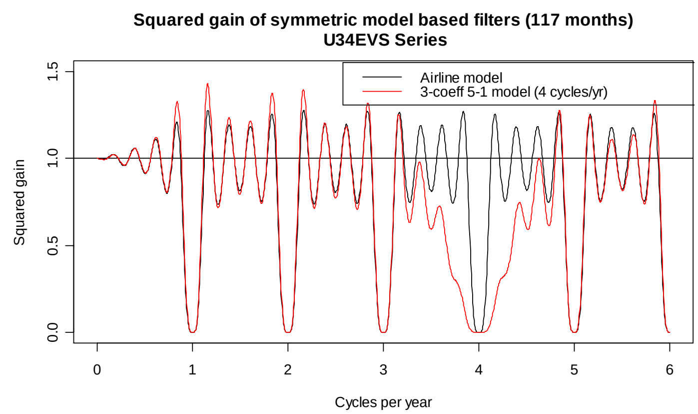


Figure 4. Squared gain of the finite central (symmetric) model-based filters for logs of Shipments (U34EVS).

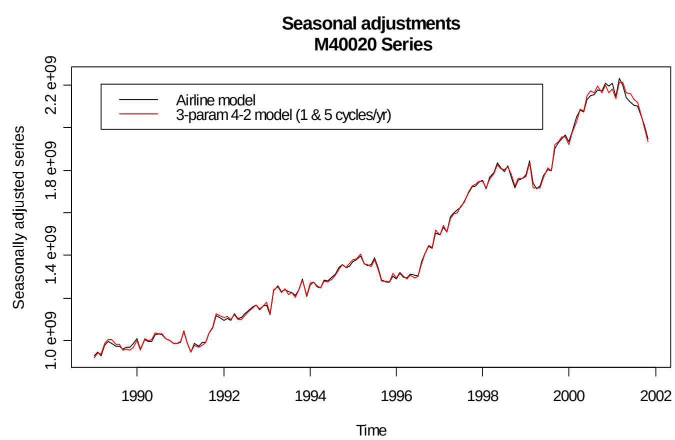


Figure 5. Airline and 4-2(3) model-based seasonal adjustments for Imports (M40020). The GMAIC 4-2(3) model has c_2 assigned to the. frequencies 1 and 5 cycles/year

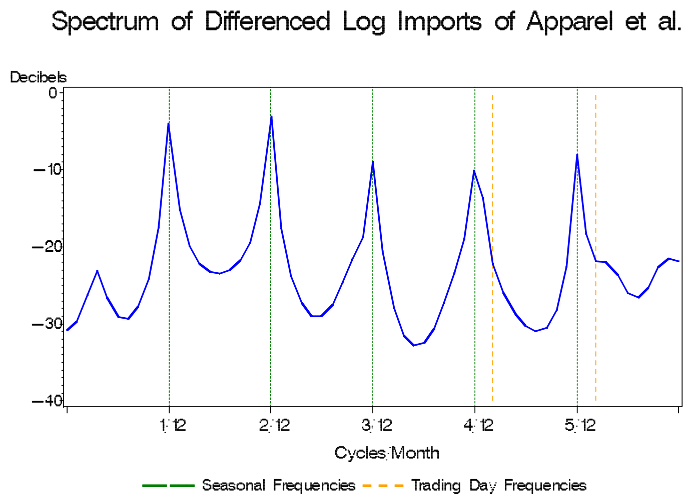


Figure 6. Spectrum of first-differenced logs of Imports (M40020). The reason for the pairing of the peaks at 1/12 and 5/12 with c_2 and the rest with c_1 isn't obvious.

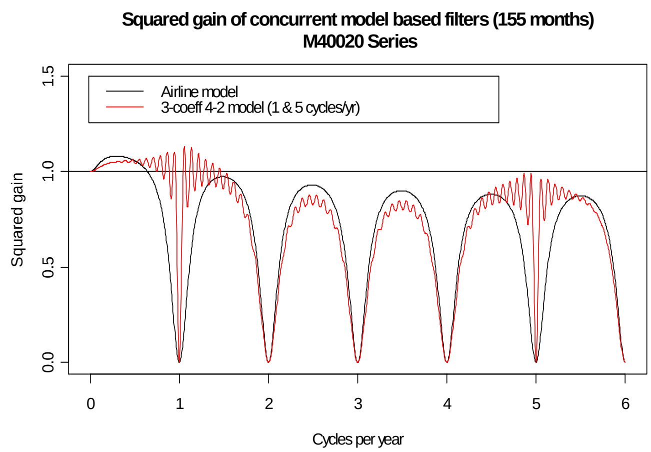
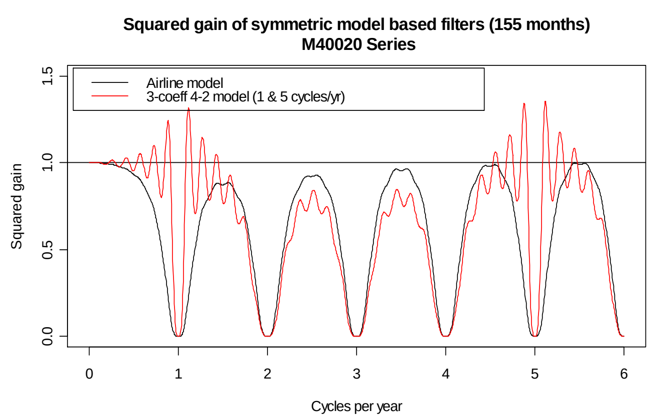


Figure 7. Squared gain of the finite concurrent model-based filters for logs of the Imports series M40020.



.

Figure 8. Squared gain of the finite symmetric model-based filter for logs of the Imports series M40020.

Fig. 3 shows the squared gains functions of the concurrent seasonal adjustment filters. These provide their model's seasonal adjustment of the most recent month, $\exp(\hat{A}_N)$. Fig. 4 is the analogue of Fig. 3 for the central month's adjustments, $\exp(\hat{A}_{[N/2]})$. The smaller value of c_2 results in the squared gains of the 5-1(3) model filters having wider troughs at the 4 cycles/year frequency than the airline model filters. The wider troughs indicate more suppression of variance components in the neighborhood of this frequency by the new model. Consequently, the 5-1(3) model's seasonal adjustment is smoother. (This is particularly visible in the last two years of the series in Fig 1.) Elsewhere, the squared gains of the airline and 5-1(3) model's filters are similar. (The rapid oscillations in the squared gains are essentially due to the high values of Θ and c_1 giving rise to filter coefficients that decay little over the relatively short length of the series. For further discussion, see Subsection 4.2.2 of Findley and Martin, 2003.)

Figs. 5-8 present graphs analogous to those of Figs. 1-4.for the preferred 4-2(3) model for the Census Bureau series M40020 of *Imports of Apparel and Other Household Textiles*. For this series the coefficient estimates are $a = 0.20$, $c_1 = 0.93$, and $c_2 = 0.99$. The coefficient c_2 is associated with the frequencies of one and five cycles per year. Its near unit value, compared with $\Theta^{1/12} = 0.94$ for the airline model, indicates that the 4-2(3) model finds the seasonal components at these frequencies much more stable than does the airline model. As a result, its squared gains have sharper troughs at these frequencies, effecting less suppression of variability, and its seasonally adjusted series is less smooth than that of the airline model. The frequency 1 cycle/year stands out in the spectrum plot in Fig. 6 as having almost the highest peak, and it has the deepest trough in the spectrum of the differenced and seasonally difference log data (not show), but the reason for its pairing via c_2 with frequency 5 cycles/year is not obvious, nor is the grouping of the other four seasonal frequencies with c_1.

Now we turn to forecasting properties.

5. Forecasting Performance

To obtain information about a model's h-step-ahead forecasting performance, some number of observations at the end of the series can be regarded as future data to be forecasted from a model fit to the earlier data. These forecasts can be compared to the actual series values (or, for series values identified as outliers, to the outlier-adjusted values). The span of modeled data can be increased one observation at a time, to produce a sequence of h-step-ahead forecast errors. Let $e_{A,h,t+h}$ denote the error of an airline model's forecast of Z_{t+h} from a model fitted to $Z_s, 1 \le s \le t$ and let $e_{G,h,t+h}$ denote the corresponding error of a specified generalized model. Given such errors for $t_0 \le t \le T-h$ for both models, we graph the differences of squared forecast errors

$$\sum_{s=t_0}^{t} \left\{ e^2_{G,h,s+h} - e^2_{A,h,s+h} \right\}, \quad t_0 \le t \le T-h, \tag{10}$$

and look for persistent upward movement or downward movement in the graph, the former indicating persistently better forecasts from the airline model, and the latter persistently better forecasts from the generalized model.

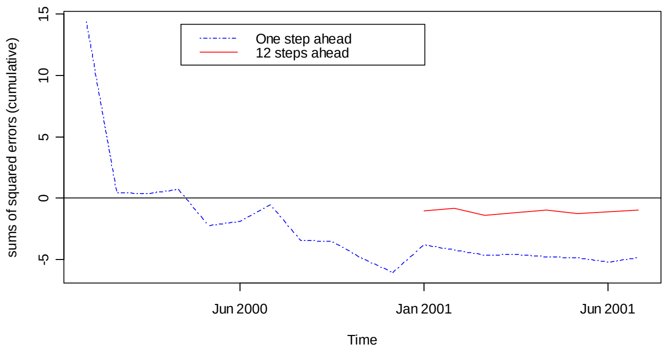


Figure 9. Out-of-sample square forecast error diagnostic (10) for series U34EVS comparing the three-coefficient 5-1 model and the airline model.

An example of this diagnostic graph is given in Figure 9, which shows the sums (10) of differences of h-step-ahead squared errors ($h = 1, 12$) of the 5-1(3) model and the airline model for the last years of the series U34EVS. The generally descending dotted line indicates that the one-step-ahead forecast performance of the 3-coefficient 5-1 model is persistently better than that of the airline model. The results for 12-step-ahead forecast performance are inconclusive, i.e. neither model is preferred. For more information about the out-of-sample forecast error diagnostic, see Findley, Monsell, Bell, Otto and Chen (1998) and Findley (2001).

We examined this diagnostic both for the 27 frequency-specific models preferred by GMAIC over the airline model and also for the 24 three-coefficient 4-2 models similarly preferred over the airline model. The results are summarized in the Tables 3 and 4. The tables show GMAIC preference does not always yield out-of-sample forecast performance as good as or better than that of the airline model, but it preponderantly does. Table 4 shows that the 4-2(3) model is the most broadly effective of the generalizations considered for forecasting.

Table 3. Comparative out-of-sample forecast performance between the airline model and the frequency-specific model most preferred over it by GMAIC for 27 series

Preferred Model	1-step forecasting	12-step forecasting
Frequency-Specific	10	7
Airline	5	3
None	12	17

Table 4. Comparative out-of-sample forecast performance between the airline model and the three-coefficient 4-2 preferred over it by GMAIC for the 24 series of Table 2.

Preferred Model	1-step forecasts	12-step forecasts
4-2(3)	13	8
Airline	3	4
None	8	12

6. Issues with estimates of c_1 or c_2 equal to one.

In addition to the 27 series discussed above, there were 18 others for which GMAIC preferred a frequency-specific model. The 18 series (and 26 preferred models among the four frequency-specific types) were excluded from Table 2 because the preferred model had one of its estimates of c_1 and c_2 equal to one. All but one of the GMAIC models excluded were four-coefficient models. (The exception was a 4-2(3) model for one series.) We now discuss some of the issues posed by unit estimates.

It is known that spurious unit coefficient estimates occur with positive probability for invertible seasonal moving average models, see Tanaka (1996). However, Tanaka's Table 8.2 (p. 313) of exact probabilities, which applies to simplified airline models with $\theta = 0$ (only Θ is estimated), shows that these models will have estimates of one with probability less than 0.05 when the true value satisfies $0 \le \Theta \le 0.9$ with series of the lengths we consider. Simulation experiments we have conducted by generating and estimating frequency-specific models, analogous those of the next Section, indicate that spurious coefficient estimates of one also occur for less than five percent with frequency specific models of each type we consider. Thus our observed percentage of unit estimates with the four-coefficient models with the 75 series is much higher than expected. (The percentage almost doubles if MAIC is used instead of GMAIC for comparisons with (2)).

A partial explanation, which Table 5 below seems to support, is that some of the series are well modeled by a frequency specific model with a unit coefficient. However, theoretical support is lacking for our use of MAIC or GMAIC for any of these 26 preferred models for which the unit coefficient estimate is correct. For such a model, the r.h.s of (4) has a factor of degree at least one that coincides with a factor of the differencing operator on the l.h.s of (4). This imparts to the model a fixed seasonal effect: the common divisor of the polynomials on both sides of the ARIMA equation, denoted $\delta^c(B)$, can be canceled from both sides and replaced by a periodic mean function $\mu(t)$ satisfying $\delta^c(B)\mu(t) = 0$. The resulting model is no longer a generalization of the airline model, and the large sample properties of maximum likelihood estimates, including rates of convergence, are quite different from the properties of such estimates for invertible models, see Tanaka (1996). Further, it is hard to conceive how such non-standard properties could cancel out in the AIC differences (7) in such a way that a shifted chi-square asymptotic distribution results.

Hence, in place of GMAIC, we turn to out-of-sample forecast error properties for confirmation of the models. Table 5 shows that for about forty percent of 26 series, the non-invertible GMAIC models have a forecast advantage over the airline model, whereas the converse result holds for about fifteen percent of the series.

Such forecasting analyses are time consuming and not easily automated. Further, it is not currently practicable to use a unit coefficient model, selected because of its forecasting advantage, for seasonal adjustment because no software is available to provide AMB seasonal adjustments from models with perfectly periodic seasonal components at some seasonal frequencies and evolving seasonality at others. Due to such complications, it appears that GMAIC-preferred three-coefficient models, which are common and seldom have unit coefficient estimates, are the models that usefully generalize the airline model for purposes of seasonal adjustment.

Table 5. Comparative out-of-sample forecast performance between the airline model and the 26 frequency-specific models preferred over it by GMAIC but excluded from Table 2 due to a unit value estimate of c_1 or c_2.

Preferred Model	1-step forecasting	12-step forecasting
Frequency-Specific	10	10
Airline	4	3
None	12	13

7. Estimation Variability of c_2 for 3-coefficient Models

We observed in simulation results not presented here that the variability of estimates of c_2 in the frequency-specified models is substantially greater than that of c_1. Intuitively, this suggests c_1 gains stability by estimating more frequency components than c_2. Here we demonstrate that the variability of c_2 is also tied to the number of frequency components it estimates by showing histograms of the estimates of c_2 for 5-1(3) and 4-2(3) models. We generated 1000 realizations of length 150 of 5-1(3) and 4-2(3) models with true coefficient values $a = 0.50$, $c_1 = 0.96$, and $c_2 = 0.93$. (These are average values of the coefficients of a set of 21 MAIC preferred 5-1(3) models, of which 13 had c_1 or $c_2 = 1$.) The histogram of the c_2 estimates of the 5-1(3) model is given in Figure 10. For 4% of the realizations, the c_2 coefficient was estimated as unity. Fig. 11 shows the histogram of c_2 estimates from the 4-2(3) model. Only 0.5% of the estimates are unity and the tails of the histogram are thinner than in Fig. 10, indicating less variability.

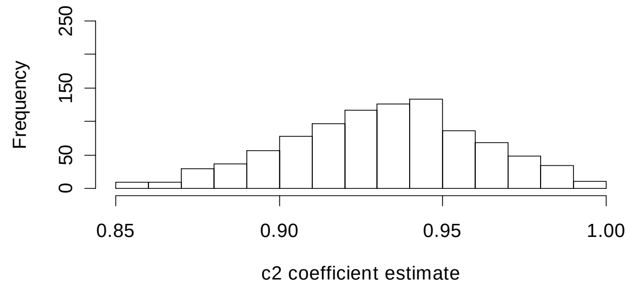


Figure 10. Distribution of c_2 estimates for 1000 realizations of the 3-coefficient 5-1 model with coefficients $a = 0.50$, $c_1 = 0.96$, and $c_2 = 0.93$. Forty coefficient estimates were one. Sixteen were less than 0.85.

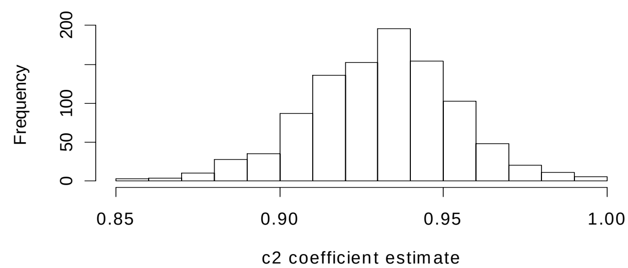


Figure 11. Distribution of c_2 estimates for 1000 realizations of the 3-coefficient 4-2 model with coefficients $a = 0.50$, $c_1 = 0.96$, and $c_2 = 0.93$. Five estimates were one.

8. Conclusions

Airline model-based AMB seasonal adjustment filters treat variance components around all seasonal frequencies in a similar way. However, spectrum estimates like that of Fig. 2 demonstrate the unsurprising fact that that seasonal economic series do not always have similar variance components at all seasonal frequencies. In this paper, we have examined generalizations of the airline model that divide the seasonal frequencies into two groups and provide for different treatment of each group with seasonal adjustment. The use of two groups chosen by our generalization of Akaike's MAIC procedure was also shown to often lead better out-of-sample forecast than the airline model.

Our three-coefficient airline model generalizations were preferred by GMAIC much more often than the four-coefficient generalizations. They are also much less likely to have coefficient estimates of one, which are problematic for seasonal adjustment at the present time. Deciding how often such unit root estimates are spurious is a topic needing further research.

Another topic for investigation is whether, selecting between three parameter models and (1), instead of using a separate set $\mathcal{F}(i)$ to define separate thresholds $\Delta_N^{(i)}$ for the 5-1(3) and 4-2(3) models, a single set $\mathcal{F}(i)$ containing all 21 three-coefficient models (22 models if (3) is included) to define a unified threshold depending only on N and on the estimated airline model coefficients: for implantation for seasonal adjustment, we plan to use a more refined GMAIC procedure in which the $\Delta_N^{(i)}(\theta,\Theta)$ values used are taken from stored table covering a grid of (θ,Θ) pairs with $\theta,\Theta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ instead of just the $\Delta_N^{(i)}(0.5, 0.5)$ values of Table 1. The $\Delta_N^{(i)}(\theta,\Theta)$ value used for a given series would be the one in the table whose N is closest to the series' length and whose (θ,Θ) is closest to the parameter vector of the estimated airline model of the series.

Acknowledgements

The authors are indebted to Kathleen McDonald-Johnson for comments that improved an earlier draft of this paper and to Tucker McElroy for computing help.

References

H. Akaike (1973), "Information theory and an extension of the maximum likelihood principle," in *Proceedings of the 2nd International Symposium on Information Theory* (B. N. Petrov and F. Csaki Eds.) pp. 267--281, Akademiai Kiado, Budapest (1973). Reprinted in *Brealthroughs in Statistics* (S. Kotz and N. L.. Johnson, Eds.) pp. 610-624 Springer-Verlag (1922) and in "Selected Papers of Hirotugu Akaike" (E. Parzen, K. Tanabe, G. Kitagawa, Eds.), pp. 199-214, Springer, New York, 1998.

Burman, J. P. (1980), "Seasonal Adjustment by Signal Extraction," *Journal of the Royal Statistical Society*, Ser. A , 143, 321-337.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control* (revised edition), Holden Day, San Francisco.

Doornik, J. A. (2001), *Object-Oriented Matrix Programming using Ox 3.0.* London: Timberlake Consultants Press.

Findley, D. F. (1999), "AIC II". in *Encyclopedia of Statistical Science, Update Volume 3*. (S. Kotz, C. R. Read, D. L. Banks Eds.) pp. 2--6, Wiley, New York, 1999.

Findley, D. F. (2001). Asymptotic Stationarity Properties of Out-of-Sample Forecast Errors of Misspecified RegARIMA Models. Proceedings of the Business and Economic Statistics Section of the American Statistical Association, CD-ROM. Also available from http://www.census.gov/ts/papers/osfeasap.pdf

Findley, D. F. and Martin, D. E. K. (2003). "Frequency domain analyses of SEATS and X-11/12-ARIMA seasonal adjustment filters for short and moderate-length time series," Research Report S2003-02, Statistical Research Division, U.S. Census Bureau, http://www.census.gov/srd/papers/pdf/rrs2003-02.pdf

Findley, D. F., Martin, D. E. K. and Wills, K. C. (2002). "Generalizations of the Box-Jenkins Airline Model," Proceedings of the American Statistical Association, Business and Economic Statistics Section [CD-ROM], Alexandria, VA: American Statistical Association.

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C. and Chen, B.-C. (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program," *Journal of Business and Economic Statistics*, **16**, 127-152.

Fischer, B. and C. Planas (2000), "Large scale fitting of regression models with ARIMA errors," *Journal of Official Statistics* 16, 173-184.

Hillmer, S. and Tiao, G. C. (1982), "An ARIMA Model-Based Approach to Seasonal Adjustment", *Journal of the American Statistical Association*, 77, 63-70.

Koopman, S. J., N. Shephard, and J. A. Doornik (1999). "Statistical algorithms for models in state space form using SsfPack 2.2." *Econometrics Journal* 2, 113-66.

Tanaka, K. (1996). *Time Series Analysis,* Wiley, New York.

Taniguchi, M. and Kakizawa, Y. (2000), *Asymptotic Theory of Statistical Inference for Time Series.* Springer-Verlag, New York.

Appendix. Estimation of the Generalized Models.

The generalized airline models are defined in terms of products of moving average factors of degrees one or two rather than in terms of the full MA polynomial of degree $s+1$. The latter is needed for the state space representation used to calculate the likelihood function and also to calculate the gain functions and seasonal adjustments, see Durbin and Koopman (2001) for more details on such calculations. The full MA polynomial could be obtained from the factors by coding a routine to carry out polynomial multiplication. However, Fast Fourier Transform functions are available in Ox and similar software, and these can be used to transform a product of polynomial factors into the coefficient sequence of the product polynomial (effectively, the convolution of the factors' coefficients). Once the full MA polynomial is available, there are routines to produce the ARIMA model's state space representation and implement filtering and smoothing algorithm for it to obtain maximum Gaussian likelihood values and AMB seasonal adjustments.