GENERALIZATIONS OF THE BOX-JENKINS AIRLINE MODEL

David F. Findley, Donald E. K. Martin and Kellie C. Wills, Census Bureau
Donald. E. K. Martin, Census Bureau, Washington DC 20233 and Howard University, Washington DC 20059

Key Words: Time Series, Seasonal Adjustment, Forecasting, Filters, Phase Delay, Squared Gain

1. INTRODUCTION

For modeling monthly totals of passengers in international air travel and other seasonal time series, Box and Jenkins (1976) developed a two-coefficient time series model of factored form that is now known as the airline model. For a seasonal time series Z_t with $s \geq 2$ observations per year, this model is given by

$$(1-B)(1-B^s)Z_t = (1-q\,B)(1-\Theta B^s)e_t^{(1)} \quad ,$$

where B denotes the backshift operator, i.e. $BZ_t = Z_{t-1}$. Throughout the paper, for any i, $e_t^{(i)}$ denotes a sequence of independent variates with mean zero and variance s_i^2. The parameters satisfy $-1 \leq q, \Theta \leq 1$, and for economic time series, usually $\Theta \geq 0$. The airline model is by far the most widely used model for monthly and quarterly macroeconomic time series. This broad usage raises concerns that this model is overused and suggests that related but more general models should be investigated as alternatives. We present a limited study of two generalizations, focusing on their usefulness for seasonal adjustment, particularly for the ARIMA-model-based (AMB) signal extraction method of Hillmer and Tiao (1982).

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

$$(1-B)^2(\sum_{j=0}^{s-1} B^j)Z_t = (1-q\,B)(1-\Theta^{1/s}B)(\sum_{j=0}^{s-1}\Theta^{j/s}B^j)e_t^{(1)}\,. \qquad (1)$$

From this representation, one sees that $\Theta^{1/s}$ occurs in both the nonseasonal and seasonal polynomials, $(1-q\,B)(1-\Theta^{1/s}B) = 1-(q+\Theta^{1/s})B + q\cdot\Theta^{1/s}B^2$ and $(\sum_{j=0}^{s-1}\Theta^{j/s}B^j)$, respectively, of (1). This may partially explain why values of the seasonal moving average parameter Θ can substantially influence the trend component of solutions of (1) (see Section 4 of Findley and Martin 2002).

Our first generalization is obtained by substituting a general second-degree moving average (MA(2)) polynomial for $(1-q\,B)(1-\Theta^{1/s}B)$ in (1), yielding what we here call the *generalized airline model*,

$$(1-B)^2(\sum_{j=0}^{s-1} B^j)Z_t = (1-a_2B+b_2B^2)(\sum_{j=0}^{s-1} c_2^j B^j)e_t^{(2)}, \quad (2)$$

in which the airline model's nonseasonal polynomial factor is generalized. When the roots of the MA(2) polynomial on the right in (2) are real, (2) is equivalent to what we call *restricted generalized airline model*,

$$(1-B)^2(\sum_{j=0}^{s-1} B^j)Z_t = (1-a_3B)(1-b_3B)(\sum_{j=0}^{s-1} c_3^j B^j)e_t^{(3)}\,. \quad (3)$$

Another generalization of the airline model is obtained by noting that the factor $(1-q\,B)(1-\Theta B^s)$ in (1) can be written as

$$1-q\,B-\Theta B^s + q\cdot\Theta B^{s+1}\,.$$

Using instead a moving average polynomial with no constraint on the coefficient of B^{s+1}, we obtain

$$(1-B)^2(\sum_{j=0}^{s-1} B^j)Z_t = (1-a_4B-b_4B^s+c_4B^{s+1})e_t^{(4)}\,. \qquad (4)$$

We call model (4) the *1-12-13 model*, because in the case of monthly data, i.e. when $s=12$, the powers of B that occur on the right are 1, 12 and 13.

Our primary interest in this research is the use of the new models for seasonal adjustment. In the next section, we present series for which there is support for use of one or more of these new models in place of the airline model. In Section 3, for three of the series identified in Section 2, we compare frequency domain properties of competing models' canonical seasonal adjustment filters, specifically their central (symmetric) and one-sided (concurrent) adjustment filters. We also compare their seasonal adjustments.

2. SERIES FOR WHICH A NEW MODEL IS COMPETITIVE WITH THE AIRLINE MODEL

We consider two categories of time series. The first is series for which the estimated airline model fails to provide an AMB decomposition of the data into a sum of a seasonal and a nonseasonal series, whereas a new model does. The second category is series for which at least one of the fitted new models is preferred over the airline model by Akaike's minimum AIC criterion and, in many cases, provides an alternative seasonal adjustment with desirable features.

2.1 Using a new model to obtain seasonal decompositions

In the AMB approach, the fitted model's (generalized) autocovariance generating function (a.g.f.) must be decomposable into a sum of a.g.f.'s of ARIMA models appropriate for modeling seasonal, trend, and irregular components, for example, a nonnegative constant a.g.f. for modeling the irregular as white noise. Most ARIMA model types have parameter values for which no such decomposition exists. With such values, the model is "inadmissible." We fit the new models (2) and (4) to two series whose fitted airline models obtained in Fiorentini and Planas (2001) are inadmissible. The series are the quarterly Index of French Production of Wines (FRWPI) (1986 quarter 1 – 2000 quarter 1), and the log of the monthly Index of Domestic Turnover in the Italian Manufactures of Textiles (ITTMTI) (November 1991 – December 1999). For both, the fitted generalized airline model (2) was admissible, but not the fitted models (1) and (4). Thus, an AMB seasonal adjustment is obtainable from (2), but not from (1) and (4). Table 1 shows model coefficients. AIC values (not shown) favor the airline model over the generalized model for both series. Other examples of statistically preferred but inadmissible models are given below. (AMB seasonal adjustments from "suboptimal" models can be satisfactory in the sense of having no residual seasonality and changing little with additional or moderately revised data.)

2.2 Series for which a new model is preferred over the airline model by AIC

We fit the airline (1), generalized airline (2) and 1-12-13 (4) models to 111 Census Bureau monthly time series. Of these, 39 are foreign trade (import/export) series, 36 are construction series, and 36 are M3 series (manufacturers' shipments, inventories and orders). Table 2 gives a breakdown of the 23% of these series for which either model (2) or (4) is preferred over the airline model by AIC. Recall that if the MA(2) roots are real, models (2) and (3) coincide. Of the 16 series for which AIC preferred the generalized airline model (2), the MA(2) roots are complex in 4 cases. All of these 16 models are admissible.

In Table 3, we list parameter estimates for the series included in Table 2 for which model (3) and, in one case also (4), is preferred by AIC. The table shows how a_3 differs from q and how b_3 and c_3 differ from $\Theta^{1/12}$. Model (3) was proposed a decade ago by William Bell who thought the $(1-B)^2$ on the left in (1) would often cause overdifferencing and therefore yield $b_3 = 1$ in (3), which Table 3 affirms. (When $B = 1$ is a zero of the moving average polynomial in (3) or (4), the model can be reduced by canceling a factor of $1-B$ from both sides of the model equation and adding a constant on the right. However, for the series for which this occurred, we found no advantage from using a reduced model.) Table 4 lists parameter estimates for the series for which the generalized airline model is preferred by AIC and has complex roots. It shows how a_2, b_2 and c_2 compare to $q + \Theta^{1/s}$, $q \cdot \Theta^{1/s}$ and $\Theta^{1/s}$, respectively. For the series with c_2 (or c_3 or Θ) equal to 1, the implied seasonal component is deterministic and not obtainable from the AMB method. Table 5 lists parameter estimates for the series for which AIC preferred the 1-12-13 model and for which this model is admissible, showing how a_4 differs from q, b_4 differs from Θ and c_4 differs from the product $q \cdot \Theta$. In these tables, many of the coefficients being compared seem to differ little (e.g. $\Theta^{1/s}$ and c_2 or c_3), but the AIC differences in favor of a generalized model with three coefficients indicate that the models' coefficient vectors (q, Θ) and (a, b, c) differ significantly.

Among the six series for which the 1-12-13 model was preferred but not admissible, five were long enough that the out-of-sample forecast diagnostic described in Findley et al. (1998) could be used to decide if the 1-12-13 model offered better forecasts than the airline model and therefore potentially better X-12-ARIMA seasonal adjustments. For three of these five, the diagnostic firmly supported the choice of the 1-12-13 model. For the other two, the new models' forecasts were persistently better in the last year but persistently worse in several preceding years.

3. FREQUENCY DOMAIN PROPERTIES OF FITTED MODEL'S SEASONAL ADJUSTMENT FILTERS

Before comparing frequency domain properties of the seasonal adjustment filters associated with the models of a few of the series identified in the last section, we give some background material. For a linear filter for Z_t with output $Y_t = \sum_j C_j Z_{t-j}$, the *frequency response function* of the filter is

$$C(l) = \sum_j C_j e^{-i\frac{2p}{12} j l}, \quad -6 < l \le 6,$$

when l is in units of cycles per year. The amplitude $G(l) = |C(l)|$ is called the *gain function* of the filter.

Table 1. Airline model and generalized airline model fits to the Fiorentini and Planas (2001) series (The roots of the MA(2) polynomials of the generalized model are complex for ITTMTI and real for FRWPI).

	Airline model		*Generalized airline model*		
Series	q	Q	a_2	b_2	c_2
ITTMTI	0.43	-0.31	1.268	0.456	0.727
FRWPI	0.55	-0.22	0.749	0.017	0.222

Table 2. Number of series for which either model (2) or model (4) is preferred over the airline model

Type of Series	**No. of series with model (2) preferred (by AIC)**	**No. of series with model (4) preferred and admissible (inadmissible)**
Foreign Trade	6 of 39	3 (5) of 39
Construction	1 of 36	0 (1) of 36
Manufacturing	9♦ of 36	1♦ (0) of 36

¨ The series U33LVS occurs in both categories

Table 3. Airline and restricted generalized airline model coefficients for series for which the generalized airline model is preferred by AIC and coincides with the restricted generalized model.

	Airline model		*Restricted generalized airline model*		
Series	q	$Q/Q^{1/12}$	a_3	b_3	c_3
M3000C	0.346	0.476/ 0.940	0.305	1.000	0.934
M3010C	0.481	0.704/ 0.971	0.468	1.000	0.967
M40020	0.210	0.457/ 0.937	0.172	1.000	0.934
M40040	0.304	0.596/ 0.958	0.296	1.000	0.952
X3022C	0.341	0.580/ 0.956	0.305	1.000	0.947
X40000	0.251	0.640/ 0.963	0.372	0.902	0.974
SL0844	0.545	1.000/ 1.000	0.732	0.834	1.000
U34KTI	-0.122	0.668/ 0.967	0.202	0.645	0.965
U39BVS	0.444	0.272/ 0.897	0.372	0.987	0.885
U33KVS	0.475	0.596/ 0.958	0.602	0.874	0.965
U32SWI	0.057	0.809/ 0.983	0.311	0.734	0.976
U33LVS	0.510	0.679/ 0.968	0.594	0.929	0.977

Table 4. Airline and generalized airline model coefficients for series for which the generalized airline model is preferred by AIC and has complex MA(2) roots

Series	q	Q	$q+Q^{1/12}$	$q \times Q^{1/12}$	$Q^{1/12}$	a_2	b_2	c_2
U36FVS	0.691	0.925	1.685	0.687	0.994	1.725	0.749	1
U34EVS	0.715	0.780	1.695	0.701	0.979	1.769	0.810	1
U36CVS	0.289	0.314	1.196	0.261	0.908	1.349	0.499	0.925
U34DVS	0.440	0.496	1.384	0.415	0.943	1.438	0.578	0.944

Table 5. Airline and 1-12-13 model coefficients for series for which the preferred 1-12-13 model is admissible

	Airline model			*1-12-13 model*		
Series	q	Q	$q \times Q$	a_4	b_4	c_4
X41020	0.418	0.635	0.266	0.357	0.659	0.425
M21610	0.559	0.732	0.409	0.484	0.735	0.496
M12060	0.800	0.380	0.304	0.846	0.405	0.251
U33LVS	0.510	0.679	0.346	0.519	0.740	0.501

A function $\phi(\lambda)$ that satisfies $C(\lambda) = G(\lambda) e^{i\frac{2\pi}{12}\phi(\lambda)}$ when $C(\lambda) \neq 0$, such as

$$\phi(\lambda) = (12/2\pi) A\mathrm{rctan}[\mathrm{Im}\left(C(\lambda)\right) / \mathrm{Re}\left(C(\lambda)\right)],$$

is called the *phase function* of the filter. When $C(\lambda) = 0$, $\phi(\lambda)$ is undefined. For the filters we consider $C(0) = 1$ so $\phi(0) = 0$, and for $\lambda \neq 0$, the *phase delay*

$$\tau(\lambda) = -\phi(\lambda)/\lambda,$$

measures the time delay (or time advance if $\tau(\lambda) < 0$) induced by the filter on the λ-frequency component of the input time series. For example, for the *k-th* power of the backshift operator, $B^k Z_t = Z_{t-k}$, for $k = 0, \pm 1, \pm 2 ...$, the frequency response function is $e^{-i\frac{2\pi}{12}k\lambda}$ so $\tau(\lambda) = k$.

The squared gain function $G(\lambda)^2$ has the important property that if Z_t is a stationary time series with spectral density $f_Z(\lambda)$, then the spectral density $f_Y(\lambda)$ of the filter output series Y_t is

$$f_Y(\lambda) = G(\lambda)^2 f_Z(\lambda). \qquad (5)$$

Thus the squared gain function measures the extent to which the filter increases, decreases, or leaves unchanged the contribution to variance of each frequency component of the input series. An approach to an analogue of (5) for nonstationary ARIMA processes is discussed in Section 3 of Findley and Martin (2002).

The gain and phase delay functions are even, $G(-\lambda) = G(\lambda)$, $\tau(-\lambda) = \tau(\lambda)$ ($\phi(\lambda)$ is odd), so only frequencies $0 \leq \lambda \leq 6$ need be considered. In fact, for the phase delay function, because we are mainly interested in the delays associated with turning points and business cycle movements, we shall only plot phase delay for periods greater than a year, i.e. for frequencies $0 \leq \lambda < 1$.

There are various perspectives from which the squared gain and phase delay plots can be evaluated (see Section 3 of Findley and Martin 2002 for a more detailed discussion than is given here). If one of the new models is considered the correct model for the data, then from the AMB perspective, the associated seasonal adjustment filter must be considered mean square optimal, regardless of its frequency domain properties. On the other hand, a user of a seasonal adjustment filter may be interested in whether the filter enhances or diminishes variation in the series at various frequencies, or in the amount of delay in detecting changes in trend (e.g. turning points) induced by the filter. From the Digital Signal Processing perspective, a seasonal adjustment filter should have a gain function that is near zero at the seasonal frequencies and one at nonseasonal frequencies, perhaps with some compromise to obtain, e.g. less phase delay. In practice, a synthesis of both perspectives is appropriate for the AMB approach due to the approximate nature of ARIMA models for economic data and to desired features (e.g. smoothing or rapid detection of turning points) with no direct connection to mean square optimality.

Figures 1-3 present the squared gain and phase delay plots of the concurrent seasonal adjustment filter and the squared gain of the symmetric seasonal adjustment filter for the airline model and the new model(s) preferred over the airline model by AIC. The figures also show the seasonal adjustments. Figure 1 is for the series U36CVS (Unfilled orders of heavy duty truck manufacturing), for which the generalized model (2) is the preferred model; Figure 2 corresponds to the series X41020 (Exports of cookware, cutlery, house and garden ware), for which the 1-12-13 model (4) is preferred; and Figure 3 is for the series U33LVS (Unfilled orders of pump and compressor manufacturing), the series for which both models (3) and (4) are preferred over the airline model.

For both the concurrent and symmetric filters, at the higher frequencies, the squared gains associated with the restricted generalized airline model (3) are closer to one than those of the corresponding airline models, representing a more neutral treatment of the corresponding frequency components and less smoothing. This is in contrast to the squared gains associated with the filters from model (4), which are significantly farther below one at these frequencies, indicating a much greater suppression of higher frequency components and hence more smoothing. The phase delays associated with model (3) are less than those of airline model, which in turn are usually less than the phase delays associated with model (4), although the differences are not large at the most important low frequencies associated with not too rapid trend movements.

The graphs of seasonal adjustments from the airline and new models are not discernibly different in many months because differences in seasonal adjustments of less than 1% are not visible on the graphs. For months when there is a visible difference, the seasonal adjustments confirm the general interpretations of the squared gain plots made above with regard to smoothing. For example, the models for which the squared gains indicate a greater suppression of higher frequency components away from seasonal frequencies are the models with smoother seasonal adjustments; see especially Figures 2-3. Comparison of the squared gain and phase delay plots of the concurrent filters shows that greater suppression of low frequency components (or less enhancement when the gain exceeds one) leads to greater phase delay of these components. For all of the series for which the 1-12-13 model was preferred, its seasonal adjustment filters provided greater smoothing but with greater phase delay at all but the lowest trend frequencies. Usually (but not always) the preferred models (2) and (3) offered less phase delay and less smoothing than the airline model.

In future work, we shall investigate more general models than (2) which also generalize the seasonal polynomial factor of the model (2). We further plan to examine empirically the costs of greater smoothing in terms of greater revisions with new data and delays in detecting trend movements.

ACKNOWLEDGEMENTS

The authors are indebted to Siem Jan Koopman and John Aston for developing software that enabled us to estimate the new models, and obtain filter coefficients and seasonal adjustments, and to John Aston, Kathleen McDonald-Johnson and Brian Monsell for comments that improved this paper.

REFERENCES

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

Findley, D. F. and Martin, D. E. K. (2002). “Frequency domain analysis of SEATS and X-11/12-ARIMA seasonal adjustment filters for short and moderate-length time series,” *Proceedings of the 3rd International Symposium on Frontiers of Time Series* (ed. Y. Kawaski), 95-120, Tokyo: Institute of Statistical Mathematics.

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(2), 127-177.

Fiorentini, G. and Planas, C. (2001). “Overcoming Nonadmissibility in ARIMA-Model-Based Signal Extraction,” *Journal of Business & Economic Statistics*, 19(4), 455-464.

Hillmer, S. C. and Tiao, G. C. (1982). “An ARIMA model-based approach to seasonal adjustment,” *Journal of the American Statistical Association*, 77, 63-70.

Figure 1

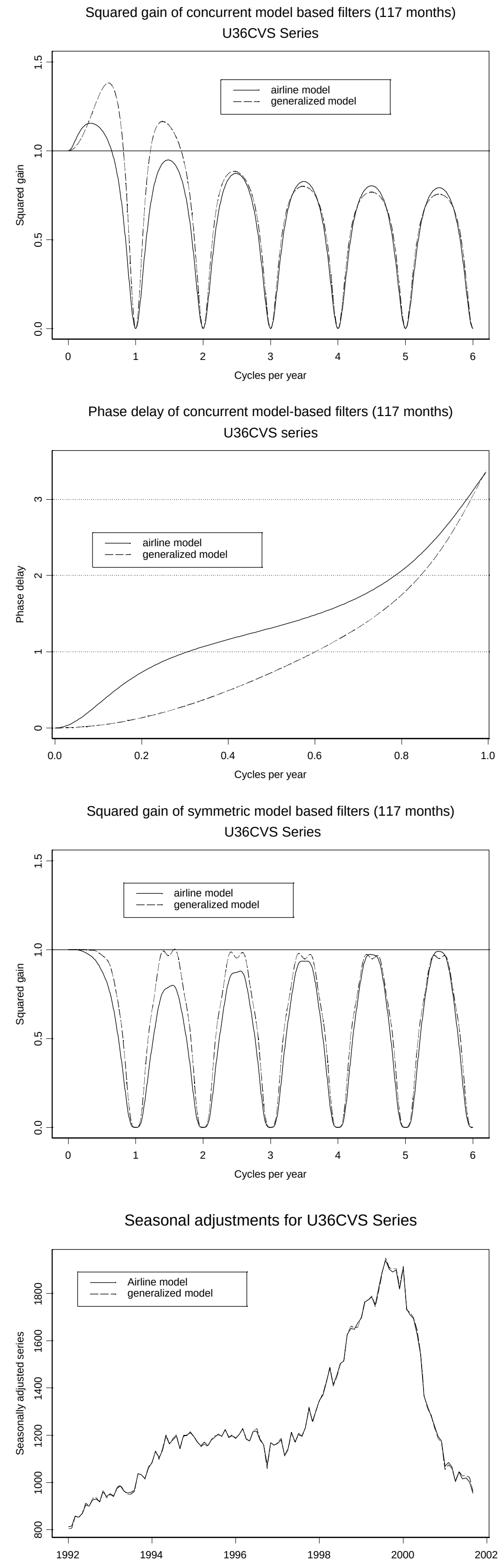

Figure 2

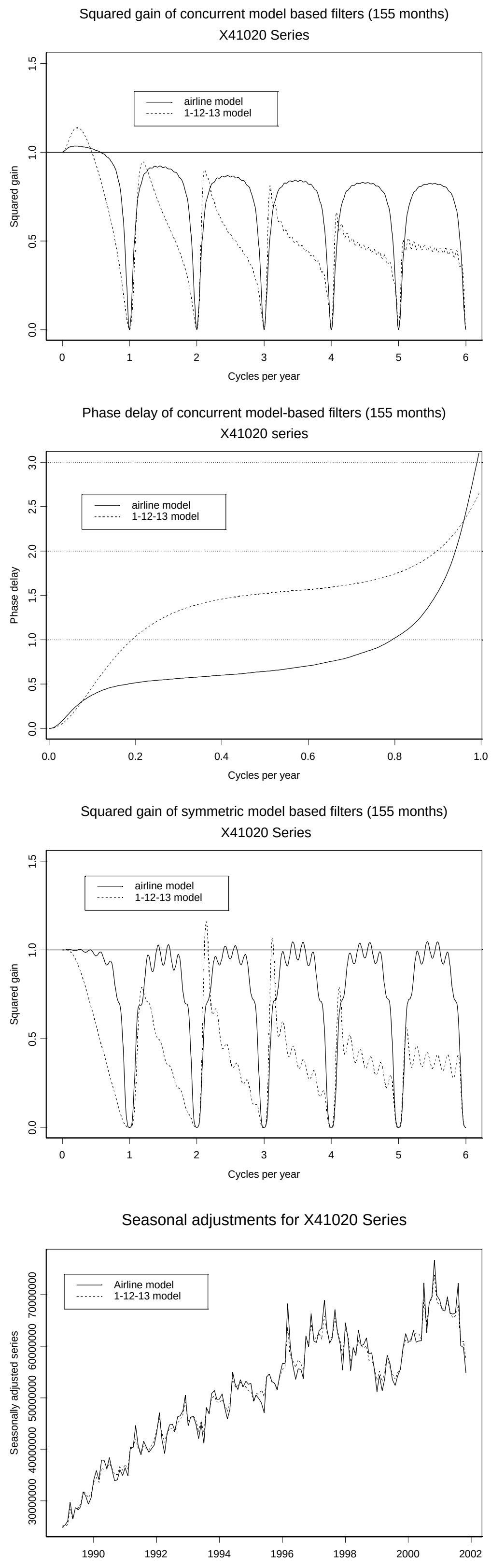

Figure 3