ELSEVIER

Journal of Econometrics 92 (1999) 325–353

JOURNAL OF Econometrics

www.elsevier.nl/locate/econbase

Estimation error and the specification of unobserved component models

Agustín Maravall[a,*], Christophe Planas[b]

[a] *Research Department, Bank of Spain, Alcala 50 28014 Madrid, Spain*
[b] *Institute for Systems, Informatics and Safety, Joint Research Centre of the European Commission, TP 361, 21020 Ispra(VA), Italy*

Received 1 June 1996; received in revised form 1 September 1998; accepted 8 October 1998

Abstract

The paper deals with the problem of identifying stochastic unobserved two-component models, as in seasonal adjustment or trend-cycle decompositions. Solutions based on the properties of the unobserved component estimation error are considered, and analytical expressions for the variance of the errors in the final, preliminary, and concurrent estimators are obtained. These expressions are straightforwardly derived from the ARIMA model for the observed series.

The estimation error variance is always minimized at a canonical decomposition (i.e., at a decomposition with one of the components noninvertible), and a simple procedure to determine that decomposition is presented. On occasion, however, the most precise final estimator may be obtained at a canonical decomposition different from the one that yields the most precise preliminary estimator. Two examples are presented. First, a simple 'trend plus cycle'-type model is used to illustrate the derivations. The second example presents results for a class of models often encountered in actual time series.

JEL classification: C22; C51

Keywords: Seasonal adjustment; Unobserved component models; Signal extraction; ARIMA models; Identification; Estimation error

We consider the problem of decomposing an observed series into the sum of two orthogonal components, each one the output of an unobserved linear stochastic process parametrized as an ARIMA model. Thus the basic model (presented in Section 1) is that of an observed ARIMA model with unobserved

* Corresponding author. Tel.: + 34 1 338 54 76; fax: 34 1 338 51 93; E-mail: maravall@bde.es.

0304-4076/99/$ - see front matter
PII: S0304-4076(98)00094-3

ARIMA components. Examples are the seasonally adjusted (SA) series plus seasonal component decomposition of economic series, the trend-plus-cycle decomposition often used in business cycle analysis, and, in general, signal-plus-noise type of decompositions. The analysis centers on minimum mean squared error (MMSE) estimators of the unobserved components.

It is well known that, given the ARIMA model for the unobserved series, the decomposition into unobserved components presents an identification problem, which stems from the fact there is in general an amount of white-noise variation that can be arbitrarily allocated between the two components (see, for example, Bell and Hillmer, 1984; Watson, 1987). This identification problem is discussed in Section 2. Several solutions have been proposed. In the so-called ARIMA Model Based (AMB) approach, which derives the models for the components from the ARIMA model for the observed series, identification is reached by assigning all noise to one of the components, so as to make the other one noninvertible. This approach has been mostly developed in the context of seasonal adjustment, and basic references are Burman (1980) and Hillmer and Tiao (1982). In the so-called structural time series (STS) approach, which directly specifies the models for the components, identification is achieved by a priori restricting the order of the MA polynomials. This approach is heavily used in econometric applications, and basic references are Engle (1978) and Harvey and Todd (1983). (Additional solutions to the identification problem can be found in Watson (1987) and Findley (1985)). The fact remains, however, that there is no universally accepted criterion to reach identification in unobserved component models. Yet, as shown in Tiao and Hillmer (1978) and Hillmer and Tiao (1982), the assumptions made to reach identification affect the properties of the component estimators. In this paper, we analyse the effect on the component estimation error. Burridge and Wallis (1985) within the STS approach, and Hillmer (1985) within the ABM approach, have provided algorithms for computing the variance of the components estimation error. We follow an alternative approach, close to the one in Watson (1987), which provides simple analytical expressions for the variances of the components estimation error, for different admissible decompositions.

When choosing between two admissible decompositions that only differ in the allocation of white noise to the components, one relevant consideration could be the precision of the associated estimators. There are, however, several types of estimators, depending on the available information. For periods close to the end of the series, preliminary estimators have to be used, which will be revised as new observations become available, until the final or historical estimator is obtained. Since it seems reasonable that an agency producing SA data, for example, would like to provide historical series as precise as possible, we begin by considering (Sections 3 and 4) the historical estimator, obtained with the complete filter.

For a given overall ARIMA model, the different admissible decompositions can be expressed as a function of a parameter α in the unit interval. The two extreme values, $\alpha = 0$ and $\alpha = 1$, correspond to the two possible canonical decompositions, each one associated with noninvertibility of one of the components. Section 4 expresses the variance of the final estimation error as a second-order polynomial in α, where the coefficients can be determined from the overall ARIMA model. The decomposition that yields the most precise component estimators is derived and it is shown that it will always be a canonical one. Which of the two canonical decompositions it happens to be depends on the stochastic properties of the series, and a simple algorithm to determine which component should be made canonical is provided. Heuristically, the rule can be interpreted as making noninvertible the most stable of the two components (i.e., adding all noise to the most stochastic component).

In Section 5 the results are extended to preliminary estimators of the components. The estimation error is, in this case, equal to the sum of the error in the historical estimator plus the so-called revision error. Since, for an agency involved in short-term policy, minimizing the error in the measurement of the signal for the most recent period seems an important feature, special attention is paid to the error in the concurrent estimator of the components. It is seen how, for all preliminary estimators, the variance of the estimation error is a polynomial of degree 2 in α, with coefficients that can be derived from the overall ARIMA model, and that this variance is always minimized at a canonical decomposition. Which one of the two canonical decompositions it is can be determined from the following rule, which applies to historical as well as to preliminary estimation. Assume we are interested in the component for period t and that x_t is the observation corresponding to this period. Specify each component in its canonical form and consider, for each, the MMSE estimation filter. Let v_0 denote the coefficient of x_t in this filter. If the component with smallest v_0 weight is made canonical, then the estimation error variance (for both components) is minimized; i.e. all noise is then assigned to the component with the largest weight. Thus, if interest centers on having the most precise historical estimator, v_0 denotes the central weight of the complete filter. If, alternatively, the most precise concurrent estimator is sought, v_0 denotes the first weight of the one-sided filter. More generally, if interest centers on minimizing the error of the estimator of the component for time t, computed at time $(t + k)$, then v_0 is the weight of x_t in the truncated filter (i.e., the filter that extends up to x_{t+k}). As shall be seen, it will often be the case that the same canonical decomposition minimizes the variance of the different types of estimators. There are, however, cases where the solutions 'switch' and, for example, one of the canonical decompositions yields the most precise final estimator, while the other one yields the most precise concurrent estimator.

Two examples are discussed in Section 6. The first one is a simplified 'trend-plus-cycle' model of the type used by economists in business-cycle analysis, and illustrates the derivation of the estimation error variances from the parameters of the 'observed' model. The second example consists of a class of models that are often found to approximate reasonably well the stochastic properties of many series: the so-called Airline Model of Box and Jenkins (1970). This example extends the results in Hillmer (1985), and presents some stylized facts often found in actual time series. Proofs of all lemmas are given in an appendix.

1. The model

We consider the problem of decomposing an observed series x_t into two unobserved components (UC), s_t and n_t, as in

$$x_t = s_t + n_t. \tag{1.1}$$

The two components are the output of the linear stochastic processes

$$\phi_s(B)s_t = \theta_s(B)a_{st}, \tag{1.2a}$$

$$\phi_n(B)n_t = \theta_n(B)a_{nt}, \tag{1.2b}$$

where $\phi_s(B) = 1 + \phi_{s1}B + \cdots + \phi_{sp_s}B^{p_s}$ denotes a polynomial in the lag operator B having all roots on or outside the unit circle, and $\theta_s(B) = 1 + \theta_{s1}B + \cdots + \theta_{sq_s}B^{q_s}$ denotes a polynomial in B with the roots on or outside the unit circle. Replacing s by n, the polynomials $\phi_n(B)$ and $\theta_n(B)$ are defined in a similar way. The model consists of Eqs. (1.1), (1.2a) and (1.2b) and some additional assumptions.

Assumption 1. The variables a_{st} and a_{nt} are independent normally distributed white-noise innovations in the components.

Important examples of the decomposition (1.1) are the 'trend + detrended series' decomposition often used in business cycle analysis, where the trend may be a random walk and the detrended series (also called cycle) a low-order stationary process, and the 'seasonal component + SA series' decomposition, where the seasonal component is often modeled as $U(B)s_t = \theta_s(B)a_{st}$, with $U(B)$ the nonstationary 'seasonal' polynomial $U(B) = 1 + B + \cdots + B^{\tau-1}$ (τ denotes the number of observations per year), and the SA series is given by a process of the type $\nabla^d n_t = \theta_n(B)a_{nt}$, with d typically 1 or 2. Since, as the examples illustrate, each component is basically characterized by its autoregressive (AR) roots, AR roots associated with different frequencies should be allocated to different components. Thus we specify the following assumption, which also avoids redundant roots in the polynomials in Eqs. (1.2a) and (1.2b).

Assumption 2. The polynomials $\phi_s(B)$ and $\phi_n(B)$ share no root in common. The same holds true for the polynomials $\phi_s(B)$ and $\theta_s(B)$, and for the polynomials $\phi_n(B)$ and $\theta_n(B)$.

Eqs. (1.1), (1.2a) and (1.2b), and Assumptions 1 and 2 imply that the bserved series x_t follows the general ARIMA process

$$\phi(B)x_t = \theta(B)a_t. \tag{1.3}$$

The AR polynomial $\phi(B)$ is given by

$$\phi(B) = \phi_s(B)\phi_n(B), \tag{1.4}$$

and the moving average (MA) part, $\theta(B)a_t$, is determined through the identity

$$\theta(B)a_t = \phi_n(B)\theta_s(B)a_{st} + \phi_s(B)\theta_n(B)a_{nt}, \tag{1.5}$$

and the constraint that the roots of $\theta(B)$ lie on or outside the unit circle. The following assumption, however will force $\theta(B)$ to be invertible.

Assumption 3. $\theta_s(B)$ and $\theta_n(B)$ share no unit root in common.

Without loss of generality, it will be assumed that $V_a = 1$, where V_a is the variance of a_t in Eq. (1.3). (It should be kept in mind, thus, that the innovation variances V_s and V_n will be implicitly expressed as a fraction of V_a). It will prove useful to define the inverse model of Eq. (1.3), given by

$$\theta(B)z_t = \phi(B)a_t. \tag{1.6}$$

Under Assumption 3, the model (1.6) is stationary, with autocovariance generating function (ACGF) given by

$$h(B,F) = \sum_{j=0}^{\infty} h_j(B^j + F^j) = \pi(B)\pi(F), \tag{1.7}$$

where $F = B^{-1}$ denotes the forward operator and $\pi(B)$ contains the coefficients of the AR expansion of Eq. (1.3), that is

$$\pi(B) = \phi(B)/\theta(B) = \sum_{j=0}^{\infty} \pi_j B^j \quad (\pi_0 = 1). \tag{1.8}$$

Notice that the inverse process will have finite variance, given by

$$h_0 = \sum_{j=0}^{\infty} \pi_j^2. \tag{1.9}$$

2. Identification of the model

Having observations on x_t, model (1.3) can be identified from the data. For the rest of the discussion, we shall assume that the ARIMA model for x_t is known. Given this overall model, there is obviously an infinite number of ways of decomposing x_t as in Eqs. (1.1), (1.2a) and (1.2b) under Assumptions 1–3.

Let p_s, p_n, q_s and q_n denote the orders of the polynomials $\phi_s(B)$, $\phi_n(B)$, $\theta_s(B)$, and $\theta_n(B)$, respectively. If the only identification restrictions that are considered are restrictions on the orders of the polynomials of Eqs. (1.2a) and (1.2b), then a necessary and sufficient condition for identification is (see Hotta, 1983).

Assumption 4a. $p_s > q_s$ or $p_n > q_n$ (or both).

Following Tiao and Hillmer (1978), one may question whether zero-coefficient restrictions are the most adequate. It will prove useful to illustrate the point with a simple UC model similar to the ones used in business cycle analysis (see, for example, Stock and Watson, 1988). Consider an annual series, the sum of a trend component, s_t, and a detrended series, n_t, where the trend is the random-walk process

$$\nabla s_t = a_{st}, \tag{2.1a}$$

and the detrended series (or 'cycle') is the stationary ARMA(1, 1) model

$$(1 + 0.7B)n_t = (1 + 0.2B)a_{nt}. \tag{2.1b}$$

Since Eq. (2.1a) satisfies Assumption 4a, given the observed series $x_t = s_t + n_t$, the model is identified. Direct inspection of Eq. (2.1b) shows that the detrended series consists of a stationary cyclical behaviour (with a two-year period) and some random noise. The Eqs. (2.1a) and (2.1b) imply that the observed series x_t can be seen as the output of the ARMA (1, 1, 2) process:

$$(1 + 0.7B)\nabla x_t = \theta(B)a_t. \tag{2.2}$$

Setting, for our example, $V_s = 5V_n$, it is easily found that $\theta(B) = (1 + 0.364B - 0.025B^2)$. Let $g_j(\omega)$ denote the spectrum or pseudospectrum (see Harvey, 1989) of process j ($j = x, s, n$), with ω being the frequency in radians; for simplicity $g_j(\omega)$ will always be expressed in units of 2π. It is easily seen that $g_s(\omega)$ has a minimum for $\omega = \pi$ equal to $g_s(\pi) = V_s/4$. It follows that if a white-noise component u_t, with variance V_u in the interval $[0, V_s/4]$, is removed from s_t and added to n_t, the resulting components also provide an acceptable decomposition of x_t. The only difference would be that the new s_t component would be smoother, while n_t would now be noisier, as evidenced by Figs. 1 and 2. It is

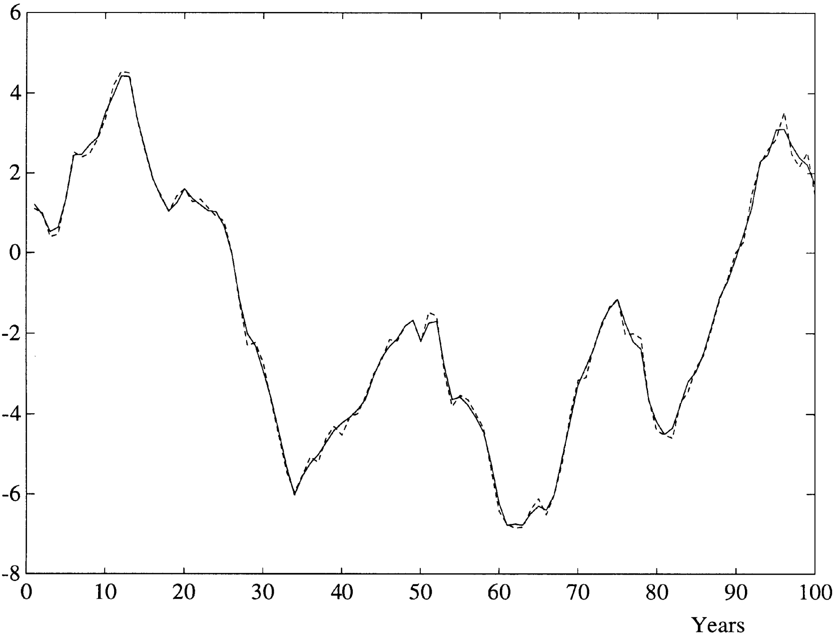

Fig. 1. Trend.

straightforward to find that the new s_t and n_t components follow processes of the type

$$\nabla s_t = (1 + \theta_s B)a_{st}, \tag{2.3a}$$

$$(1 + 0.7B)n_t = (1 + \theta_n B)a_{nt}. \tag{2.3b}$$

For a given model (2.2) for the observed series, different decompositions of the type (2.3) would provide admissible decompositions that would differ in the way the noise contained in the series is allocated to the two components. Consider an analyst interested in whatever is in the series that cannot be attributed to the trend, including the noise. He would choose the decomposition that leaves all noise in the detrended series, for which V_u is equal to its maximum value $V_s/4$. (Identification by using the 'minimum extraction' principle was first proposed by Box et al. (1978) and Pierce (1978)). Since the requirement that it should not be possible to decompose s_t into a smoother component plus white noise implies that $g_s(\pi) = 0$, and since the time domain equivalent of this spectral zero is the presence of the factor $(1 + B)$ in the MA part of the component model, s_t will

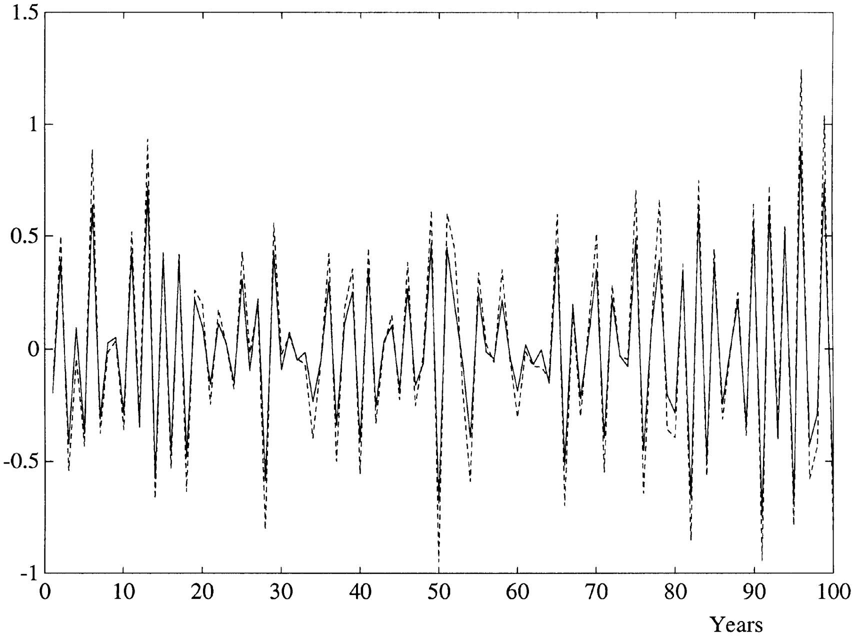

Fig. 2. Cycle.

follow the noninvertible model

$$\nabla s_t = (1 + B)a_{st},$$

and the model for n_t will be as in Eq. (2.3b). Alternatively, a similar type of reasoning may lead to the transfer of noise from n_t to s_t. Assume a time series observed with a twice-a-year frequency. Then model (2.3b) could be seen as a seasonal component and one may wish to remove a semester effect as smooth as possible. The noise would be added to the SA series and the chosen decomposition would consist of a noninvertible seasonal component n_t, with $g_n(0) = 0$, and hence

$$(1 + 0.7B)n_t = (1 - B)a_{nt};$$

the model for s_t would be as in Eq. (2.3a). Therefore, the minimum extraction requirement yields two canonical solutions, both of which could be justified; each one is characterized by noninvertibility of one of the two components. Fig. 3 displays the spectra of the observed series and its first decomposition (2.1). Fig. 4 shows the spectra of the two associated canonical decompositions.

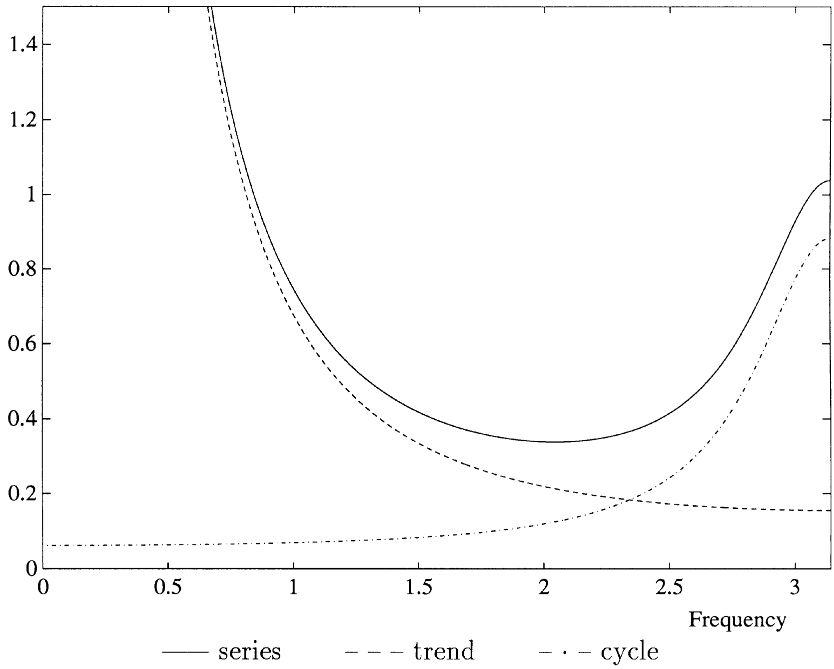

Fig. 3. Spectral decomposition.

In the general case of Eqs. (1.1), (1.2a) and (1.2b), assume an identified model with (without loss of generality) $p_s > q_s$. Let s_t be invertible; since then $g_s(\omega) > 0$ for all ω, a white-noise component, with variance in the interval $[0, \min g_s(\omega)]$, can be removed from s_t and assigned to n_t. Given that adding noise to an ARMA $(p, p-k)$ model, for $k > 0$, yields an ARMA (p, p) model, we replace Assumption 4a by the more general one.

Assumption 4. $p_s \geqslant q_s$ or $p_n \geqslant q_n$ (or both).

Given Eq. (1.3), the ARIMA model for the observed variable, the class of admissible decompositions is given by the pair of components s_t and n_t satisfying Eqs. (1.1), (1.2a), (1.2b), (1.4) and (1.5) and Assumptions 1–4. We require, of course, nonnegative spectra $g_s(\omega)$ and $g_n(\omega)$. Identification of a unique model can then be reached with the following assumption:

Assumption 5. For $\omega \in [0, \pi]$, either $\min g_s(\omega) = 0$ or $\min g_n(\omega) = 0$ (or both).

Identification is, in this case, obtained by forcing a component to be noninvertible. This noninvertible component will be denoted a 'canonical' component,

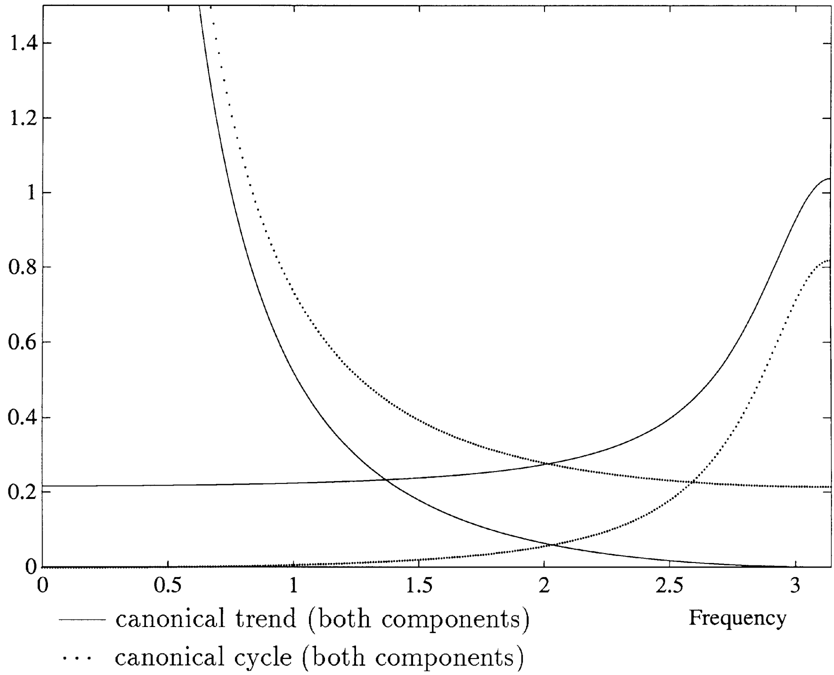

Fig. 4. Canonical decompositions.

and the associated decomposition a canonical decomposition. Obviously, in the two-component case there will be two canonical decompositions. One of them puts all additive white noise in the component n_t, the other one, in the component s_t. Any admissible decomposition can be seen as something in between, whereby some noise is allocated to n_t and some to s_t.

As shown in Hillmer and Tiao (1982), canonical components display some important features. In particular, any other admissible component is equal to the canonical one plus added noise, and hence the canonical requirement makes the component as smooth as possible. On the negative side, Maravall (1986) shows how canonical components can produce large revisions in the preliminary estimators of the component. Besides, the existence of two canonical solutions reflects two possible choices: it seems reasonable, for example, that, in order to avoid noise-induced overreaction, the monetary authority be interested in a noise-free SA series. On the other hand, it sounds also reasonable that an analyst wishes to leave as much variation as possible in the SA series, in which case the seasonal component would be noise-free. Therefore, both canonical solutions could, in principle, be rationalized.

Our characterization of a canonical decomposition is somewhat different from the standard one used in the AMB approach. In the latter, three components (trend, seasonal, and irregular) are considered. Since the irregular component picks up highly transitory variation, all white noise is assigned to it and the decomposition is then unique, with the trend and seasonal components both noninvertible. But moving to our two-component decomposition, two options are open: the irregular can be assigned to the trend or to the seasonal and, on a priori grounds, both options can be justified.

3. MMSE estimator and its error

The properties of the component estimator will depend on the admissible decomposition selected. In order to explore this dependence, we consider first the case of a complete realization of the process, i.e., the case of a series x_t with t going from $-\infty$ to ∞. It is well known (see, for example, Maravall, 1995) that in terms of the model parameters, the minimum MSE estimator of s_t is given by the expression

$$\hat{s}_t = v(B,F)x_t = V_s \frac{\theta_s(B)\theta_s(F)\phi_n(B)\phi_n(F)}{\theta(B)\theta(F)} x_t, \tag{3.1}$$

which, under appropriate assumptions concerning the starting conditions (Bell, 1984) extends to nonstationary series. In the stationary case, $v(B,F)$ is the Wiener–Kolmogorov filter and in general, we shall refer to Eq. (3.1) as the WK filter. Direct inspection shows that the WK filter is the ACGF of the model

$$\theta(B)z_t = \theta_s(B)\phi_n(B)b_t, \tag{3.2}$$

with b_t white noise with variance V_s. Since $\theta(B)$ is invertible, the model is stationary and its ACGF will converge. Unless the model for the series is a pure AR, the filter (3.1) will extend from $-\infty$ to ∞. Its convergence however guarantees that, in practice, it could be approximated by a finite filter and, in most applications, the estimator for the central periods of the series can be safely seen as generated by the WK filter (3.1). This estimator obtained with the 'complete' filter will be denoted 'historical' or 'final' estimator and shall be the one of interest until Section 5.

Given the overall ARIMA model, the effect of different admissible decompositions will show up in the MA part of (3.2), through the polynomial $\theta_s(B)$ and the variance V_s. To see the dependence of the estimation error on the chosen decomposition, we use the following result from Pierce (1979).

Lemma 1. Let e_t denote the estimation error $e_t = s_t - \hat{s}_t = \hat{n}_t - n_t$. Then e_t can be seen as the output of the ARMA model

$$\theta(B)e_t = \theta_s(B)\theta_n(B)d_t, \tag{3.3}$$

where d_t is white noise with variance (V_s, V_n).

Notice that Assumption 3 guarantees that e_t will always have a finite variance.

4. Historical estimation error and admissible decompositions

As mentioned in Section 2, each admissible decomposition is characterized by a particular allocation of the noise to the two components. Let s_t and n_t denote an admissible decomposition of x_t; then $g_x(\omega) = g_s(\omega) + g_n(\omega)$. Let, for $\omega \in [0, \pi]$, $V_u^s = \min g_s(\omega)$, and $V_u^n = \min g_n(\omega)$. The total amount of additive noise in x_t that can be distributed between the components is equal to $V_u = V_u^s + V_u^n$. Following an approach similar to Watson (1987), we shall express each admissible decomposition in terms of a parameter α that reflects the particular noise allocation. Denote by s_t^0 and n_t^0 the decomposition with s_t canonical and n_t with maximum noise, and let $g_s^0(\omega)$ and $g_n^0(\omega)$ be the associated spectra of the components (these functions are straightforward to obtain from the overall ARIMA model; see, for example, Burman, 1980). Since any admissible component s_t^α is equal to s_t^0 plus an amount of noise with variance in the interval $[0, V_u]$, the spectra of s_t^α and n_t^α can be expressed as

$$g_s^\alpha(\omega) = g_s^0(\omega) + \alpha V_u, \tag{4.1a}$$

$$g_n^\alpha(\omega) = g_n^0(\omega) - \alpha V_u \tag{4.1b}$$

with $\alpha \in [0, 1]$. The two canonical decompositions (one with s_t canonical, the other with canonical n_t) can be seen as the two extreme cases $\alpha = 0$ and $\alpha = 1$. The time domain equivalent of Eqs. (4.1a) and (4.1b) can be expressed as

$$\theta_s^\alpha(B)\theta_s^\alpha(F)V_s^\alpha = \theta_s^0(B)\theta_s^0(F)V_s^0 + \alpha\phi_s(B)\phi_s(F)V_u, \tag{4.2a}$$

$$\theta_n^\alpha(B)\theta_n^\alpha(F)V_n^\alpha = \theta_n^0(B)\theta_n^0(F)V_n^0 - \alpha\phi_n(B)\phi_n(F)V_u, \tag{4.2b}$$

where the superindex denotes the admissible decomposition under consideration. The two expressions in Eqs. (4.2a) and (4.2b) also hold in the nonstationary case. Our aim is to derive an expression that relates the variance of the component estimation error, $V(e_t^\alpha)$, to the parameter α. For $0 \leqslant \alpha \leqslant 1$ denote the estimators of the components by

$$\hat{s}_t^\alpha = v_s^\alpha(B,F)x_t, \quad \hat{n}_t^\alpha = v_n^\alpha(B,F)x_t, \tag{4.3}$$

where the WK filter is $(k = s, n)$ $v_k^\alpha(B,F) = \sum_{j=0}^{\infty} v_{k,j}^\alpha(B^j + F^j)$. Thus $\hat{s}_t^0$ and $\hat{n}_t^0$ correspond to the decomposition with canonical $\hat{s}_t$, and $\hat{s}_t^1$ and $\hat{n}_t^1$ to the one with canonical n_t.

Lemma 2. Let $e_t^\alpha = s_t^\alpha - \hat{s}_t^\alpha = \hat{n}_t^\alpha - n_t^\alpha$. *Then,*

$$V(e_t^\alpha) = V(e_t^0) + (1 - 2v_{s,0}^0)V_u\alpha - h_0 V_u^2 \alpha^2, \tag{4.4}$$

where e_t^0 *is the error in* $\hat{s}_t^0$, $v_{s,0}^0$ *is the central weight of* $v_s^0(B,F)$, *and* h_0 *is given by Eq.* (*1.9*).

Lemma 2 expresses the variance of the component estimation error as a second-order polynomial in α, with coefficients that can be obtained from the 'observed' ARIMA model. Considering that $V(e_t^0)$ is the variance of model (3.3) and $v_{s,0}^0$ is the variance of model (3.2), both for the case of canonical s_t, and that h_0 is the variance of the inverse model (1.6), the three coefficients of Eq. (4.4) can be easily computed as the variance of ARMA models with the AR polynomial always equal to $\theta(B)$. (Simple ways to compute the variance of an ARMA process are found in Box et al. (1978) and Wilson (1979).)

Since Eq. (4.4) implies that $V(e_t^\alpha)$ is a parabola in α with a finite maximum, within the interval $0 \leqslant \alpha \leqslant 1$ its minimum will always be at one of the two boundaries. Hence,

Corollary 1. $V(e_t^\alpha)$ *is always minimized at a canonical decomposition.*

Up to now, the two components s_t and n_t have been treated indistinguishably. Without loss of generality, we denote by s_t the component with the largest central weight in the WK filter that provides its canonical component estimator.

Standardization rule 1. $v_{s,0}^0 \geqslant v_{n,0}^1$.

Now it becomes possible to identify which of the two canonical decompositions has minimum estimation error.

Lemma 3. Under the standardization rule 1, *the historical estimator MSE is minimized for the decomposition with canonical* n_t.

Lemma 3 provides a simple procedure to determine which canonical decomposition yields minimum component historical estimation error. For each of the two components compute the central weight of the WK filter that yields the estimator of the component in its canonical form. Then, set as canonical component the one with the smallest weight (i.e., add all noise to the one with the largest weight). Notice that, from the two canonical specifications, the central weights of the WK filter can be simply computed as the variance of the ARMA model (3.2). Two remarks seem worth adding:

(a) Since $v_{k,0}$ measures the contribution of observation x_t to the component estimator, the precision of the estimator is maximized by assigning all additive noise to the component for which that contribution is largest.
(b) In the important application to seasonal adjustment, if s_t denotes the seasonal component and n_t the SA series, it is often the case that $v^0_{s,0} < v^0_{n,0}$ and hence the most precise estimator of s_t and n_t are obtained with a canonical seasonal component decomposition.

5. Preliminary estimation error, revisions, and admissible decompositions

Up to now we have considered estimation of the components for an infinite realization of the series. Since the WK filter converges in both directions, as mentioned in Section 3, it can be safely truncated and, for most series lengths, the estimator for the central periods can be seen as the one obtained with the complete filter (the historical or final estimator). While it seems reasonable that a data-producing agency, wishing to produce historical series as precise as possible, minimizes the error in the final estimator, it also seems reasonable that someone involved in short-term monitoring or policy-making would seek to minimize the error in the estimator for the most recent periods, in order to avoid error-induced actions. Given that for the most recent observation the WK filter cannot be applied, a preliminary estimator has to be used instead. We proceed to consider the error in this preliminary estimator.

We shall assume that the series is long enough for the weights of the filter to have converged in the direction of the past. In the vast majority of practical applications this is not a restrictive assumption, and it allows us to associate the finite-sample effect on the preliminary estimator with the unavailability of future observations. Let $\hat{s}_{t|t+k}$ denote the estimator of s_t when the last observation is x_{t+k}. We can write the error in the preliminary estimator, $d_{t|t+k} = s_t - \hat{s}_{t|t+k}$ as

$$d_{t|t+k} = e_t + r_{t|t+k},$$

where $e_t = s_t - \hat{s}_t$ is the error in the final estimator $\hat{s}_t$ discussed in the previous sections, and $r_{t|t+k} = \hat{s}_t - \hat{s}_{t|t+k}$ is the 'revision error' in the preliminary estimator. As seen in Pierce (1980), under assumptions 1–3, the two errors, e_t and $r_{t|t+k}$, are independent. Furthermore, in the stationary case, from Eqs. (1.3) and

(3.1), the estimator $\hat{s}_t$ can be expressed as

$$\hat{s}_t = \xi_s(B,F)a_t = \cdots + \xi_{s,-1}a_{t-1} + \xi_{s,0}a_t + \xi_{s,1}a_{t+1} + \cdots + \xi_{s,k}a_{t+k} + \xi_{s,k+1}a_{t+k+1} + \cdots = \xi_s^-(B)a_{t+k} + \xi_s^+(F)a_{t+k+1}, \quad (5.1)$$

where the weights $\xi_{s,j}$ are determined from the identity

$$\phi_s(B)\theta(F)\xi_s(B,F) = V_s\theta_s(B)\theta_s(F)\phi_n(F). \quad (5.2)$$

The preliminary estimator $\hat{s}_{t|t+k}$ can be obtained by taking conditional expectations at time $t+k$ in Eq. (5.1), yielding $\hat{s}_{t|t+k} = \xi_s^-(B)a_{t+k}$ since $E_T a_{T+j} = 0$ for $j > 0$. Therefore, the revision in the concurrent estimator can be expressed as

$$r_{t|t+k} = \xi_s^+(F)a_{t+k+1} = \sum_{j=k+1}^{\infty} \xi_{s,j}a_{t+j}. \quad (5.3)$$

Under suitable assumptions concerning the starting values, expression (5.3) remains valid for the nonstationary case. Since Eq. (5.2) implies that the filter $\xi_s^+(F)$ converges, expression (5.3), properly truncated, can be used to compute the ACGF of $r_{t|t+k}$, in particular

$$V(r_{t|t+k}) \simeq \sum_{j=t+k+1}^{M} \xi_{s,j}^2, \quad (5.4)$$

where M is the truncation point. For the admissible decomposition associated with α, let the components preliminary estimation error and revision error be denoted $d_{t|t+k}^{\alpha}$ and $r_{t|t+k}^{\alpha}$, respectively. In the previous section we looked at the dependence of e_t^{α} on α. Now we look at the dependence of $r_{t|t+k}^{\alpha}$ and of $d_{t|t+k}^{\alpha}$, on α. From Eqs. (1.8), (3.1) and (5.2) it is seen that

$$\pi(B)\xi_s^{\alpha}(B,F) = v_s^{\alpha}(B,F). \quad (5.5)$$

Equating coefficients of B^0 in Eq. (5.5), it is obtained that

$$v_{s,0}^{\alpha} = \sum_{j=0}^{\infty} \xi_{s,j}^{\alpha}\pi_j, \quad (5.6)$$

where $v_{s,0}^{\alpha}$ is the coefficient of x_t in the estimator (4.3). Denote by $v_{s,0}^{\alpha}(k)$ and by $h_0(k)$ the sum of the first $(k+1)$ terms in the r.h.s. of Eq. (5.6) and of Eq. (1.9). respectively; thus,

$$v_{s,0}^{\alpha}(k) = \xi_{s,0}^{\alpha} + \pi_1\xi_{s,1}^{\alpha} + \cdots + \pi_k\xi_{s,k}^{\alpha}, \quad (5.7)$$

$$h_0(k) = 1 + \pi_1^2 + \cdots + \pi_k^2. \quad (5.8)$$

Lemma 4. The variance of the revision error in the preliminary estimator $\hat{s}^{\alpha}_{t|t+k}$ *is given by*

$$V(r^{\alpha}_{t|t+k}) = V(r^{0}_{t|t+k}) + 2[v^{0}_{s,0} - v^{0}_{s,0}(k)]V_u\alpha + [h_0 - h_0(k)]V_u^2\alpha^2, \tag{5.9}$$

where the superscript 0 *denotes the decomposition with* s_t *canonical.*

Consequently, as a function of α, the variance of the revision in a preliminary estimator is a parabole with a finite minimum and hence, within a finite interval, the maximum will be attained at one of the boundaries. Thus:

Corollary 2. For $\alpha \in [0,1]$, $V(r^{\alpha}_{t|t+k})$ *is maximized at one of the two canonical decompositions.*

Corollary 2 generalizes the result in Maravall (1986), and shows an unpleasant feature of the canonical decompositions: they may imply relatively large revisions. However, when α_m, the value of α that minimizes Eq. (5.9), falls outside the interval [0,1], while one of the two canonical decompositions maximizes the variance of the revision error, the other canonical decomposition minimizes this variance. Our main concern, however, is not the revision error *per se*, but the total error in the preliminary estimator of the signal. The dependence of the variance of this error on the particular admissible decomposition selected is shown in the following lemma, obtained by using expressions (4.4) and (5.9) in $V(d^{\alpha}_{t|t+k}) = V(e^{\alpha}_t) + V(r^{\alpha}_{t|t+k})$.

Lemma 5. The variance of the error in the preliminary estimator $\hat{s}_{t|t+k}$ *is given by the second-degree polynomial in* α

$$V(d^{\alpha}_{t|t+k}) = V(d^{0}_{t|t+k}) + (1 - 2v^{0}_{s,0}(k))V_u\alpha - h_0(k)V_u^2\alpha^2, \tag{5.10}$$

where $d^{0}_{t|t+k}$ *is the error that corresponds to the canonical* s_t.

As a function of α, thus, the variance of the total error is a concave parabole, which implies the following result.

Corollary 3. For $\alpha \in [0,1]$, $V(d^{\alpha}_{t|t+k})$ *is minimized at a canonical decomposition.*

As a consequence, when the effects of the historical estimation error and of the revision error are aggregated, it still remains true that a canonical specification yields the most precise preliminary estimators of the components. In order to determine which one of the two canonical decomposition displays that property we modify the Standardization rule 1 as follows. For a particular admissible decomposition, express the preliminary estimator as

$$\hat{s}^{\alpha}_{t|t+k} = v^{\alpha}_{s}(B, F, k)x_t, \tag{5.11}$$

where $v^{\alpha}_{s}(B, F, k)$ denotes the asymmetric truncated filter (centered at x_t). The decomposition with s_t canonical yields $\hat{s}^{0}_{t|t+k}$ and $\hat{n}^{0}_{t|t+k}$, while that with n_t canonical yields $\hat{s}^{1}_{t|t+k}$ and $\hat{n}^{1}_{t|t+k}$. It will be convenient to consider the filter that yields the preliminary estimator of u_t, that is

$$\hat{u}_{t|t+k} = v_u(B, F, k)\, x_t. \tag{5.12}$$

The parameters $v^{0}_{s,0}(k)$ and $h_0(k)$, defined by Eqs. (5.7) and (5.8), have a simple interpretation in terms of the filters that provide the preliminary estimators of the components, as shown by the following lemma.

Lemma 6. (a) $v^{0}_{s,0}(k)$ *is the weight of* B^0 *in the filter* $v^{0}_{s}(B, F, k)$.
(b) $h_0(k)$ *is the weight of* B^0 *in the filter* $v_u(B, F, k)$.

Without loss of generality, denoted by s_t the component with the largest weight for x_t in the filter that provides the preliminary estimator of the component in its canonical form, i.e.:

Standardization rule 2. $v^{0}_{s,0}(k) \geqslant v^{1}_{n,0}(k)$.

Lemma 7. Among all admissible decompositions, the most precise preliminary estimators are obtained for the decomposition with canonical n_t.

Corollary 4. Let Eqs. (1.1), (1.2*a*) *and* (1.2*b*) *represent the admissible decompositions of a given ARIMA model under Assumptions* 1–3. *To select the decomposition with smallest MSE in a preliminary estimator of* s_t *and* n_t,

(a) *compute the weight of* x_t *in the two filters that provide the preliminary estimators of the components specified in their canonical form*;
(b) *choose the decomposition with canonical component the one with the smallest weight.*

Since when $k \to \infty$, from Eqs. (5.7) and (5.8), $v^{0}_{s,0}(k) \to v^{0}_{s,0}$ and $h_0(k) \to h_0$, expression (5.10) becomes (4.4), in agreement with the fact that, for $k \to \infty$, the preliminary estimator becomes the historical one. A particular case of considerable importance is when $k = 0$. The associated estimator, $\hat{s}_{t|t}$, is denoted the 'concurrent' estimator. Obviously, to use the most precise concurrent estimator (i.e., the estimator of the signal for the most recent period) could be a reasonable choice for an agency involved in short-term economic policy and monitoring. By setting $k=0$ in Eqs. (5.7) and (5.8), it is seen that $h_0=1$, and $v^0_{s,0}(0)=\xi^0_{s,0}$. Expression (5.10) becomes $V(d^{\alpha}_{t|t})=V(d^0_{t|t})+(1-2\xi^0_{s,0})V_u\alpha-V_u^2\alpha^2$, the standardization rule 2 becomes $\xi^0_{s,0}\geqslant\xi^1_{n,0}$, and the decomposition with most precise estimators is the one that sets n_t canonical and adds all noise to s_t.

Although historical and preliminary estimators have minimum MSE at a canonical specification, the particular canonical specification may well not be the same. Thus there may be models for which the historical seasonally adjusted series is best estimated with a canonical seasonal component, while the concurrent seasonally adjusted series is best estimated with a canonical trend. The switching of solutions will, of course, happen when the two standardization rules give different results. Focusing attention on historical and concurrent estimation, from Lemmas 2 and 5, the following corollary is obtained.

Corollary 5. Under the standardization rule 1, when $\xi^0_{s,0}<(1-V_u)/2$, the historical estimation error is minimized with a canonical n_t and the concurrent estimation error is minimized with a canonical s_t. Otherwise, n_t canonical minimizes both types of errors. (Under the standardization rule 2 (with $k=0$), replacing ξ^0_{s0} with v^0_{s0}, and V_u with h_0V_u in the inequality, the same result holds.)

The possible switching of solutions is an inconvenient feature since, in practice, it could mean that agencies producing historical series and agencies involved in short-term policy would use different seasonally adjusted series. Perhaps the most sensible procedure would be, in the case of switching solutions, to publish the most percise historical estimator, and to use the most precise concurrent esimator for internal short-term policy making.

6. Examples

6.1. Trend-plus-cycle model

We begin with the same example of Section 2. The model is that of Eq. (2.2) with $\theta(B)=(1+0.364B-0.025B^2)$, and accepts a 'trend-plus-cycle' decomposition, where the admissible decompositions are given by components of the type (2.3). The identity (1.5) becomes

$$(1+0.364B-0.025B^2)a_t=(1-0.7B)(1+\theta_sB)a_{st}+(1-B)(1+\theta_nB)a_{nt}. \tag{6.1}$$

Since Eq. (6.1) is an identity among MA(2) processes, the associated system of covariance equations consists of 3 equations while the unknowns are the

4 parameters θ_s, θ_n, V_s, and V_n, so that the model is not identified. As seen before, a way to reach identification is by adding the restriction $\theta_s = 0$, which yields the decomposition (2.1), with $V_s = 5V_n = 0.621$ (model (2.2) is standardized by setting $V_a = 1$). From this initial decomposition, it is found that for $\omega \in [0,\pi]$, $\min g_s(\omega) = g_s(\pi) = V_s/4 = 0.155$. Similarly, $\min g_n(\omega) = g_n(0) = 0.062$, and hence the amount of additive noise that can be exchanged between the components is $V_u = 0.217$, the sum of these two minima.

Starting from the decomposition (2.1), if we substract from $g_s(\omega)$ its minimum 0.155, the resulting spectra can be factorized to obtain the model for the canonical signal (for a simple algorithm to factorize a spectrum see Maravall and Mathis, 1994). This model is found to be

$$\nabla s_t^0 = (1 + B)a_{st}^0, \quad V_s^0 = 0.155. \tag{6.2a}$$

Since the noise removed from s_t is added to n_t, factorizing the spectrum $(g_n(\omega) + 0.155)$ yields the model for the component n_t^0,

$$(1 + 0.7B)n_t^0 = (1 + 0.433B)a_{nt}^0, \quad V_n^0 = 0.301. \tag{6.2b}$$

From models (2.2) and (6.2), expression (3.2), (3.3), and (1.6) indicate that the variance $V(e_t^0)$, the central weight of the WK filter for $\hat{s}_t^0$, $v_{s,0}^0$, and the coefficient h_0 of Lemma 2 are the variances of the processes

$$\theta(B)z_t = (1 + B)(1 - 0.443B)b_t, \quad V_b = V_s^0 V_n^0 = 0.047,$$

$$\theta(B)z_t = (1 + B)(1 + 0.7B)b_t, \quad V_b = V_s^0 = 0.155,$$

$$\theta(B)z_t = (1 + 0.7B)(1 - B)b_t, \quad V_b = V_a = 1,$$

respectively. This yields $V(e_t^0) = 0.101$, $v_{s,0}^0 = 0.441$, $h_0 = 1.653$, and using Eq. (4.4), for any admissible decomposition

$$V(e_t^\alpha) = 0.101 + 0.026\,\alpha - 0.078\,\alpha^2. \tag{6.3}$$

The historical estimation error variance is seen to be minimized for $\alpha = 1$, that is, for the decomposition with canonical n_t, in which case $V(e_t^1) = 0.049$. This could have been found from Lemma 3. The decomposition with canonical n_t is found by removing $\min g_n(\omega) = 0.062$ from $g_n(\omega)$, and adding it to $g_s(\omega)$ in the initial decomposition (2.1). Factorizing the resulting spectra yields the models.

$$\nabla s_t^1 = (1 - 0.084)a_{st}^1, \quad V_s^1 = 0.739,$$

$$(1 + 0.7B)n_t^1 = (1 - B)a_{nt}^1, \quad V_n^1 = 0.018.$$

Proceeding as before, $v^1_{n,0}$ is the variance of the model $\theta(B)z_t = (1-B)^2 b_t$, with $V_b = V^1_n = 0.018$, and hence equal to 0.200. Thus, since $v^0_{s,0} = 0.441 > v^1_{n,0}$, the notation for the components conforms to the standardization rule 1 and Lemma 3 can be directly applied. For this example, thus, the MSE of the historical estimators of the two components are minimized when the cycle is made canonical.

Concerning preliminary estimation, we focus on the concurrent estimator and its one-period revision. In order to obtain the error variances for any admissible decomposition, from Eq. (5.10), we need the parameters $V(d^0_{t|t+k})$, $v^0_{s,0}(k)$ and $h_0(k)$, for $k = 0, 1$. The first parameter $V(d^0_{t|t+k})$ is equal to the sum of $V(e^0_t)$, already computed, plus $V(r^0_{t|t+k})$, which can be computed through Eq. (5.4). For this we need the coefficients in $F^j (j = 0, 1, \ldots, M)$ of the filter $\xi^0_s(B, F)$, given by Eq. (5.2). For this example.

$$\xi^0_s(B,F) = V^0_s \frac{(1+B)(1+F)(1+0.7F)}{(1-B)(1+0.364F-0.025F^2)} = V^0_s \eta(B,F).$$

In order to express the filter $\eta(B,F)$ as the sum of a filter in B and a filter in F, we first write the numerator and denominator of $\eta(B,F)$ as $(1+B)(0.7+1.7B+B^2)F^2$ and $(1-B)(-0.025+0.364B+B^2)F^2$, respectively, and then obtain the partial fractions decomposition:

$$\eta(B,F) = \frac{c_0}{1-B} + \frac{c_1 + c_2 B + c_3 B^2}{-0.025 + 0.364B + B^2}. \tag{6.4}$$

The coefficients c_0, c_1, c_2 and c_3 are determined by removing denominators in Eq. (6.4), and equating coefficients of B^0, B, B^2, and B^3 in the left- and right-hand side of the resulting identity. This yields a linear system of equations with solution $c_0 = 5.078$, $c_1 = 0.827$, $c_2 = 1.378$, and $c_3 = -1$. The filter $\eta(B,F)$ can then be expressed as $\eta(B,F) = \eta^-(B) + \eta^+(F)$, where $\eta^-(B) = 5.078(1-B)^{-1}$ and $\eta^-(F) = (-1 + 1.378F + 0.827F^2)(1 + 0.364F - 0.025F^2)^{-1}$. Thus, multiplying by V^0_s, $\xi^0_{sj} = 0.788$ for $j < 0$; $\xi^0_{s0} = 0.633$, $\xi^0_{s1} = 0.270$, $\xi^0_{s1} = 0.026$, $\xi^0_{s3} = -0.003$, $\xi^0_{s4} = 0.002$, $\xi^0_{s5} = -0.001$, and $\xi^0_{sj} \simeq 0$ for $j > 5$. Expression (5.4) yields $V(r^0_{t|t}) = 0.074$, and hence $V(d^0_{t|t}) = V(e^0_t) + V(r^0_{t|t}) = 0.175$. For the one-period revision of the concurrent estimator, since $V(d^0_t) = V(d^0_{t|t+1}) + (\xi^0_{s,1})^2$, it follows that $V(d^0_{t|t+1}) = 0.103$. The coefficients $h_0(k)$ are found through Eq. (5.8), with $\pi(B) = (1 + 0.713B)\nabla/\theta(B)$. In particular $\pi_0 = 1$, $\pi_1 = -0.664$, and hence $h(0) = 1$, $h(1) = 1.441$. Finally, from Eq. (5.7), $v^0_{s,0}(0) = 0.633$, and $v^0_{s,0}(1) = 0.453$. Plugging these values in Eq. (5.10), it is obtained that

$$V(d^\alpha_{t|t}) = 0.175 - 0.057\alpha - 0.047\alpha^2, \tag{6.5}$$

$$V(d^\alpha_{t|t+1}) = 0.103 + 0.020\alpha - 0.068\alpha^2. \tag{6.6}$$

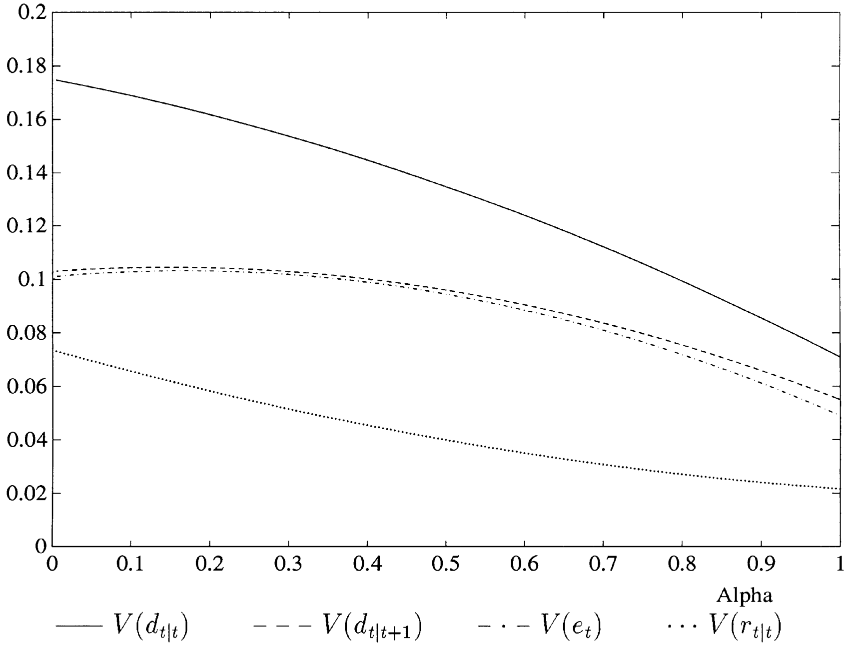

Fig. 5. Estimation error variance for different admissible decomposition (first example).

Expression (5.9) provides also the variance of the revision error; for the concurrent estimator

$$V(r^{\alpha}_{t|t}) = 0.074 - 0.083\alpha + 0.031\alpha^2, \tag{6.7}$$

The four variances (6.3), (6.5), (6.6), and (6.7) are represented in Fig. 5. For this example, consideration of different estimators does not produce any switching of solutions, and the specification with canonical n_t ($\alpha = 1$) always minimizes the estimation error variance. (It is straightforward to find that $\xi^1_{n0} = 0.150 < \xi^0_{s,0} = 0.633$, and hence the standardization rule 2 is also satisfied). The variances of the concurrent, one-period revision, and final estimation errors are given in Table 1. The use of a canonical n_t component instead of a canonical s_t cuts in less than half the variance of the error.

6.2. The 'airline model'

We consider a class of models appropriate for monthly or quarterly series with trend and seasonality. The model is given by

$$\nabla\nabla_\tau x_t = (1 + \theta_1 B)(1 + \theta_\tau B^\tau)a_t, \tag{6.8}$$

Table 1
Estimation error variance: Trend plus cycle example

	Concurrent estimator	One-period revision	Final estimator
Canonical seasonal component ($\alpha = 0$)	0.175	0.103	0.101
Canonical seasonally adjusted series ($\alpha = 1$)	0.070	0.055	0.049

Table 2
Variance of error in final estimator: Airline model. s_t: seasonal component; n_t: SA series

θ_1	Model Spec.	$\theta_{12} = 0$	$\theta_{12} = -0.25$	$\theta_{12} = -0.5$	$\theta_{12} = -0.75$
0.75	Canonical s_t	0.410	0.504	0.436	0.259
	Canonical n_t	0.407	0.504	0.439	0.267
0.50	Canonical s_t	0.308	0.377	0.327	0.195
	Canonical n_t	0.300	0.376	0.337	0.220
0.25	Canonical s_t	0.226	0.274	0.239	0.144
	Canonical n_t	0.210	0.271	0.255	0.190
0	Canonical s_t	0.164	0.197	0.173	0.106
	Canonical n_t	0.138	0.186	0.191	0.168
− 0.25	Canonical s_t	0.121	0.143	0.129	0.081
	Canonical n_t	0.082	0.119	0.139	0.146
− 0.50	Canonical s_t	0.096	0.113	0.106	0.070
	Canonical n_t	0.042	0.070	0.095	0.118
− 0.75	Canonical s_t	0.077	0.118	0.116	0.076
	Canonical n_t	0.019	0.036	0.054	0.074

where τ is the number of observations per year and, as before, $V_a = 1$. Following the work of Box and Jenkins (1970), model (6.8) is often referred to as the 'airline model'. On the one hand, it is often encountered in practice; it also provides a convenient reference example, since the parameters θ_1 and θ_τ are directly related to the stability of the trend and of the seasonal component. In particular, a value of the parameter θ_1 (or θ_τ) close to -1 indicates the presence of a stable trend (or seasonal) component. For $-1 < \theta_1 < 1$ and $-1 < \theta_\tau < \theta^*$, where θ^* is a small positive value (see Figs. 7 and 8), the model accepts a decomposition as in Eqs. (1.1), (1.2a) and (1.2b) with assumptions 1–4. If s_t denotes the seasonal component and n_t the SA series, then the components follow models of the type $U(B)s_t^\alpha = \theta_s^\alpha(B)a_{st}^\alpha$, and $\nabla^2 n_t^\alpha = \theta_n^\alpha(B)a_{nt}^\alpha$, where $\theta_s^\alpha(B)$ and $\theta_n^\alpha(B)$ are, in general, polynomials in B of order $\tau - 1$ and 2, respectively.

We know that the estimators with minimum MSE are obtained at a canonical specification. Tables 2 and 3 present the final and concurrent estimation error

Table 3
Variance of error in concurrent estimator: Airline model. s_t: seasonal component; n_t: SA series

θ_1	Model Spec.	$\theta_{12}=0$	$\theta_{12}=-0.25$	$\theta_{12}=-0.5$	$\theta_{12}=-0.75$
0.75	Canonical s_t	1.257	1.151	0.905	0.521
	Canonical n_t	1.261	1.557	0.913	0.532
0.50	Canonical s_t	0.956	0.873	0.685	0.393
	Canonical n_t	0.964	0.888	0.710	0.433
0.25	Canonical s_t	0.699	0.641	0.505	0.292
	Canonical n_t	0.710	0.665	0.551	0.369
0	Canonical s_t	0.491	0.458	0.367	0.215
	Canonical n_t	0.498	0.483	0.426	0.327
− 0.25	Canonical s_t	0.333	0.323	0.269	0.164
	Canonical n_t	0.326	0.336	0.324	0.292
− 0.50	Canonical s_t	0.228	0.239	0.214	0.139
	Canonical n_t	0.193	0.217	0.234	0.244
− 0.75	Canonical s_t	0.149	0.205	0.207	0.143
	Canonical n_t	0.097	0.120	0.141	0.161

variance associated with the two canonical decompositions, for $\tau = 12$, and for different values of θ_1 and θ_{12} within the admissible region. For both types of errors, the variance is large for models whose spectra are dominated by a very stochastic trend (values of θ_1 close to 1). On the other hand, the estimation error variance is small when the model contains relatively stable components. It is further seen that the variance of the final estimation error accounts for (roughly) between 1/3 and 1/2 of the variance of the concurrent estimation error; the revision error is, thus, typically larger than the final estimation error.

Using, as an example, $\theta_1 = -0.34$ and $\theta_{12} = -0.42$, the coefficients of expressions (4.4), (5.9), and (5.10) can be derived from the overall ARIMA model in a manner similar to that illustrated in the previous example. The variances of the errors are found equal to

$$V(r^{\alpha}_{t|t}) = 0.138 - 0.018\alpha + 0.057\alpha^2,$$

$$V(d^{\alpha}_{t|t}) = 0.263 + 0.081\alpha - 0.051\alpha^2,$$

$$V(d^{\alpha}_{t|t-12}) = 0.153 + 0.065\alpha - 0.094\alpha^2,$$

$$V(e^{\alpha}_{t}) = 0.125 + 0.099\alpha - 0.108\alpha^2,$$

and they are represented in Fig. 6. This example illustrates a case of 'switching solutions': while the final estimation error is minimized with the decomposition with canonical SA series, the concurrent estimation error is minimized with the

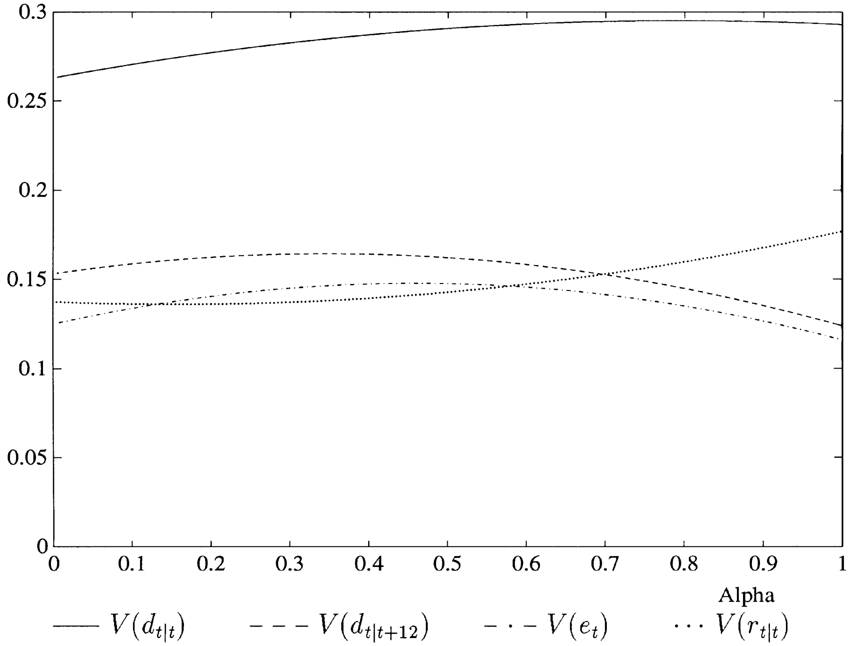

Fig. 6. Estimation error variance for different admissible decomposition (second example).

Table 4
Estimation error variance: Airline model example

	Concurrent estimator	12-period revision	Final estimator
Canonical seasonal component ($\alpha = 0$)	0.263	0.153	0.125
Canonical seasonally adjusted series ($\alpha = 1$)	0.293	0.124	0.116

decomposition with a canonical seasonal component. Still, as seen in Table 4, the difference between the errors associated with the two canonical decompositions is relatively small.

For the monthly and quarterly Airline Model, Figs. 7 and 8 display the lines that separate the regions of the admissible parameter space where a canonical seasonal minimizes the final and concurrent estimation error, from that where the minimum is achieved with a canonical SA series. The area between the continuous and dotted lines represents the region of switching solutions. Be that

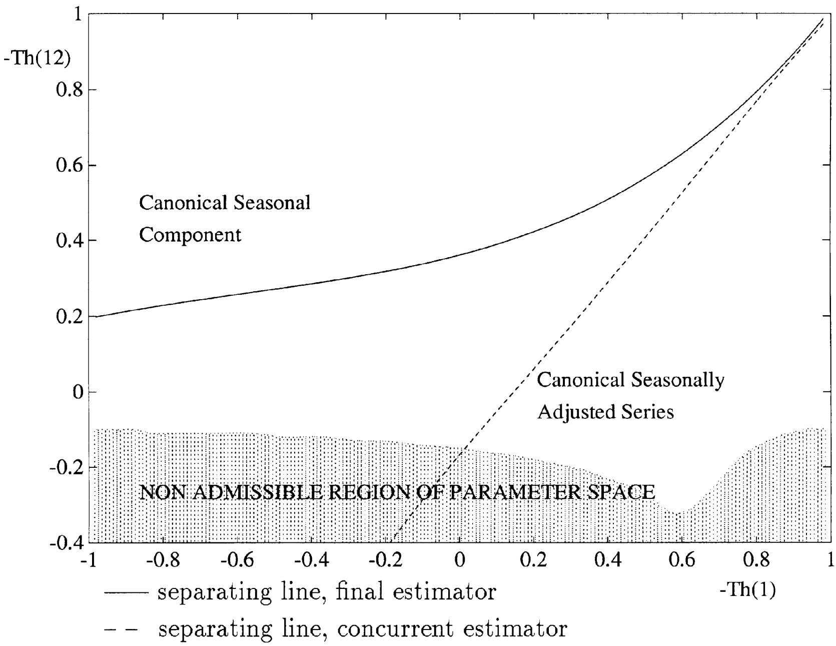

Fig. 7. Decomposition with minimum MSE: Airline monthly model.

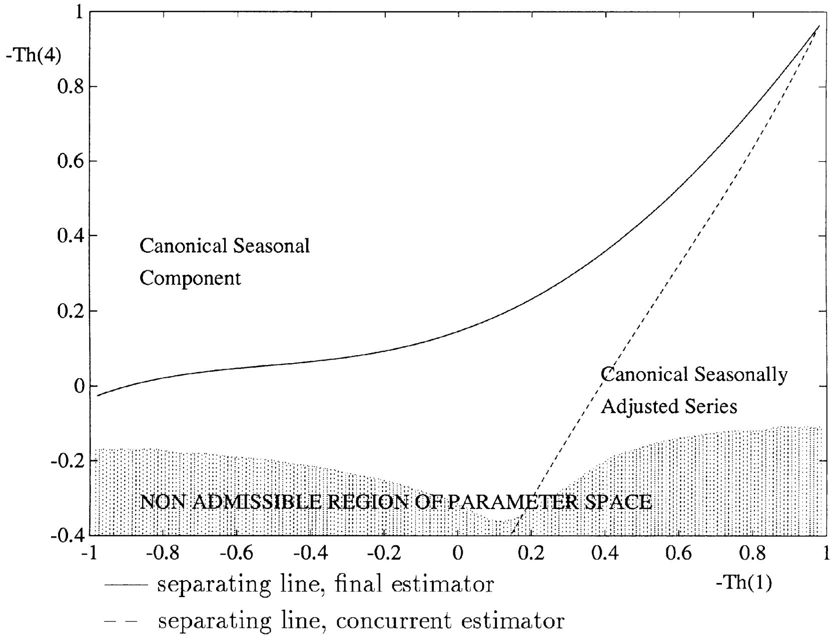

Fig. 8. Decomposition with minimum MSE: Airline quarterly model.

as it may, what both figures show is that stable trends imply the use of a canonical trend, while stable seasonals imply the use of a canonical seasonal component.

Acknowledgements

Thanks are due to two referees for their helpful comments, which certainly improved our understanding of the problem.

Appendix A.

Proof of Lemma 2. From Lemma 1,

$$\mathrm{ACGF}(e_t^\alpha) = V_n^\alpha V_s^\alpha \frac{\theta_s^\alpha(B)\theta_s^\alpha(F)\theta_n^\alpha(B)\theta_n^\alpha(F)}{\theta(B)\theta(F)}. \tag{A.1}$$

Considering Eqs. (4.2a) and (4.2b), the numerator can be rewritten as

$$\theta_s^0(B)\theta_s^0(F)\theta_n^0(B)\theta_n^0(F)V_s^0V_n^0 + \alpha V_u[-\theta_s^0(B)\theta_s^0(F)\phi_n(B)\phi_n(F)V_s^0 + \theta_n^0(B)\theta_n^0(F)\phi_s(B)\phi_s(F)V_n^0] - \alpha^2 V_u^2\phi(B)\phi(F),$$

where use has been made of (1.4). Therefore, Eq. (A.1) becomes

$$\mathrm{ACGF}(e_t^\alpha) = \mathrm{ACGF}(e_t^0) + \alpha V_u[-v_s^0(B,F) + v_n^0(B,F)] + \alpha^2 V_u^2\pi(B)\pi(F). \tag{A.2}$$

Since $v_n^0(B,F) - v_s^0(B,F) = 1 - 2v_s^0(B,F)$ and considering Eq. (1.7), equating the coefficient of B^0 in both sides of Eq. (A.2) yields expression (4.4). □

Proof of Lemma 3. The series x_t can be decomposed as in

$$x_t = s_t^0 + n_t^1 + u_t, \tag{A.3}$$

where s_t^0 and n_t^1 are the two canonical components, and u_t is white noise with variance V_u. The WK filter for u_t is given by

$$v_u(B,F) = V_u \frac{\phi(B)\phi(F)}{\theta(B)\theta(F)} = V_u\pi(B)\pi(F), \tag{A.4}$$

equal thus to the ACGF of the inverse model (1.7), scaled by V_u. It follows that $V_u h_0$ is the central coefficient of the WK filter for u_t. Therefore,

$$v^1_{s,0} = v^0_{s,0} + V_u h_0 \tag{A.5}$$

is the central weight of the WK filter associated with the decomposition that assigns all white noise to s_t (i.e., with the decomposition with n_t canonical). Under the standardization rule 1, $v^0_{s,0} \geqslant v^1_{n,0}$ or, adding $v^0_{s,0} + V_u h_0$ to both sides of the inequality, $2v^0_{s,0} + V_u h_0 \geqslant v^1_{n,0} + v^1_{s,0} = 1$, where use has been made of Eq. (A.5) and of $x_t = \hat{n}^1_t + \hat{s}^1_t$. Therefore,

$$1 - 2v^0_{s,0} - V_u h_0 \leqslant 0. \tag{A.6}$$

Given that $V(e^\alpha_t)$, as a function of α, is a concave parabola, its minimum will happen at one of the boundaries ($\alpha = 0$ or $\alpha = 1$). From Eq. (4.4), $V(e^1_t) - V(e^0_t) = V_u[1 - 2v^0_{s,0} - V_u h_0]$, so that, considering Eq. (A.6), $V(e^1_t) \leqslant V(e^0_t)$. □

Proof of Lemma 4. Expression (5.5) can be rewritten $\phi(B)\xi^\alpha_s(B,F) = v^\alpha_s(B,F)\theta(B) = [v^0_s(B,F) + \alpha V_u \frac{\phi(B)\phi(F)}{\theta(B)\theta(F)}]\theta(B) = \phi(B)\xi^0_s(B,F) + \alpha V_u \phi(B)\pi(F)$, or simplifying, $\xi^\alpha_s(B,F) = \xi^0_s(B,F) + \alpha V_u \pi(F)$. From Eq. (5.3), $r^\alpha_{t|t+k} = \sum_{j=k+1}^{\infty} [\xi^0_{s,j} + \alpha V_u \pi_j] a_{t+j}$, and hence $V(r^\alpha_{t|t+k}) = V(r^0_{t|t+k}) + 2\alpha V_u (\sum_{j=k+1}^{\infty} \xi^0_{sj}\pi_j) + \alpha^2 V_u^2 \sum_{j=k+1}^{\infty} \pi_j^2$.
Considering Eqs. (1.9), (5.6), (5.7) and (5.8), expression (5.9) is obtained. □

Proof of Lemma 6. From Eqs. (1.3), (5.11) and (1.8),

$$\pi(B)\hat{s}_{t|t+k} = v_s(B,F,k)a_t, \tag{A.7}$$

where the superscript α has been deleted for notational simplicity. Taking conditional expectations at time $t + k$, expression (5.1) yields

$$\hat{s}_{t|t+k} = \xi_s(B,F,k)a_t, \tag{A.8}$$

where $\xi_s(B,F,k)$ is the filter $\xi_s(B,F)$ truncated at F^k, and use has been made of the property $E_{t+k}a_T = 0$ when $T > t + k$. Comparing Eqs. (A.7) and (A.8), it is seen that $v_s(B,F,k) = \xi_s(B,F,k)\pi(B)$. Equating the coefficients of B^0 at both sides of this identity, if c_0 denotes that of the l.h.s.,

$$c_0 = \sum_{i=0}^{k} \xi_{s,i}\pi_i. \tag{A.9}$$

For the canonical specification of s_t, Eqs. (5.7) and (A.9) imply $c_0 = v^0_{s,0}(k)$ which proves part (a). Part (b) is proved in an identical manner, by noticing that

Eqs. (5.12) and (A.4) imply $\hat{u}_{t|t+k} = a_t + \pi_1 a_{t+1} + \cdots + \pi_k a_{t+k} = \pi(F,k)a_t$, and hence $\xi_u(B,F,k) = \pi(F,k)$. □

Proof of Lemma 7. Starting with the decomposition (A.3), the lemma is proved in the same way as Lemma 3, replacing the v- by the $v(k)$-coefficients, and h_0 by $h_0(k)$. □

References

Bell, W.R., 1984. Signal extraction for nonstationary time series. Annals of Statistics 12, 646–664.

Bell, W.R., Hillmer, S.C., 1984. Issues involved with the seasonal adjustment of economic time series. Journal of Business and Economic Statistics 2, 291–320.

Box, G.E.P., Hillmer, S.C., Tiao, G.C., 1978. Analysis and modeling of seasonal time series. In: Zellner, A. (Ed.), Seasonal Analysis of Economic Time Series. U.S. Dept. of Commerce-Bureau of the Census, Washington, DC, pp. 309–334.

Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.

Burman, J.P., 1980. Seasonal adjustment by signal extraction. Journal of the Royal Statistical Society A 143, 321–337.

Burridge, P., Wallis, K.F., 1985. Calculating the variance of seasonally adjusted series. Journal of the American Statistical Association 80, 541–552.

Engle, R.F., 1978. Estimating Structural Models of Seasonality. In: Zellner, A. (Ed.), Seasonal Analysis of Economic Time Series. U.S. Dept. of Commerce-Bureau of the Census, Washington, DC, pp. 281–297.

Findley, D.F., 1985. An Analysis and Generalization of M. Watson's Minimax Procedure for Determining the Component Estimates of a Seasonal Time Series. Statistical Research Division, U.S. Bureau of the Census, Washington, DC.

Harvey, A.C., 1989. Forcasting, Structural Times Series Models and the Kalman Filter. Cambridge University Press, Cambridge.

Harvey, A.C., Todd, P.H.J., 1983. Forecasting economic time series with structural and Box–Jenkins models: a case study. Journal of Business and Economic Statistics 1, 299–306.

Hillmer, S.C., 1985. Measures of variability for model-based seasonal adjustment procedures. Journal of Business and Economic Statistics 3, 60–68.

Hillmer, S.C., Tiao, G.C., 1982. An ARIMA-model based approach to seasonal adjustment. Journal of the American Statistical Association 77, 63–70.

Hotta, L.K., 1983. Identification of unobserved components models. Journal of Time Series Analysis 10, 259–270.

Maravall, A., 1986. Revisions in ARIMA signal extraction. Journal of the American Statistical Association 81, 736–740.

Maravall, A., 1995. Unobserved components in economic time series. In: Pesaran, H., Wickens, M. (Eds.), The Handbook of Applied Econometrics, vol. 1. Basil Blackwell, Oxford.

Maravall, A., Mathis, A., 1994. Encompassing univariate models in multivariate time series: a case study. Journal of Econometrics 61, 197–233.

Pierce, D.A., 1978. Seasonal adjustment when both deterministic and stochastic seasonality are present. In: Zellner, A. (Ed.), Seasonal Analysis of Economic Time Series. U.S. Dept. of Commerce-Bureau of the Census, Washington, DC, pp. 242–269.

Pierce, D.A., 1979. Signal extraction error in nonstationary time series. Annals of Statistics 7, 1303–1320.

Pierce, D.A., 1980. Data revisions in moving average seasonal adjustment procedures. Journal of Econometrics 14, 95–114.
Stock, J.H., Watson, M.W., 1988. Variable trends in economic time series. Journal of Economic Perspectives 2, 147–174.
Tiao, G.C., Hillmer, S.C., 1978. Some consideration of decomposition of a time series. Biometrika 65, 497–502.
Watson, M.W., 1987. Uncertainty in model-based seasonal adjustment procedures and construction of minimax filters. Journal of the American Statistical Association 82, 395–408.
Wilson, G.T., 1979. Some efficient computational procedures for high order ARMA models. Journal of Statistical Computation and Simulation 8, 301–309.