*Journal of Forecasting*
*J. Forecast.* **30**, 597–621 (2011)
Published online 28 April 2010 in Wiley Online Library
(wileyonlinelibrary.com) **DOI**: 10.1002/for.1175

# A Nonparametric Method for Asymmetrically Extending Signal Extraction Filters†

TUCKER McELROY*
*Statistical Research Division, US Census Bureau, Washington, DC, USA*

ABSTRACT
Two important problems in the X-11 seasonal adjustment methodology are the construction of standard errors and the handling of the boundaries. We adapt the 'implied model approach' of Kaiser and Maravall to achieve both objectives in a nonparametric fashion. The frequency response function of an X-11 linear filter is used, together with the periodogram of the differenced data, to define spectral density estimates for signal and noise. These spectra are then used to define a matrix smoother, which in turn generates an estimate of the signal that is linear in the data. Estimates of the signal are provided at all time points in the sample, and the associated time-varying signal extraction mean squared errors are a by-product of the matrix smoother theory. After explaining our method, it is applied to popular nonparametric filters such as the Hodrick–Prescott (HP), the Henderson trend, and ideal low-pass and band-pass filters, as well as X-11 seasonal adjustment, trend, and irregular filters. Finally, we illustrate the method on several time series and provide comparisons with X-12-ARIMA seasonal adjustments. 

KEY WORDS ARIMA model; nonstationary time series; seasonal adjustment; X-11

## INTRODUCTION

A long-standing problem in the seasonal adjustment community is the determination of signal extraction error estimates for the X-11 filters for seasonal, nonseasonal, trend, and irregular; see President's Committee to Appraise Employment and Unemployment Statistics (1962) and the discussion in Bell and Kramer (1996). This uncertainty is sometimes assessed through revision error (Pierce, 1980), but more properly signal extraction mean squared error (MSE) is the correct quantity to measure (Bell and Hillmer, 1984). Secondly, there is the so-called boundary problem—the question of how to asymmetrically extend the X-11 filters to the boundaries of the sample. For model-based approaches

* Correspondence to: Tucker McElroy, Statistical Research Division, US Census Bureau, 4700 Silver Hill Road, Washington, DC 20233-9100, USA. E-mail: Tucker.s.mcelry@census.gov
† *Disclaimer*. This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the author and not necessarily those of the US Census Bureau.

(MBA) to seasonal adjustment, both the signal extraction MSEs and asymmetric filter extensions are automatic byproducts of the filter calculations (Bell, 1984). This paper sets forth a non-MBA method for addressing these two problems; our methodology uses the 'recast' concept of Kaiser and Maravall (2005)—which is based on the 'implied models' concept of Bell and Hillmer (1984)—adapted to a non-MBA context.

Several approaches to determining X-11 MSEs have been proposed, including Wolter and Monsour (1981), Pfeffermann (1994), and Bell and Kramer (1996). These methods explicitly recognize the contribution of sampling error. There are also a number of papers on matching X-11 to MBA filters—Cleveland and Tiao (1976), Burridge and Wallis (1985), and Depoutot and Planas (1998) attempt to find a closest MBA filter by matching filter weights. As pointed out in Bell and Hillmer (1984), filters do not fully determine a model, and the full model is still needed to calculate MSEs. Bell (2005) suggests that one could compute the X-11 MSEs given a model and component decomposition; Chu *et al.* (2007) carry out this idea by finding the Box–Jenkins Airline model with seasonal adjustment filter that is closest to that of the X-11 filter, where the metric used considers the frequency response functions of the filters. The boundary problem can he solved via forecast and backcast extension of the data; this is the method of X-11-ARIMA (Dagum, 1980) using a fitted ARIMA model to do the forecasts. Before the idea of forecast extension, asymmetric X-11 filters were generated in an ad hoc fashion. The forecast extension approach is more defensible, but typically the forecasts are generated from a fitted model.

The above methods are model-based. In contrast, our approach here relies on a nonparametric spectrum estimate and certain properties of the X-11 filters. This spectrum estimate, together with the X-11 frequency response function, will define the target signal spectrum and a corresponding matrix smoother (by 'matrix smoother' we mean a suite of time-varying signal extraction filters). The matrix smoother consists of all the various asymmetric filters, so that the boundary problem is addressed. From the theory of matrix smoothers discussed below, we can also obtain X-11 MSEs at each time point in the sample. Thus the method addresses both of the problems outlined above without using model-based forecast extension, and thus avoids implicit problems in forecasting due to misspecified models.

Our method follows the basic strategy outlined in Kaiser and Maravall (2005). The next section discusses this theory, with details on how a matrix smoother and error covariance matrix can be generated from the spectra for signal and noise. In the third section we discuss several illustrations of our methodology on popular nonparametric filters such as the Hodrick–Prescott (HP), the Henderson trend, and the ideal low-pass. A technical difficulty stems from the fact that X-11 filters do not have frequency response function bounded between 0 and 1, which is a requirement of the method; in the fourth section we discuss an approximation to the X-11 filter that resolves these difficulties. The recast X-11 matrix smoother is then constructed and illustrated on four seasonal time series.

## RECASTING

The original motivation of Kaiser and Maravall (2005) was to provide a model-based interpretation to the use of certain *ad hoc* filters (in particular, the HP) that were applied to seasonally adjusted data. By combining the given symmetric filter and the ARIMA model for the observed time series, one could obtain models for the various components such that their estimation via MBA would exactly coincide with sequential application of the original filters. As a by-product, one could obtain extensions of the *ad hoc* filters to the sample boundary, as well as signal extraction MSEs. In this



**DOI**: 10.1002/for

paper we replace the MBA of Kaiser and Maravall (2005) with a nonparametric approach; this is described below. Essentially, we replace the primitive of a known ARIMA model for the data with the knowledge of a spectral density estimate of the series. In cases where there is uncertainty regarding the correct model specification, our nonparametric approach is especially appealing.

Consider a nonstationary time series $Y_t$ that can be written as the sum of two possibly nonstationary components $S_t$ and $N_t$, the signal and the noise:

$$Y_t = S_t + N_t \tag{1}$$

Following Bell (1984), we let $Y_t$ be an integrated process such that $W_t = \delta(B)Y_t$ is weakly stationary. Here $B$ is the backshift operator and $\delta(z)$ is a polynomial with all roots located on the unit circle of the complex plane (also, $\delta(0) = 1$ by convention), and $F = B^{-1}$ is the forward shift operator. This $\delta(z)$ is referred to as the differencing operator of the series, and we assume it can be factored into relatively prime polynomials $\delta^S(z)$ and $\delta^N(z)$ (i.e., polynomials with no common zeroes), such that the series $U_t = \delta^S(B)S_t$ and $V_t = \delta^N(B)N_t$ are mean zero weakly stationary time series, which are uncorrelated with one another. Note that $\delta^S = 1$ and/or $\delta^N = 1$ are included as special cases. (In these cases either the signal or the noise or both are stationary.) We let $d$ be the order of $\delta$, and $d_S$ and $d_N$ are the orders of $\delta^S$ and $\delta^N$; since the latter operators are relatively prime, $\delta = \delta^S \cdot \delta^N$ and $d = d_S + d_N$.

We have the following relationship between a spectral density $f$ and its associated autocovariance function $\gamma$: $\gamma(h) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\lambda)e^{i\lambda h}\mathrm{d}\lambda$. The autocovariances can be arranged into a symmetric Toeplitz matrix $\Sigma$ via $\Sigma_{jk} = \gamma(j-k)$. If the Fourier transform of the autocovariance sequence is $f$, then we refer to the Toeplitz matrix by $\Sigma(f)$. Also, if $X$ is a random vector we will sometimes write $\Sigma_X$ to denote its covariance matrix (when the components of $X$ represent observations on a sample drawn from a stationary process with spectrum $f$, then $\Sigma_X = \Sigma(f)$).

Let $\Psi(B)$ denote our generic filter, which we suppose to be symmetric such that the frequency response function $\Psi(e^{-i\lambda})$ is real-valued, where $\lambda \in [-\pi, \pi]$. Now for Wiener–Kolmogorov (WK) filters (see Bell, 1984) we have

$$0 \le \Psi\left(e^{-i\lambda}\right) \le 1 \tag{2}$$

although this condition does not need to hold for a generic filter. In fact, this condition actually precludes most of the X-11 filters, but there are many other examples of filters that satisfy this condition (e.g. HP, Henderson, ideal low-pass and band-pass). Later we will consider modifications of the X-11 filters that satisfy (2), which is essential to our method. Now following the approach of Kaiser and Maravall (2005), the relation of signal pseudo-spectrum $f_S$ to data pseudo-spectrum $f_Y$ should be

$$f_S(\lambda) = \Psi\left(e^{-i\lambda}\right) f_Y(\lambda) \tag{3}$$

which is exactly true if $\Psi(B)$ is a WK filter. Note that (2) is necessary for (3) to be mathematically meaningful. Although in reality $\Psi(B)$ may not be a WK filter, we nevertheless assume (3) to be true so as to define $f_S$ in terms of $f_Y$ and $\Psi$. Next, it follows that

$$f_U(\lambda) = \left|\delta^S\left(e^{-i\lambda}\right)\right|^2 f_S(\lambda) = \frac{\Psi\left(e^{-i\lambda}\right)}{\left|\delta^N\left(e^{-i\lambda}\right)\right|^2} f_W(\lambda)$$



**DOI**: 10.1002/for

where $f_U$ is the spectral density of $\{U_t\}$. Since both $f_U$ and $f_W$, the spectral density of $\{W_t\}$, are bounded, we must have that $\delta^N(B)$ and $\delta^N(F)$ divide $\Psi(B)$. That is, this is a condition such that (3) is valid. Now from $f_Y = f_S + f_N$ (which follows from (1) under the assumption that $\{U_t\}$ and $\{V_t\}$ are not correlated with one another) and (3) we obtain

$$f_N(\lambda) = \left(1 - \Psi\left(e^{-i\lambda}\right)\right) f_Y(\lambda)$$

and thus for $f_V$ the spectral density of $\{V_t\}$:

$$f_V(\lambda) = \left|\delta^N\left(e^{-i\lambda}\right)\right|^2 f_N(\lambda) = \frac{1 - \Psi\left(e^{-i\lambda}\right)}{\left|\delta^S\left(e^{-i\lambda}\right)\right|^2} f_W(\lambda)$$

We conclude that $\delta^S(B)$ and $\delta^S(F)$ must divide $1 - \Psi(B)$. In summary, we make the following assumptions in addition to (2) and (3):

$$\Gamma(B) = \frac{\Psi(B)}{\delta^N(B)\delta^N(F)} \quad \Phi(B) = \frac{1 - \Psi(B)}{\delta^S(B)\delta^S(F)} \tag{4}$$

where $\Gamma(B)$ and $\Phi(B)$ are bounded on the unit circle of the complex plane. It now follows from (3) and (4) that

$$f_U(\lambda) = \Gamma\left(e^{-i\lambda}\right) f_W(\lambda) \tag{5}$$

$$f_V(\lambda) = \Phi\left(e^{-i\lambda}\right) f_W(\lambda) \tag{6}$$

Now in both (5) and (6) we have our target spectra as the product of known (or computable) bounded functions multiplying $f_W$. Hence our estimates of these spectra are obtained by plugging in an estimate of $f_W$, which can be model-based or nonparametric, as desired. Ultimately, we need to compute the autocovariance matrices $\Sigma_U$ and $\Sigma_V$, whose $(j, k)$th entries are given by $\gamma_U(j - k)$ and $\gamma_V(j - k)$, the autocovariance functions of $\{U_t\}$ and $\{V_t\}$ respectively. These are estimated as follows:

$$\hat{\gamma}_U(h) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Gamma\left(e^{-i\lambda}\right) \hat{f}_W(\lambda) e^{i\lambda h} \mathrm{d}\lambda$$

$$\hat{\gamma}_V(h) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi\left(e^{-i\lambda}\right) \hat{f}_W(\lambda) e^{i\lambda h} \mathrm{d}\lambda$$

where $\hat{f}_W$ is an estimate of $f_W$. If $\hat{f}_W$ is the periodogram (with a continuous argument), then the above autocovariance estimates will be consistent and asymptotically normal under very mild conditions on the data—this follows from Lemma 3.1.1 of Taniguchi and Kakizawa (2000). Note that no smoothing of the periodogram is needed in this case, because the integration essentially performs this automatically. If $\hat{f}_W$ is a model-based estimate, then consistency will follow from having consistent parameter estimates, assuming that the true spectral density $f_W$ belongs to the model class that is selected.

With regard to efficient computation of the autocovariance estimates, it is not necessary to do any Riemann integration. When the periodogram is used, it is easily derived that (see McElroy and Holan, 2009)

$$\hat{\gamma}_U(h) = \frac{1}{n-d} W' \Sigma\left(\Gamma(e^{-i\cdot})e^{ih\cdot}\right) W$$

$$\hat{\gamma}_V(h) = \frac{1}{n-d} W' \Sigma\left(\Phi(e^{-i\cdot})e^{ih\cdot}\right) W$$

where $W = (W_1, W_2, \ldots, W_{n-d})'$. Now if $g(e^{-i\lambda})$ is a polynomial in in $e^{\pm i\lambda}$, then the corresponding Toeplitz matrix is easy to compute. In particular, suppose $g(e^{-i\lambda}) = \Sigma_{s=-r}^{q} g_s e^{-i\lambda s}$ for some positive integers $r$ and $q$. Then

$$\Sigma_{jk}\left(g(e^{-i\cdot})e^{ih\cdot}\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{s=-r}^{q} g_s e^{-i\lambda s} e^{i\lambda(j-k+h)} \mathrm{d}\lambda == g_{j-k+h}$$

allows us to find the matrix entries. If instead $\hat{f}_W$ corresponds to an ARMA process with parameters estimated from the data, the autocovariances are simply computed when $\Gamma(B)$ and $\Phi(B)$ have finite order. Note that $\Gamma(e^{-i\lambda})$ and $\Phi(e^{-i\lambda})$ are both real and positive by (4), and hence there exists spectral factorizations such that $\Gamma(e^{-i\lambda}) = |\Theta(e^{-i\lambda})|^2 \zeta^2$ and $\Phi(e^{-i\lambda}) = |\Xi(e^{-i\lambda})|^2 \xi^2$. Then supposing that $\hat{f}_W$ corresponds to the process $\phi(B)W_t = \theta(B)_{\epsilon_t}$ (with $\epsilon_t$ white noise of appropriate variance), then $\hat{\gamma}_U$ is the autocovariance of the process $\phi(B)W_t = \zeta\Theta(B)\theta(B)_{\epsilon_t}$, while $\hat{\gamma}_V$ is the autocovariance of the process $\phi(B)W_t = \xi\Xi(B)\theta(B)_{\epsilon_t}$. Thus the autocovariance estimates are very easy to calculate in practice. Now these autocovariance estimates fill out the entries of $\hat{\Sigma}_U$ and $\hat{\Sigma}_V$, which in turn can be plugged into the WK matrix smoother discussed below.

A matrix smoother is a matrix $F$ (it will be clear from context that this cannot be confused with the forward shift operator $F$) that is applied to a vector of observed data $Y = (Y_1, Y_2, \ldots, Y_n)'$, producing a vector estimate $\hat{S} = FY$. Whereas a filter denotes a collection of coefficients used to form linear combinations with the data, the term 'matrix smoother' refers to a collection of such filters that produce estimates at every time point. Hence a single row of $F$ is a filter, whereas all the rows taken together is a matrix smoother. (This terminology derives from the state space literature; see Durbin and Koopman, 2001.) Since $\hat{S}$ is intended as an estimate of $S$, the error associated with the matrix smoother is

$$\epsilon = \hat{S} - S = (F-1)S + FN$$

where 1 denotes the identity matrix. If $1 - F$ and $F$ do not remove the nonstationarity in signal and noise respectively, the error will grow unboundedly with sample size; hence the matrix smoother is generally assumed to satisfy

$$1 - F = G\Delta^S \quad F = H\Delta^N$$

for some matrices $G$ and $H$, with $\Delta^S$ and $\Delta^N$ as defined in McElroy (2008). These latter matrices accomplish the differencing of the polynomials $\delta^S$ and $\delta^N$ row by row. In general, let $\Lambda(g)$ be given by $\Lambda_{jk}(g) = g_{j-k+p}$, where $g$ is a polynomial of degree $p$ with coefficients $g_l$ (with the convention that



**DOI**: 10.1002/for

$g_l = 0$ if $l < 0$ or $l > p$). This matrix has dimension $(n - p) \times n$ by definition. Then define $\Delta^N = \Lambda(\delta^N)$, $\Delta^S = \Lambda(\delta^S)$, and $\Delta = \Lambda(\delta)$. Hence we can write

$$W = \Delta Y \quad U = \Delta^S S \quad V = \Delta^N N$$

where $W$, $U$, $V$, $S$, and $N$ are column vectors for $W_t$, $U_t$, $V_t$, $S_t$, and $N_t$. Thus the error process is $\epsilon = HV - GU$ for a general matrix smoother. We are generally interested in the error covariance matrix $\Sigma_t$, whose diagonal entries are the signal extraction MSEs (i.e., they are $\mathbb{E}(\hat{S}_\epsilon - S_t)^2$):

$$\Sigma_\epsilon = H\Sigma(f_V)H' + G\Sigma(f_U)G'$$

Now $G$ and $H$ are given to us by the definition of the matrix smoother, whereas $f_V$ and $f_U$ are determined either by model estimation or by the recasting methods described above.

The WK matrix smoother, defined below, has the property that $FY$ is the minimum MSE linear estimate of $S$ under some assumptions (McElroy, 2008), and is identical to the Kalman smoother or state space smoother (Durbin and Koopman, 2001). Its formula is

$$F = \Sigma_\epsilon \Delta^{N'} \Sigma_V^{-1} \Delta^N$$

$$\Sigma_\epsilon^{-1} = \Delta^{S'} \Sigma_U^{-1} \Delta^S + \Delta^{N'} \Sigma_V^{-1} \Delta^N$$

Clearly, given the estimates of $\Sigma_U$ and $\Sigma_V$ from the recast method, one can easily calculate these matrices. The error covariance matrix is just $\Sigma_\epsilon$ (this does not take parameter uncertainty into account, as is typical in the literature). Other matrix smoothers can be found in Pollock (2000, 2002).

## ILLUSTRATIONS

This section contains several extended examples that are of popular interest: the HP filter, the Henderson filter, ideal low-pass and band-pass filters, and some X-11 seasonal filters.

### HP filtering

The HP filter is popular in econometrics, both as a low-pass trend filter and as a cycle filter; to produce estimates of cycles, the complement of the HP filter is used. The filter is defined by

$$\Psi(B) = \frac{q}{q + (1 - B)^2(1 - F)^2}$$

where $q$ is a smoothness parameter, which can be interpreted as a signal-to-noise ratio (SNR). It follows from the condition (4) that it will be appropriate to take $\delta^N(B) = 1$ and $\delta^S(B) = (1 - B)^2$, so that the noise corresponds to a stationary component and the signal is an $I(2)$ trend. (Of course, one can also let the signal be an $I(1)$ trend, or even be a stationary component; the resulting spectral calculations for these cases are left to the reader.) Then $\Gamma(B) = \Psi(B)$ and $\Phi(B) = (q + |1 - B|^4)^{-1}$ (we use the shorthand $\bar{B} = F$), and

$$f_U(\lambda) = \frac{q}{q + |1 - e^{-i\lambda}|^4} f_W(\lambda) \quad f_V(\lambda) = \frac{1}{q + |1 - e^{-i\lambda}|^4} f_W(\lambda)$$

 
**DOI**: 10.1002/for

**Henderson trend**

The analysis for the Henderson trend filter $H(B)$ is extremely similar to that of the HP, since $1 - H(B)$ contains a factor of $(1 - B)^4$. This is always true, no matter the length of the Henderson, since all of these filters pass cubic polynomials. Hence the signal is an $I(2)$ trend with $\delta^S(B) = (1 - B)^2$, while the noise is stationary with $\delta^N(B) = 1$. Then $\Gamma(B) = \Psi(B) = H(B)$, and

$$\Phi(B) = \frac{1 - H(B)}{|1 - B|^4}$$

for any length of Henderson. Below, we consider $\Phi^q(B)$ for the lengths $q = 5, 7, 9, 13, 15, 17, 23$. Only $q = 9, 17, 23$ are used in X-12-ARIMA, but $q = 15, 17$ have been used by the Australian Bureau of Statistics (see Findley *et al.*, 1998, for a discussion):

$$\Phi^5(B) = 0.07343$$

$$\Phi^7(B) = 0.05874(B + F) + 0.17622$$

$$\Phi^9(B) = 0.04072(B^2 + F^2) + 0.17277(B + F) + 0.32826$$

$$\Phi^{13}(B) = 0.01935(B^4 + F^4) + 0.10526(B^3 + F^3) + 0.30495(B^2 + F^2) + 0.60014(B + F) + 0.82520$$

$$\begin{aligned}\Phi^{15}(B) = {} & 0.01373(B^5 + F^5) + 0.07942(B^4 + F^4) + 0.24943(B^3 + F^3) \\ & + 0.55209(B^2 + F^2) + 0.93283(B + F) + 1.19115\end{aligned}$$

$$\begin{aligned}\Phi^{17}(B) = {} & 0.00996(B^6 + F^6) + 0.06021(B^5 + F^5) + 0.19972(B^4 + F^4) \\ & + 0.47500(B^3 + F^3) + 0.89046(B^2 + F^2) + 1.35820(B + F) + 1.64924\end{aligned}$$

$$\begin{aligned}\Phi^{23}(B) = {} & 0.00428(B^9 + F^9) + 0.02803(B^8 + F^8) + 0.10214(B^7 + F^7) \\ & + 0.27202(B^6 + F^6) + 0.58803(B^5 + F^5) + 1.08709(B^4 + F^4) \\ & + 1.76721(B^3 + F^3) + 2.55807(B^2 + F^2) + 3.29197(B + F) + 3.67926\end{aligned}$$

Since the choice of $\Phi^q(B)$ is determined by the user, we see that the Henderson filter is defined via

$$H^q(B) = 1 - \Phi^q(B)|1 - B|^4$$

Thus the spectra are

$$f_U(\lambda) = H^q(e^{-i\lambda}) f_W(\lambda) \quad f_V(\lambda) = \Phi^q(e^{-i\lambda}) f_W(\lambda)$$

**Ideal low-pass and band-pass**

Next we discuss the ideal low-pass filter given by $\Psi(e^{-i\lambda}) = 1_{[-\omega,\omega]}(\lambda)$, with $\omega \in (0, \pi)$. Note that we are no longer in the class of ARIMA-type filters, which are rational functions in $B$ and $F$. Now the signal is obviously a trend, and we can let $\delta^S(B) = (1 - B)^d$ for any desired $d$, since



**DOI**: 10.1002/for

$$\Phi(e^{-i\lambda}) = \frac{1-\Psi(e^{-i\lambda})}{|1-e^{-i\lambda}|^{2d}} = 1_{[-\omega,\omega]^c}(\lambda)|1-e^{-i\lambda}|^{-2d}$$

which is a bounded function of $\lambda$ since the pole of the differencing operator is multiplied by the value zero. Moreover, the noise differencing operator can consist of practically anything whose zeroes all lie on the unit circle at angles between $\omega$ and $\pi$. Letting $\delta^N(B)$ denote such an operator, we have

$$\Gamma(e^{-i\lambda}) = \frac{\Psi(e^{-i\lambda})}{|\delta^N e^{-i\lambda}|^2} = 1_{[-\omega,\omega]}(\lambda)|\delta^N(e^{-i\lambda})|^{-2}$$

which again is a bounded function. Now we can compute the needed spectra, which turn out to be

$$f_U(\lambda) = 1_{[-\omega,\omega]}(\lambda)|\delta^N(e^{-i\lambda})|^{-2} f_W(\lambda) \quad f_V(\lambda) = 1_{[-\omega,\omega]^c}(\lambda)|1-e^{-i\lambda}|^{-2d} f_W(\lambda)$$

The autocovariances are found by numerical integration as follows:

$$\gamma_U(h) = \frac{1}{2\pi}\int_{-\omega}^{\omega}|\delta^N(e^{-i\lambda})|^{-2} f_W(\lambda)e^{i\lambda h}\mathrm{d}\lambda$$

$$\gamma_V(h) = \frac{1}{2\pi}\int_{[-\omega,\omega]^c}|1-e^{-i\lambda}|^{-2d} f_W(\lambda)e^{i\lambda h}\mathrm{d}\lambda$$

A simple adjustment of these ideas allows us to handle the band-pass filter as well. So now $\Psi(e^{-i\lambda}) = 1_{A\cup -A}(\lambda)$, where $A \subset (0, \pi]$. The signal can be nonstationary, with operator $\delta^S(B)$ with zeroes on the unit circle and with angles lying only in $A \cup -A$; similarly $\delta^N(B)$ can be any operator with zeroes on the unit circle with angles only in $A^c \cap (-A)^c$. The formulas for the autocovariances are then

$$\gamma_U(h) = \frac{1}{2\pi}\int_{A\cup -A}|\delta^N(e^{-i\lambda})|^{-2} f_W(\lambda)e^{i\lambda h}\mathrm{d}\lambda$$

$$\gamma_V(h) = \frac{1}{2\pi}\int_{A^c\cap(-A)^c}|\delta^S(e^{-i\lambda})|^{-2} f_W(\lambda)e^{i\lambda h}\mathrm{d}\lambda$$

**Seasonal adjustment**

We consider the scenario of seasonal adjustment, where the nonstationary operators are $U(B) = 1 + B + \ldots + B^{11}$ for the seasonal (for monthly data) and $(1 - B)$ for the trend. First suppose that our given generic seasonal adjustment filter is

$$\mu(B) = \frac{1}{24}U(B)(1+B)F^6$$

which is known in the X-11 literature as the $2 \times 12$ trend filter, or 'crude trend' filter. Since the frequency response function of $\mu(B)$ is negative at some frequencies, we will instead set



**DOI**: 10.1002/for

$\Psi(B) = \mu(B)\mu(F)$ (but of course $\mu(B) = \mu(F)$). We see that $\delta^N(B) = U(B)$ and $\Gamma(B) = (24)^{-2}|1 + B|^2$. Also

$$1-\Psi(B)=(1-\mu(B))(1+\mu(B))=|1-B|^2\,\Xi^{\mu}(B)(1+\mu(B))$$
$$\Xi^{\mu}(B)=\left(F^5+4F^4+9F^3+16F^2+25F+36+25B+16B^2+9B^3+4B^4+B^5\right)/24$$

so that $\delta^S(B) = 1 - B$ and $\Phi(B) = \Xi^{\mu}(B)(1 + \mu(B))$; since $\Phi(1) = 73/6$, no more unit root factors can be extracted, so the trend can be at most $I(1)$. Note that we could also include the $1 + B$ factor in $\delta^N(B)$, since it has unit roots, but that would imply that the frequency $\pi$ seasonal unit root occurs twice in the pseudo-spectrum, whereas the other seasonal unit roots only occur once. This is a strange scenario, so it is more natural to let $\delta^N(B) = U(B)$. Finally, we have

$$f_U(\lambda)=\frac{|1+e^{-i\lambda}|^2}{576}f_W(\lambda)\quad f_V(\lambda)=\Xi^{\mu}\left(e^{-i\lambda}\right)\left(1+\frac{U\left(e^{-i\lambda}\right)\left(1+e^{-i\lambda}\right)}{24}\right)f_W(\lambda)$$

An alternative seasonal adjustment filter that does not need to be squared is given by

$$\nu(B)=\frac{1}{144}U(B)U(F)$$

which satisfies (2). Letting $\Psi(B) = \nu(B)$, we have $\Gamma(B) = 1/144$ and $\delta^N(B) = U(B)$. In addition:

$$\frac{1-\Psi(B)}{(1-B)(1-F)}=\frac{a(B)a(F)}{144}$$

$$a(z)=0.093z^{10}+0.297z^9+0.634z^8+1.123z^7+1.788z^6+2.652z^5$$
$$+3.738z^4+5.070z^3+6.672z^2+8.570z+10.787$$

Hence $\delta^S(B) = 1 - B$ and $\Phi(B) = a(B)a(F)/144$, which has no poles on the unit circle. Then the implied spectra are

$$f_U(\lambda)=f_W(\lambda)/144\quad f_V(\lambda)=\frac{\left|a\left(e^{-i\lambda}\right)\right|^2}{144}f_W(\lambda)$$

**Seasonal estimation**

Here we consider the various seasonal moving averages of X-11, i.e., the $3 \times p$ filters where $p = 3, 5, 7, 9$ Let $\nu_j(B)=\frac{1}{j}\frac{B^{12j}-1}{B^{12}-1}B^{-12(j-1)/2}$, so that

$$\lambda_p(B)=\nu_3(B)\nu_p(B)$$

is the $3 \times p$ seasonal filter, by definition. Note that $p$ is always an odd integer. We use the notation $\lambda_p$ for the filter, following the treatment of Bell and Monsell (1992); this should not be confused with the frequency argument $\lambda$. Now $\nu_j(B)$ has all of its many roots on the unit circle; indeed, letting



**DOI**: 10.1002/for

$Z = B^{12}$, we know that $Z^j - 1$ has all $12j$ roots located at the $12j$th roots of unity, i.e., $e^{i\pi k/(12j)}$ for $k = 1, \ldots, 12j$, so that $(Z^j - 1)/(Z - 1)$ has roots of the form $e^{i\pi k/(12j)}$ for $k$s that are not a multiple of $j$. Letting $g_k$ denote the $k$th cyclotomic polynomial (the monic polynomial with zeroes given by the distinct $k$ roots of unity), we have the following by Proposition 8.2 of Hungerford (1974):

$$\frac{Z^3-1}{Z-1} = g_9(B)g_{18}(B)g_{36}(B)$$

$$\frac{Z^5-1}{Z-1} = g_5(B)g_{10}(B)g_{15}(B)g_{20}(B)g_{30}(B)g_{60}(B)$$

$$\frac{Z^7-1}{Z-1} = g_7(B)g_{14}(B)g_{21}(B)g_{28}(B)g_{42}(B)g_{84}(B)$$

$$\frac{Z^9-1}{Z-1} = g_9(B)g_{18}(B)g_{27}(B)g_{36}(B)g_{54}(B)g_{108}(B)$$

The first few cyclotomic polynomials are given by $g_1(x) = x - 1$, $g_2(x) = x + 1$, $g_3(x) = x^2 + x + 1$, and $g_4(x) = x^2 + 1$; the others can be determined recursively if desired. Thus these seasonal filters $\lambda_p$ suppress various frequencies of the type $e^{i\pi k/36}$ and $e^{i\pi k/(12p)}$, with $k$ such that the seasonal frequencies $e^{i\pi j/12}$ (with $j = 1, \ldots, 6$) are not suppressed. Since there is no natural nonstationary noise process to associate to these unit roots, we will generally suppose that the noise (i.e., the nonseasonal) is stationary.

Setting $\Psi(B) = \lambda_p(B)\lambda_p(F)$ (and note that $\lambda_p(B) = \lambda_p(F)$), we have $\delta^N(B) = 1$ and $\Gamma(B) = |\lambda_p(B)|^2$. On the other hand:

$$\begin{aligned} 1-\Psi(B) &= (1-\lambda_p(B)) + \lambda_p(B)(1-\lambda_p(F)) \\ &= (1-\nu_3(B)) + \nu_3(B)(1-\nu_p(B)) + \nu_3(B)\nu_p(B)[(1-\nu_3(F)) + \nu_3(F)(1-\nu_p(F))] \end{aligned}$$

Letting $p = 2q + 1$ we obtain

$$1-\nu_{2q+1}(B) = \frac{|1-Z|^2}{2q+1}\left(\sum_{j=2}^{q}\binom{j}{2}[Z^{q+1-j} + \bar{Z}^{q+1-j}] + \binom{q+1}{2}\right)$$

(when $q = 1$, treat the empty summation as zero), which is symmetric in $B$ and $F$. Let $\Xi^p(B) = (1 - \nu_{2q+1}(B))/|1 - B|^2$, and with $\delta^S(B) = 1 - Z$ we get

$$\Phi^p(B) = \frac{1}{3} + \nu_3(B)\Xi^p(B) + \nu_3(B)\nu_p(B)\left[\frac{1}{3} + \nu_3(F)\Xi^p(F)\right] = \left[\frac{1}{3} + \nu_3(B)\Xi^p(B)\right](1+\lambda_p(B))$$

Hence the implied spectra are

$$f_U(\lambda) = |\lambda_p(e^{-i\lambda})|^2 f_W(\lambda) \quad f_V(\lambda) = \Phi^p(e^{-i\lambda})f_W(\lambda)$$

## RECASTING X-11

Recasting the X-11 filters for seasonal adjustment, trend, and irregular components is an important application of this work. However, a direct approach fails because none of these X-11 filters have frequency response function bounded between 0 and 1. Nevertheless it is possible to proceed, as the various component filters (i.e., the Henderson trend, the $2 \times 12$, and the seasonal moving averages) do satisfy the necessary properties when modified, as demonstrated above. We begin with the basic definition of the X-11 filters below, and show how they can be modified so that the methods described above (under 'Recasting') can be applied. Below, we construct the corresponding matrix smoothers and apply the method to several seasonal time series (with regression effects removed).

### Symmetrizing the X-11 filters

The general philosophy behind X-11 is that we first obtain a crude trend estimate and subtract this from the data; what is left consists of seasonal and irregular, and whatever is left of the trend (since we have crudely detrended). The next step is to apply a seasonal filter, which, however, assumes that only seasonal and irregular dynamics are present, effectively ignoring residual trend behavior. This in turn (after some renormalization) is subtracted from the data, resulting in a first estimate of the deseasonalized data. There is some seasonality left over, since the first round of seasonal adjustment will not be perfect; therefore the whole process can be repeated. This iteration scheme could be carried on indefinitely, but in X-11 it is only repeated once, and in the second iteration different filters can be used to do the trend and seasonal estimation parts of the algorithm. Detailed references on this procedure include Shiskin *et al.* (1967), Shiskin (1978), and Ladiray and Quenneville (2001); see also Bell and Hillmer (1984) for a historical discussion. Following the notation of Bell and Monsell (1992) and the section 'Illustrations' of this paper, we let $\mu$ denote the $2 \times 12$ 'crude trend' filter, $\lambda_{p_1}$ the first $3 \times p_1$ seasonal filter, $\lambda_{p_2}$ the second $3 \times p_2$ seasonal filter, and $H_q$ the Henderson trend of order $q$. We refer to the triple $(p_1, p_2, q)$ as a specification of the X-11 filters. Then the seasonal filter $\omega_S$ is defined via

$$\omega_S = (1-\mu)\lambda_{p_2}[1 - H_q(1-(1-\mu)\lambda_{p_1}(1-\mu))]$$

where the juxtaposition of filters is interpreted as polynomial multiplication, since each filter is a polynomial in $B$. The seasonal adjustment filter $\omega_N$, or nonseasonal filter, is

$$\omega_N = 1 - \omega_S$$

The trend and irregular filters are then given by

$$H_q\omega_N \quad (1-H_q)\omega_N$$

respectively, by definition. The notation for the four components is Seasonal (S), Nonseasonal (N), Trend (T), and Irregular (I). Thus the estimate of $S$ is given by $\omega_S Y$, and the estimate of $N$ is given by $\omega_N Y$. (We no longer use $S$ and $N$ for signal and noise, so $\delta^S$ and $\delta^N$ now denote differencing operators for the seasonal and nonseasonal components, respectively.)

As noted above, $\mu^2(B) = \mu(B)\mu(F)$ and $\lambda_p^2(B) = \lambda_p(B)\lambda_p(F)$ satisfy (2), which suggests replacing each *constituent* filter (i.e., $\mu$, $\lambda_p$, and $H_q$) by its symmetrization, or square, within $\omega_S$, $\omega_N$, etc.



**DOI**: 10.1002/for

Then, suppressing the frequency argument of the functions for clarity of presentation, we obtain utilizing (3):

$$f_S = (1-\mu^2)\lambda_{p_2}^2[1-H_q^2(1-(1-\mu^2)\lambda_{p_1}^2(1-\mu^2))]f_Y$$

$$f_N = \{1-(1-\mu^2)\lambda_{p_2}^2[1-H_q^2(1-(1-\mu^2)\lambda_{p_1}^2(1-\mu^2))]\}f_Y$$

$$f_T = H_q^2\{1-(1-\mu^2)\lambda_{p_2}^2[1-H_q^2(1-(1-\mu^2)\lambda_{p_1}^2(1-\mu^2))]\}f_Y$$

$$f_I = (1-H_q^2)\{1-(1-\mu^2)\lambda_{p_2}^2[1-H_q^2(1-(1-\mu^2)\lambda_{p_1}^2(1-\mu^2))]\}f_Y$$

Finally, we note that all of the constituent filters $\mu$, $\lambda_p$, and $H_q$ have squared magnitude bounded between 0 and 1, and hence the same property is true of one minus the squared gain function. Thus, by using recursion, we see that each of the four pseudo-spectra given above is equal to $f_Y$ multiplied by a function bounded between zero and one. We next determine the $\Gamma(B)$ and $\Phi(B)$ filters for each component, as defined in (4).

The seasonal component should have all trend nonstationarity removed. Now $1 - e^{-i\lambda}$ can be factored out of $1 - H_q^2(e^{-i\lambda})$ four times and out of $1 - \mu^2(e^{-i\lambda})$ twice. Since we can write

$$f_S = (1-\mu^2)\lambda_{p_2}^2\left[(1-H_q^2)+H_q^2(1-\mu^2)^2\lambda_{p_1}^2\right]f_Y$$

we see that a total of six factors of $1 - e^{-i\lambda}$ can be pulled out. Therefore we can set $\delta^N(B) = (1 - B)^d$ with $d = 0, 1, 2, 3$ as the nonseasonal (noise) differencing operator; then

$$\Gamma_S(e^{-i\lambda}) = \frac{\left\{(1-\mu^2)\lambda_{p_2}^2\left[(1-H_q^2)+H_q^2(1-\mu^2)^2\lambda_{p_1}^2\right]\right\}(e^{-i\lambda})}{|1-e^{-i\lambda}|^{2d}}$$

where this is a bounded function. We can determine $\Phi_S(B)$ by examining the nonseasonal component, since actually $\Phi_S = \Gamma_N$; it should have $U(B)$ as a noise differencing operator. By manipulation we obtain

$$f_N = \{(1-\lambda_{p_2}^2)+\mu^2\lambda_{p_2}^2(1-H_q^2)+\lambda_{p_2}^2H_q^2(1-\lambda_{p_1}^2)+\mu^2(3-3\mu^2+\mu^4)\lambda_{p_1}^2\lambda_{p_2}^2H_q^2\}f_Y$$

The first and third terms each admit two factors of $1 - e^{-i12\lambda}$ (see 'Seasonal estimation', above); the second term has two factors of $U(e^{-i\lambda})$ and four factors of $1 - e^{-i\lambda}$, which come from $\mu^2$ and $1 - H_q^2$ respectively. The fourth term has two factors of $U(e^{-i\lambda})$, but no factors of $1 - e^{-i\lambda}$. Hence the noise differencing operator is $\delta^S(B) = U(B)^D$ with $D = 0, 1$. Thus

$$\Gamma_N(e^{-i\lambda}) = \frac{\{(1-\lambda_{p_2}^2)+\mu^2\lambda_{p_2}^2(1-H_q^2)+\lambda_{p_2}^2H_q^2(1-\lambda_{p_1}^2)+\mu^2(3-3\mu^2+\mu^4)\lambda_{p_1}^2\lambda_{p_2}^2H_q^2\}(e^{-i\lambda})}{|U(e^{-i\lambda})|^{2D}}$$

which is a bounded function. The trend component's pseudo-spectrum is just $H_q^2$ times that of the nonseasonal, and in particular $\Gamma_T = H_q^2\Gamma_N$. For the irregular, the differencing operator must combine $U(B)^D$ and $(1 - B)^d$. However, since the irregular pseudo-spectrum is $1 - H_q^2$ times the nonseasonal's



DOI: 10.1002/for

pseudo-spectrum, we see that $d \leq 2$ must hold. That is, the implied model for the nonseasonal is consistent with $I(3)$ data, but then the irregular will not be stationary; we must restrict to at most an $I(2)$ process. We divide these differencing operators into $\Psi_I = \Psi_N - \Psi_T$, which yields

$$\Gamma_I(e^{-i\lambda}) = \frac{(1 - H_q^2(e^{-i\lambda}))\Psi_N(e^{-i\lambda})}{|1-e^{-i\lambda}|^{2d}|U(e^{-i\lambda})|^{2D}} = (1 + H_q(e^{-i\lambda}))\Phi^q(e^{-i\lambda})|1-e^{-i\lambda}|^{4-2d}\Gamma_N(e^{-i\lambda})$$

Now we can explicitly write down the spectra for the differenced components $U^S$, $U^N$, $U^T$, and $I$:

$$f_{U^S}(\lambda) = \Gamma_S(e^{-i\lambda}) f_W(\lambda)$$

$$f_{U^N}(\lambda) = \Gamma_N(e^{-i\lambda}) f_W(\lambda)$$

$$f_{U^T}(\lambda) = H_q^2(e^{-i\lambda})\Gamma_N(e^{-i\lambda}) f_W(\lambda)$$

$$f_I(\lambda) = \Gamma_I(e^{-i\lambda}) f_W(\lambda)$$

We note that these $\Gamma$ functions are all polynomials in $B$ and $F$, which facilitates calculating estimates of the autocovariance functions. They depend crucially on the choices of $d = 0, 1, 2$ and $D = 0, 1$ (though in practice $D = 1$ and $d = 1, 2$ are the most common possibilities). Note that our notation for $d$ and $D$ differs from that used for SARIMA models. For an illustration consider Figure 1, which depicts the functions $\Psi(e^{-i\lambda})$ for the various components, using three different specifications: (3, 3, 23), (3, 9, 9), and (15, 15, 9) (the reason for these choices is explained in below). By (3) the plotted functions can be multiplied by the data's true pseudo-spectrum to obtain the spectra for the components. Note that the functions are bounded between zero and one as planned, and have the 'right' spectral shapes. It is interesting that 'non-seasonality' in the nonseasonal and irregular components is indicated by spectral troughs at seasonal frequencies, rather than just monotonic behavior.

### X-11 matrix smoothers

We now describe the details of implementing X-11 matrix smoothers. The key is to determine the functions $\Gamma_S(B)$, $\Gamma_N(B)$, and $\Gamma_I(B)$, which are polynomials in $B$ and $F$. Based on the calculations in the previous subsection, we have the following formulas:

$$\begin{aligned}\Gamma_S(B) = {} & |1-B|^{6-2d}(1+\mu(B))\Xi^{\mu}(B)\lambda_{p2}^2(B)\Big[(1+H_q(B))\Phi^q(B) + \\ & H_q^2(B)(1+\mu(B))^2(\Xi^{\mu}(B))^2\lambda_{p1}^2(B)\Big]\end{aligned}$$

$$\begin{aligned}\Gamma_N(B) = {} & |U(B)|^{2-2D}\Big(|1-B|^2\left[\Phi^{p_2}(B) + \Phi^{p_1}(B)\lambda_{p_2}^2(B)H_q^2(B)\right] \\ & + 24^{-2}|1+B|^2\lambda_{p_2}^2(B)\left[1 - H_q^2(B) + (3 - 3\mu^2(B) + \mu^4(B))\lambda_{p_1}^2(B)H_q^2(B)\right]\Big)\end{aligned}$$

$$\Gamma_T(B) = H_q^2(B)\Gamma_N(B)$$

$$\Gamma_I(B) = |1-B|^{4-2d}\Phi^q(B)(1+H_q(B))\Gamma_N(B)$$



**DOI**: 10.1002/for

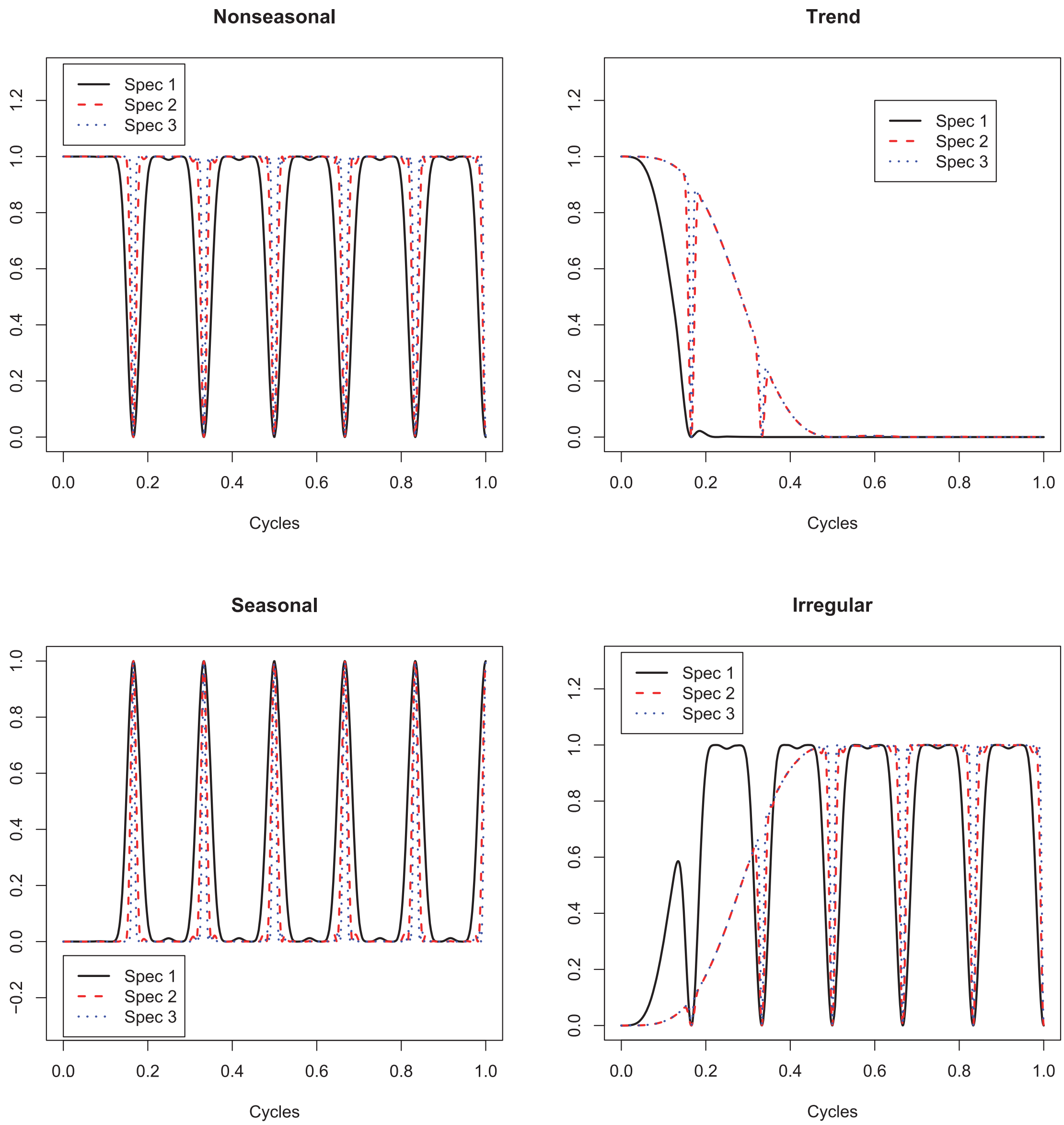


Figure 1. X-11 recast frequency response functions for nonseasonal, trend, seasonal, and irregular components, with three specifications: Spec 1 is (3, 3, 23); Spec 2 is (3, 9, 9); Spec 3 is (15, 15, 9). This figure is available in colour online at www.interscience.wiley.com/journal/for



**DOI**: 10.1002/for

The following notations are used in the above formulas:

$$\Xi^{\mu}(B) = \left(F^5 + 4F^4 + 9F^3 + 16F^2 + 25F + 36 + 25B + 16B^2 + 9B^3 + 4B^4 + B^5\right)/24$$

$$\Phi^{q}(B) = \frac{1 - H^{q}(B)}{|1-B|^2}$$

$$\Phi^{p}(B) = \left(\frac{1}{3} + \nu_3(B)\Xi^{p}(B)\right)(1 + \lambda_p(B))$$

The explicit formulas for $\Phi^q(B)$ for various $q$ are given above ('Henderson trend'). In order to compute the autocovariance matrices needed for the matrix smoothers (see 'Recasting' section above for formulas), we must determine $\Sigma(f\ e^{ih\cdot})$ for various $h$, and for $f$ equal to $f_{U^S}$, $f_{U^N}$, $f_{U^T}$, and $f_I$. As discussed above, there are several ways of estimating the data pseudo-spectral density; we considered the periodogram, an AR spectrum estimator (with order selection via AIC), and a model-based estimate coming from a fitted Box–Jenkins Airline model (Box and Jenkins, 1976).

As an illustration of these techniques, we examine US Retail Sales of Shoe Stores data from the monthly Retail Trade Survey of the Census Bureau (with outliers and trading day effects removed), from 1984 to 1998, which will be referred to as the *Shoe* series. In order to focus on the effect of specification on the components, we first consider only the periodogram-based estimates of the spectrum. Since the auto-model procedure of X-12-ARIMA indicates that $d = 2$, $D = 1$ for this series (which is quite common for seasonal data at the US Census Bureau; see Findley *et al.*, 1998), we use these values along with the three specifications (3, 3, 23), (3, 9, 9), and (15, 15, 9) to generate components estimates and RMSEs, estimating the autocovariances via the periodogram method. Figure 2 displays component estimates along with time-varying RMSEs for nonseasonal (which also applies to the seasonal) and trend; of course, RMSEs can be produced for the irregular but are not displayed here. Discrepancies between the component estimates are difficult to discern by the naked eye, but essentially the nonseasonals and trends coming from the first specification are smoother (this is to be expected since a 23-term Henderson is used, versus a 9-term one for the others); also the irregular for the third specification has the most variability. More obvious are the differences in RMSE, which have the characteristic 'U'-shape (heightened variability at the sample boundaries); the third specification has the highest error, but the first specification is lowest. Interestingly, the first specification is the closest 'fit' to the Shoe series in the sense of Chu *et al.* (2007) discussed below.

The procedure was implemented in R (R Development Core Team, 2005), which is a great advantage for the portability of the code. Computationally speaking, the algorithm's slowest aspect is the inversion of $n \times n$ matrices required for the matrix smoothers. Since these matrices depend on covariance estimates obtained from the data, it is not possible to obtain the same quantities using state space methods. Other aspects of the code involve polynomial additions and multiplications, and a few matrix multiplications. Hence the matrix inversions, which are of order $n^3$ flops, dominate the computational complexity. For the four series considered in this paper, runs of the program took anywhere from a split second (for series of around 150 observations) to almost a minute (for series of over 500 observations) on an old government computer.

 
**DOI**: 10.1002/for

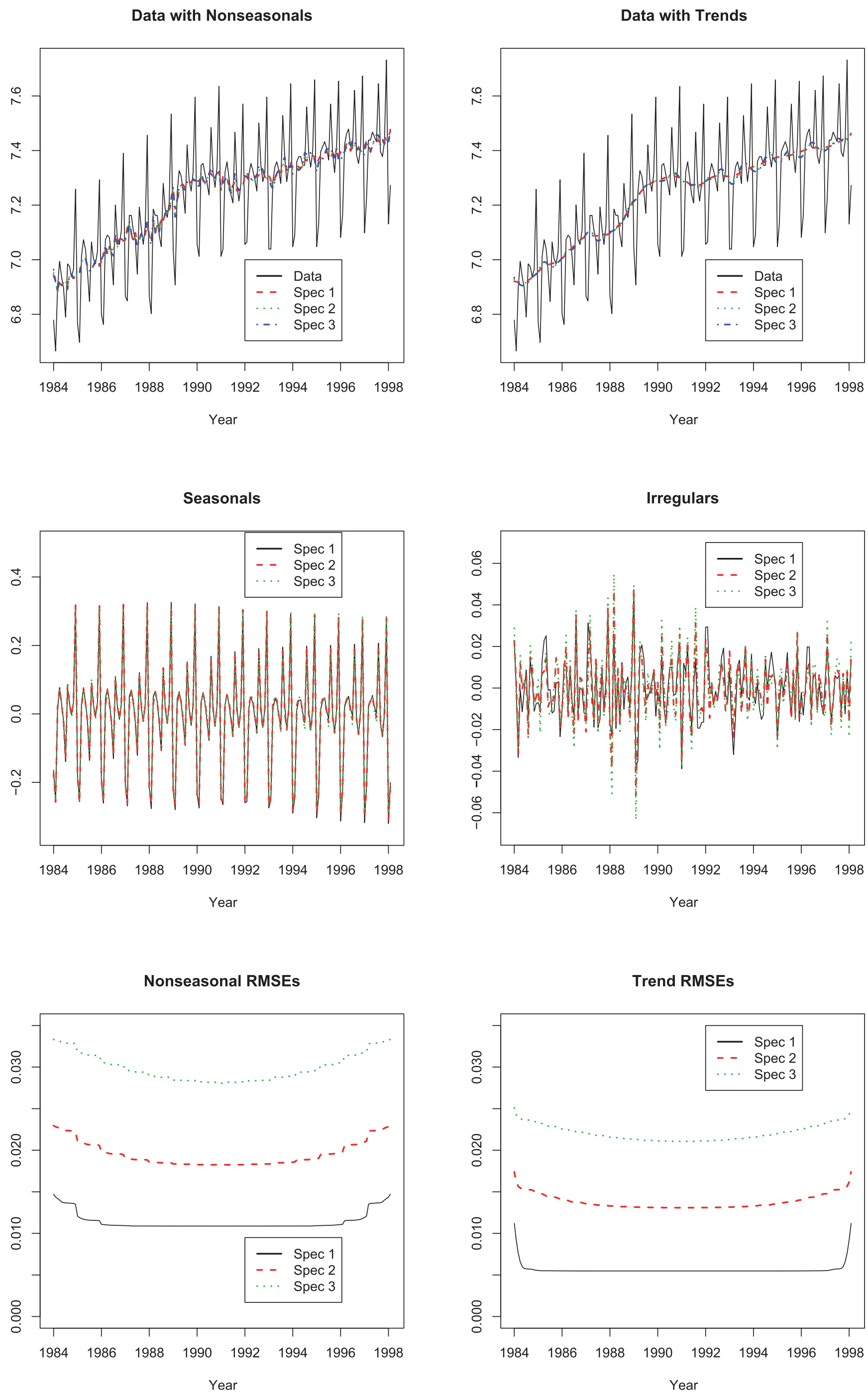


Figure 2. Component estimates from X-11 recast method for Shoe series, along with time-varying RMSE for nonseasonal and trend. Three specifications are used: Spec 1 is (3, 3, 23); Spec 2 is (3, 9, 9); Spec 3 is (15, 15, 9). This figure is available in colour online at www.interscience.wiley.com/journal/for

**Comparisons to X-11 filters**

Given that we can construct the above X-11 matrix smoother, how close is the approximation? The X-11 filter is a single symmetric filter, whereas our method constructs a bank of time-varying filters; it make sense to compare the center (or symmetric) filter of our method to X-11. In order to make the results independent of sample size, we take $\Psi(B)$ as an infinite sample approximation to the central filter. These functions have already been plotted in Figure 1, but that was a comparison of three specifications, whereas now we want to compare to X-11 frequency response functions. Figures 3–5 compare each of the three specifications in turn to X-11, considering nonseasonal, trend, seasonal, and irregular. There is a close similarity for the nonseasonal and seasonal components, but some discrepancy occurs for trend and irregular for specifications two and three; perhaps this is due to the use of the shorter Henderson. In all cases, the X-11 frequency response function can exceed unity and takes on negative values as well, whereas the recast filter provides an approximation bounded between zero and one.

Finally, it makes sense to compare the RMSEs produced by the recast method in practice. We consider the following four time series: *Shoe*, *Emp*, *Start*, and *Order*. The *Shoe* series is described above; the second series refers to Employed males, aged 16–19, covering the period January 1976 through October 2006. The *Start* series title is 'US Total New Privately Owned Housing Units Started; Thousands; NSA, Census Bureau', and the *Order* series title is 'Manufacturing: Nondefense Capital Goods: New Orders: Millions of Dollars: NSA, Census Bureau'. All of these series were first adjusted for trading day and outlier effects using the X-12-ARIMA program (Findley *et al.*, 1998). The adjusted series were then run through the X-11 recast matrix smoother (utilizing a log transformation for *Emp* and *Order*). In order to find appropriate X-11 specifications, we fitted Airline models to all four series (a larger collection of series was originally considered, but these four can reasonably be modeled by the Airline model and furthermore facilitate the illustration of our main points).

The significance of the Airline model is that it is the linchpin to MBA techniques that match model-based filters to X-11; the most current literature on this is Chu *et al.* (2007), which provides a grid of Airline model parameters and their corresponding 'closest' X-11 specification (that is, the seasonal adjustment filters determined by the given Airline model parameters is closest in the authors' particular metric to the listed X-11 specification, among all other X-11 specifications). We note that the *automdl* routine of X-12-ARIMA prefers an Airline model for *Shoe* and *Emp*, whereas a $(0\ 1\ 1)(1\ 1\ 1)_{12}$ is preferred for *Start* and a $(1\ 1\ 1)(0\ 1\ 1)_{12}$ for *Order*; nevertheless the AIC for the Airline model was not much greater for these two series, and was a reasonable fit overall. The estimated parameters for the series are (first value is the nonseasonal moving average parameter, whereas the second is the seasonal moving average parameter, using Box–Jenkins conventions): (0.57, 0.34) for *Shoe*; (0.30, 0.70) for *Emp*; (0.27, 0.89) for *Start*; (0.51, 0.73) for *Order*. Using the chart in Chu *et al.* (2007), we obtain specification (3, 3, 23) for *Shoe*, (3, 9, 9) for *Emp* and *Order*, and (15, 15, 3) for *Start*. For the coherency of results, we have chosen to analyze these three specifications in the previous figures.

We next produce four types of RMSEs for each series, focusing on the Nonseasonal and Trend components: we produce the three types of spectral density estimates (periodogram, AR spectrum, and model-based) as described earlier. Here the model-based spectrum is that of the corresponding Airline model, and thus may well be more accurate than the other two methods (assuming a correct model specification). As a fourth technique, we supply the RMSE corresponding to the bi-infinite MBA filter (this value is computed by SEATS, the model-based seasonal adjustment program of the Bank of Spain—Maravall and Caparello (2004)), which in general lies



**DOI**: 10.1002/for

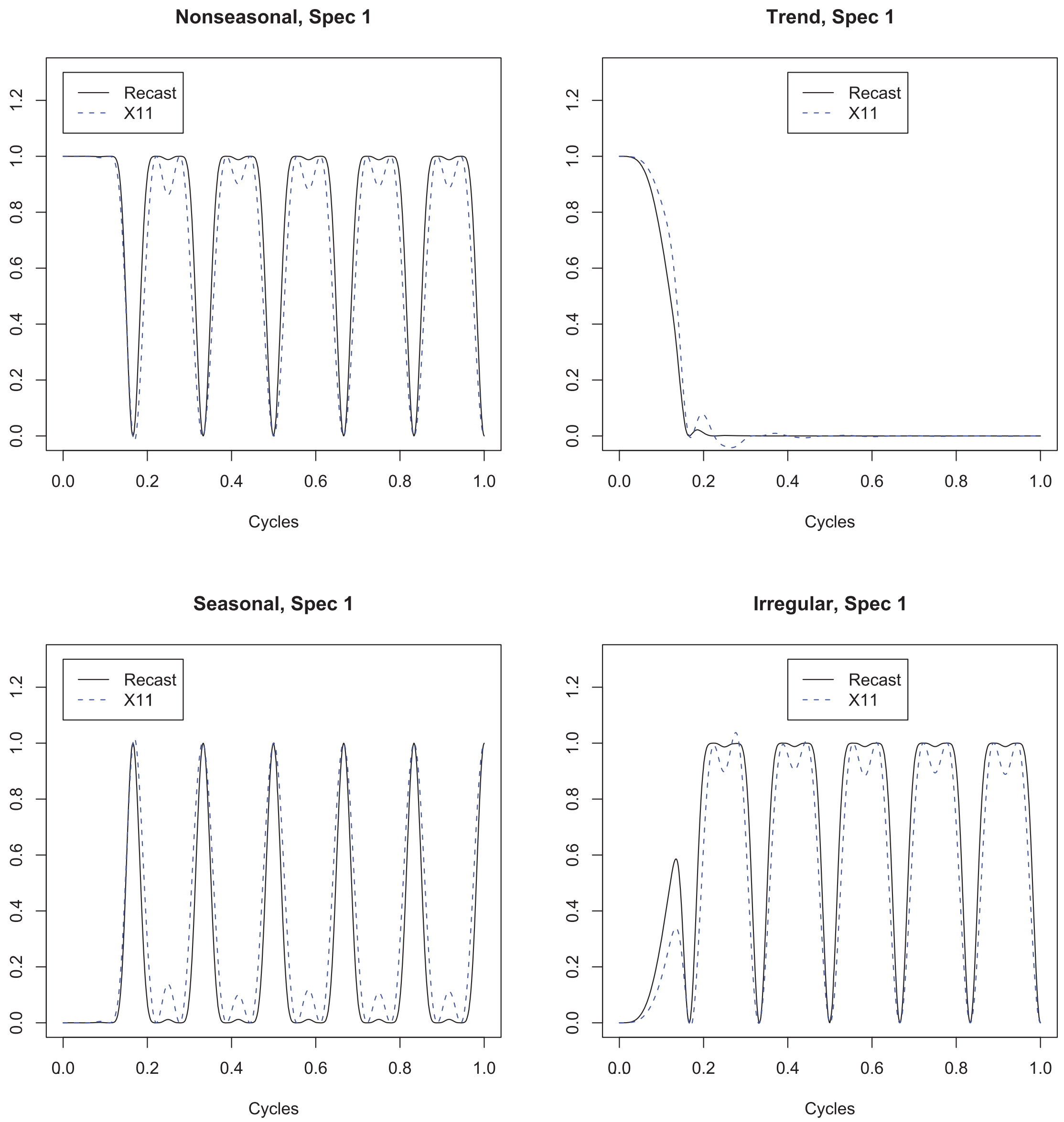


Figure 3. Comparisons of frequency response functions for X-11 versus X-11 recast, for nonseasonal, trend, seasonal, and irregular components, with specification (3, 3, 23). This figure is available in colour online at www.interscience.wiley.com/journal/for

underneath the time-varying MSE curve—this is referred to as WK. This forms a crude basis of comparison with a completely different method: for each data set we take the fitted airline model, from which signal extraction estimates and RMSE values can be obtained. These are roughly comparable to the recasting techniques, because the X-11 specifications were chosen to be close to the



**DOI**: 10.1002/for

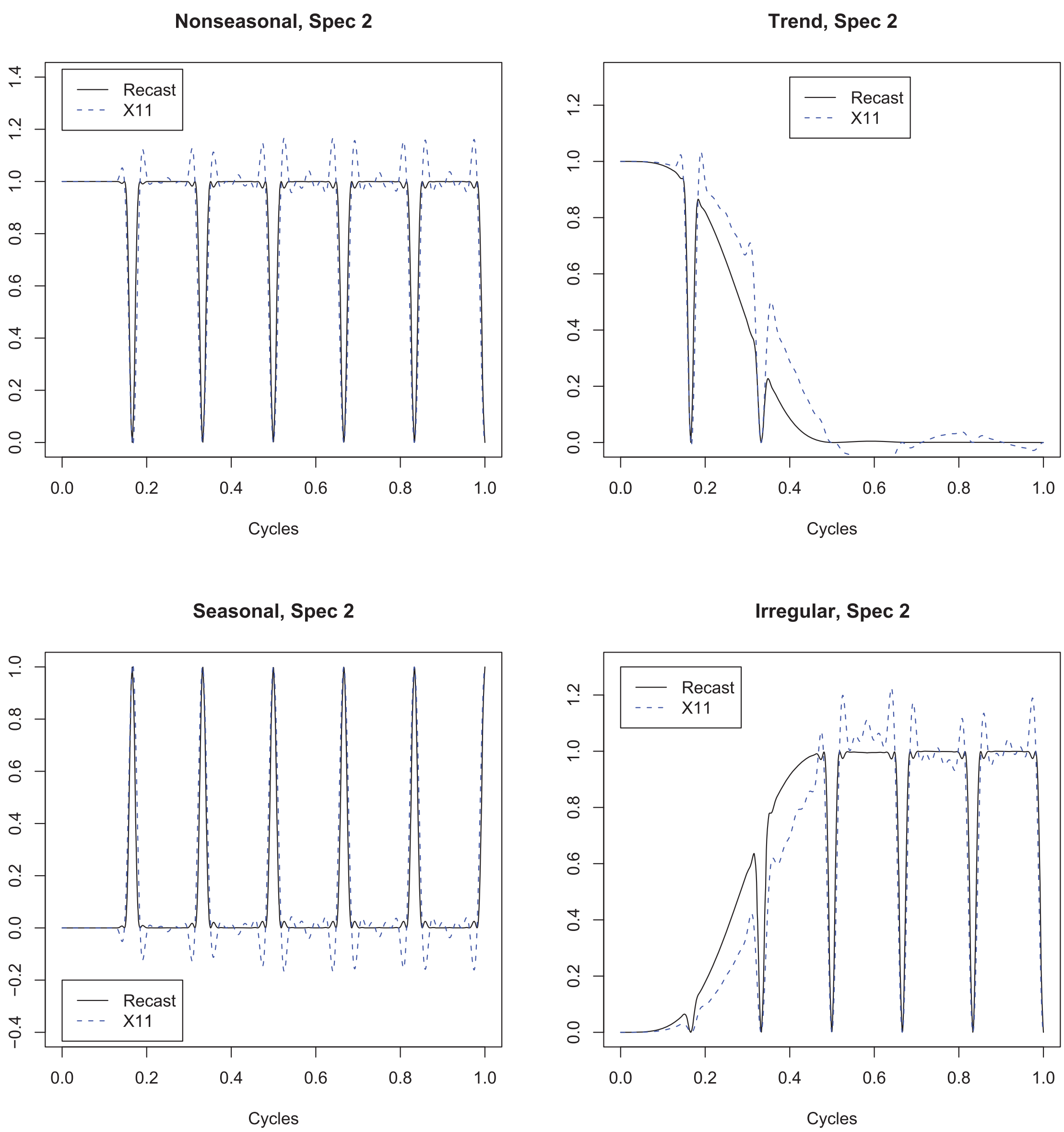


Figure 4. Comparisons of frequency response functions for X-11 versus X-11 recast, for nonseasonal, trend, seasonal, and irregular components, with specification (3, 9, 9). This figure is available in colour online at www.interscience.wiley.com/journal/for

airline models according to the method of Chu, Tiao, and Bell (2007). Note that SEATS also produces the time-varying RMSE based on a semi-infinite sample (Maravall and Caparello, 2004). Figures 6 and 7 display these comparisons. The general pattern is that RMSE is highest for the AR spectrum and lowest for MBA spectrum, and the WK line is close to the others. Note that the WK



**DOI**: 10.1002/for

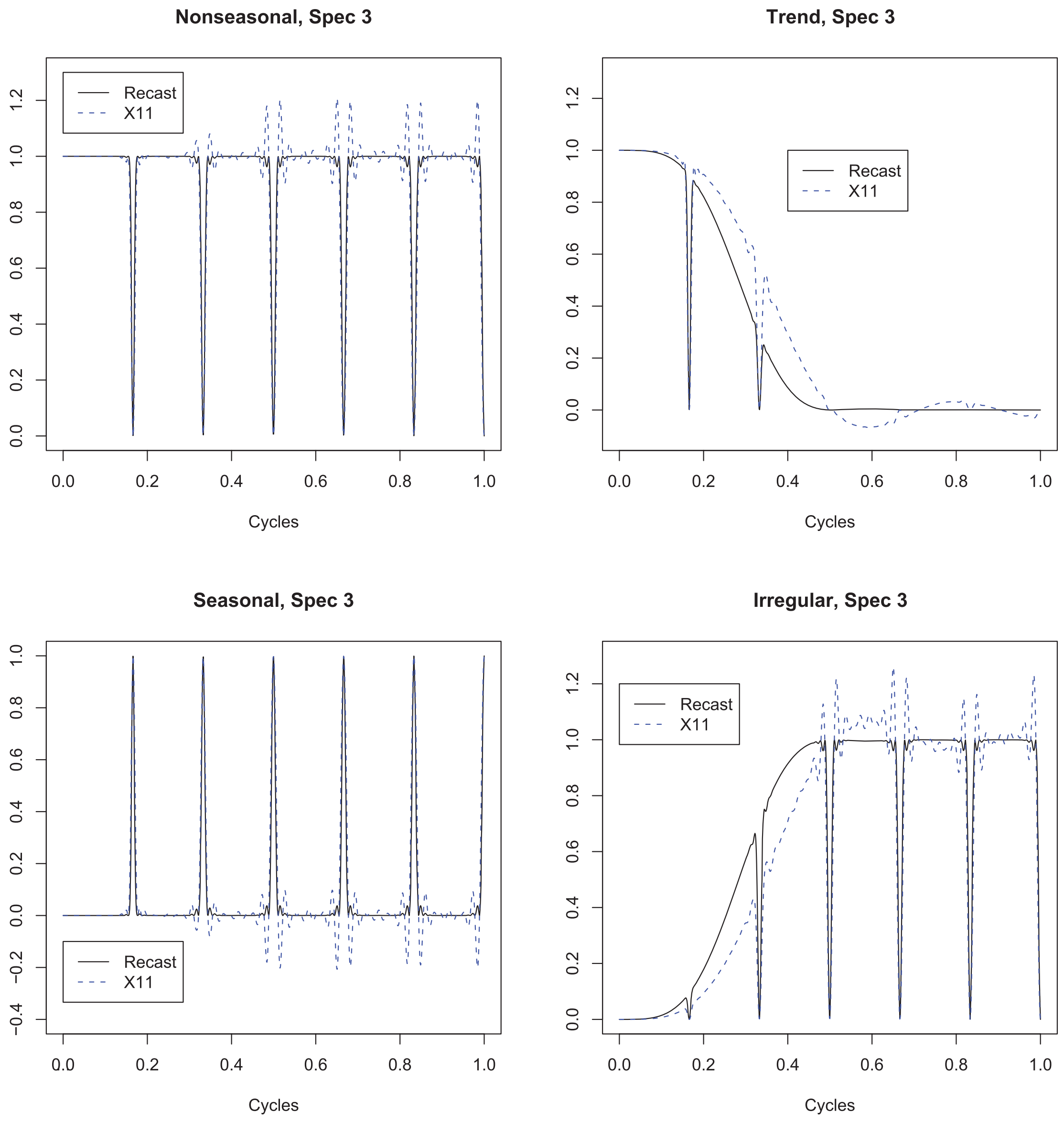


Figure 5. Comparisons of frequency response functions for X-11 versus X-11 recast, for nonseasonal, trend, seasonal, and irregular components, with specification (15, 15, 9). This figure is available in colour online at www.interscience.wiley.com/journal/for



**DOI**: 10.1002/for

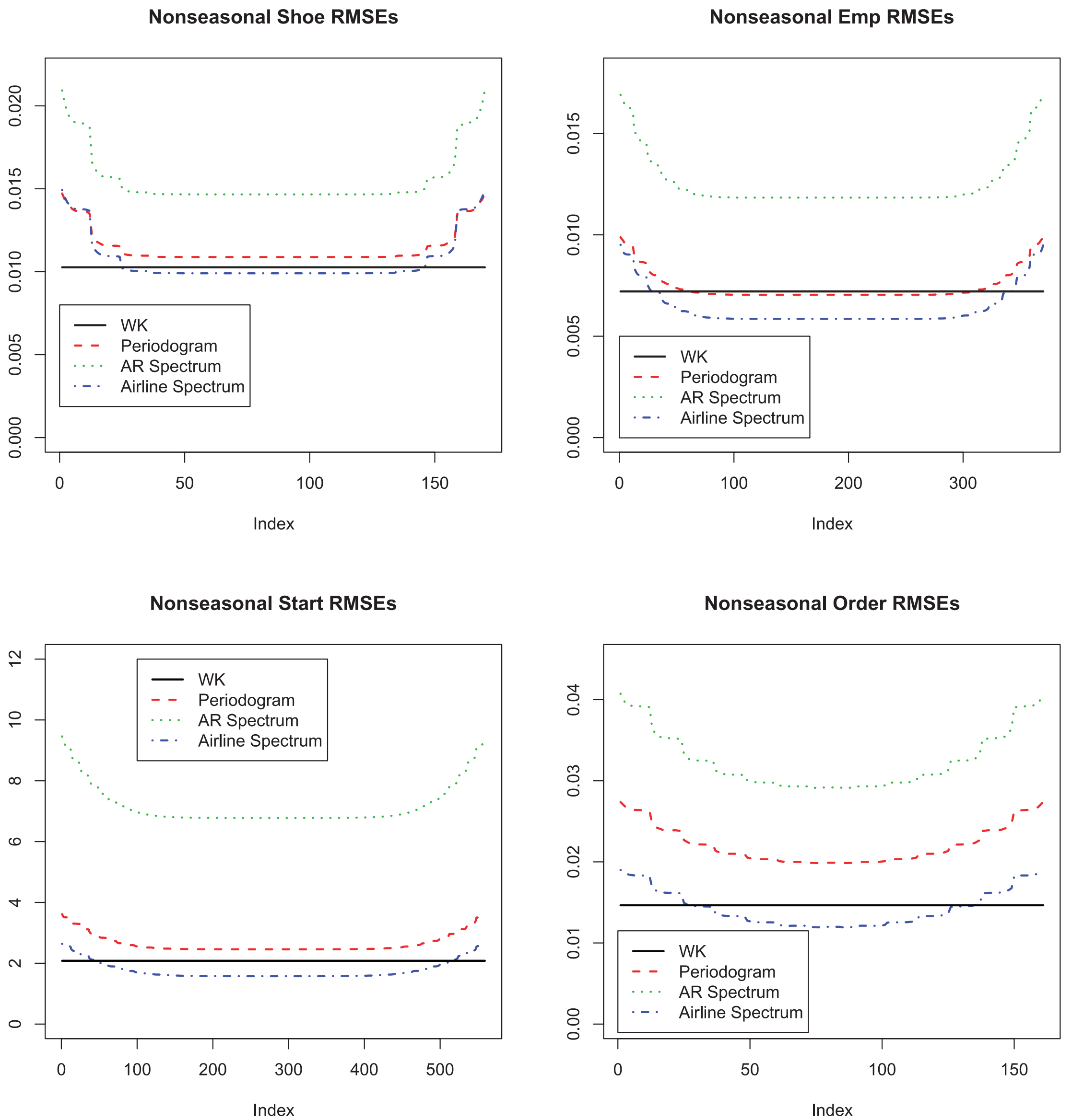


Figure 6. Comparisons of nonseasonal RMSE from X-11 recast for Shoe, Emp, Start, and Order series. Three methods of computing the spectrum are utilized: periodogram, AR spectral estimation, and Airline model spectrum. Also plotted is the RMSE from the WK filter corresponding to the fitted Airline model. This figure is available in colour online at www.interscience.wiley.com/journal/for



**DOI**: 10.1002/for

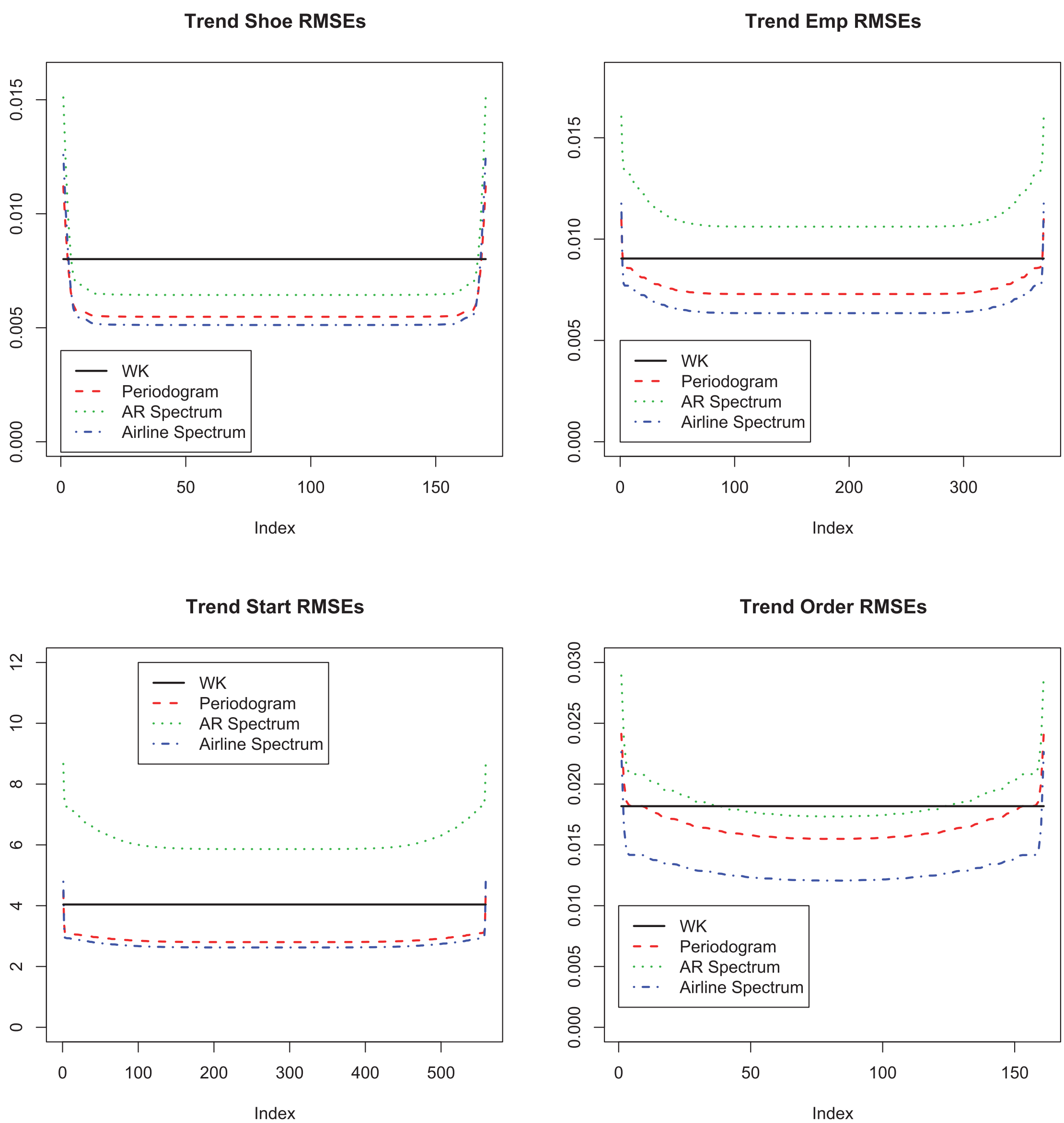


Figure 7. Comparisons of trend RMSE from X-11 recast for Shoe, Emp, Start, and Order series. Three methods of computing the spectrum are utilized: periodogram, AR spectral estimation, and Airline model spectrum. Also plotted is the RMSE from the WK filter corresponding to the fitted Airline model. This figure is available in colour online at www.interscience.wiley.com/journal/for



**DOI**: 10.1002/for

RMSE need not underlie the Airline RMSE, since the functions $\Gamma$ and $\Phi$ multiplying the pseudo-spectrum estimate in (5) and (6) are generated by the particular specification of the recast method, and thus do not correspond to exact Airline model quantities; nevertheless, we might expect the WK RMSE and Airline RMSE to be close, which is the case. Also, it is wrong to say the Airline method is better than the AR method just because the RMSE is lower—this RMSE might be understated (i.e., an underestimate of the true quantity). Our own perspective is that the periodogram method is the most statistically robust, since it will be consistent even with heavy-tailed data and in cases where it is difficult to correctly specify an ARIMA model.

## CONCLUSIONS

We have developed a nonparametric method for both extending X-11 filters to the boundary as well as obtaining time-varying signal extraction MSEs. It is known that one can extend a given bi-infinite filter to finite asymmetric filters via using forecast and backcast extensions of the data. Our method is implicitly doing this, but with the forecasts determined by the differencing polynomial and by the spectrum estimate of the differenced series. This nonparametric approach has the practical advantage that it does not require the correct identification of a model. (However, our method can be integrated with an MBA if desired by using model-based spectral estimates.) Moreover, there is no need to derive ARIMA models for the various components as in the original implied models method of Kaiser and Maravall (2005). Thus, in situations where large numbers of series are being seasonally adjusted by practitioners who are not experts in time series modeling—such as at federal statistical agencies—there is some appeal to this non-MBA method of filter extension. The trade-off is that nonparametric methods will not be as accurate as MBA when the model is correctly specified.

Our approach has been developed very generally, so that our method can be easily extended to many other filters, such as those discussed above under 'Illustrations' (e.g., Henderson, ideal lowpass). Our implementation of the method with the X-11 filters is very flexible: any combination of seasonal and Henderson filters is allowed, and six choices of unit root specifications are allowed as well—and highly portable, being written in R. One proviso is that our algorithm does not handle Box–Cox transforms, unit root identification, or estimation of regression effects, which are important aspects of any decent seasonal adjustment method. The first two aspects would typically be handled by the practitioner through time series plots and autocorrelation plots (which are very easy to do in R as well); the latter aspect is more complex and requires a more sophisticated program such as X-12-ARIMA.

## ACKNOWLEDGEMENTS

The author thanks Agustin Maravall and David Findley for stimulating discussions on this topic, and Bill Bell for careful reading of the manuscript. He also thanks an anonymous referee for helpful comments.



**DOI**: 10.1002/for

## REFERENCES


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

Bell W. 2005. Some consideration of seasonal adjustment variances. In *Proceedings of the American Statistical Association: Section on Survey Research Methods* (CD-ROM). American Statistical Association: Alexandria, VA; 2747–2758.

Bell W, Hillmer S. 1984. Issues involved with the seasonal adjustment of economic time series (with discussion). *Journal of Business and Economic Statistics* **2**: 291–320.

Bell W, Kramer M. 1996. *Toward variances for X-11 seasonal adjustments*. SRD Research Report No. RR-96/07, US Census Bureau. Available: http://www.census.gov/srd/www/byname.html [5 March 2010].

Bell W, Monsell B. 1992. *X-11 symmetric linear filters and their transfer functions*. SRD Research Report No. RR-92/15, US Census Bureau. Available: http://www.census.gov/srd/www/bynam [5 March 2010].

Box G, Jenkins G. 1976. *Time Series Analysis*. Holden-Day: San Francisco, CA.

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

Chu Y, Tiao G, Bell W. 2007. *A mean squared error criterion for comparing X-12-ARIMA and model-based seasonal adjustment filters*. SRD Research Report No. RRS2007/10, US Census Bureau. Available: http://www.census.gov/srd/www/byname.html [5 March 2010].

Cleveland W, Tiao G. 1976. Decomposition of seasonal time series: a model for the Census X-11 program. *Journal of the American Statistical Association* **71**: 581–587.

Dagum E. 1980. *The X-11-ARIMA Seasonal Adjustment Method*, No. 12-564E. Statistics Canada: Ottawa.

Depoutot R, Planas C. 1998. *Comparing seasonal adjustment and trend extraction filters with application to a model-based selection of X11 linear filters*. Eurostat Working Paper, No. 9/1998/A/9, Luxembourg.

Durbin J, Koopman S. 2001. *Time Series Analysis by State Space Methods*. Oxford University Press: Oxford.

Findley DF, Monsell BC, Bell WR, Otto MC, Chen BC. 1998. New capabilities and methods of the X-12-ARIMA seasonal adjustment program (with discussion). *Journal of Business and Economic Statistics* **16**: 127–177.

Hungerford T. 1974. *Algebra*. Springer: New York.

Kaiser R, Maravall A. 2005. Combining filter design with model-based filtering: an application to business-cycle estimation. *International Journal of Forecasting* **21**: 691–710.

Ladiray D, Quenneville B. 2001. *Seasonal Adjustment with the X-11 Method*. Springer: New York.

Maravall A, Caparello G. 2004. *Program TSW: Revised reference manual*. Working Paper 2004, Research Department, Bank of Spain (http://www.bde.es).

McElroy T. 2008. Matrix formulas for nonstationary ARIMA signal extraction. *Econometric Theory* **24**: 1–22.

McElroy T, Holan S. 2009. A local spectral approach for assessing time series model misspecification. *Journal of Multivariate Analysis* **100**: 604–621.

Pfeffermann D. 1994. A general method for estimating the variances of X-11 seasonally adjusted estimators. *Journal of Time Series Analysis* **15**: 85–116.

Pierce D. 1980. Data revision with moving average seasonal adjustment procedures. *Journal of Econometrics* **14**: 95–114.

Pollock D. 2000. Trend estimation and de-trending via rational square-wave filters. *Journal of Econometrics* **99**: 317–334.

Pollock D. 2002. Circulant matrices and time-series analysis. *International Journal of Mathematical Education in Science and Technology* **33**: 213–230.

President's Committee to Appraise Employment and Unemployment Statistics. 1962. *Measuring Employment and Unemployment*. US Government Printing Office: Washington, DC.

R Development Core Team. 2005. *R: A Language and environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (http://www.R-project.org).

Shiskin J. 1978. Seasonal adjustment of sensitive indicators. In *Seasonal Analysis of Economic Time Series*, Zellner A (ed.). US Department of Commerce, Bureau of the Census: Washington DC; 97–103.

Shiskin J, Young A, Musgrave J. 1967. *The X-11 variant of the Census Method II seasonal adjustment program*. Technical Paper No. 15, US Department of Commerce, Bureau of Economic Analysis, Washington, DC.

Taniguchi M, Kakizawa Y. 2000. *Asymptotic Theory of Statistical Inference for Time Series*. Springer: New York.

Wolter K, Monsour N. 1981. On the problem of variance estimation for a deseasonalized series. In *Current Topics in Survey Sampling*, Krewski D, Platek R, Rao J (eds). Academic Press: New York; 199–226.

*Author's biography*:
**Tucker McElroy** graduated from Columbia University in 1996 with a B.A. in mathematics, and earned a PhD in pure mathematics from the University of California, San Diego in 2001, followed by a postdoctoral fellowship at the same institution. In 2003 he began work at the Census Bureau, where he is currently a principal researcher in the area of seasonal adjustment. His research interests include filtering and model-fitting.

*Author's address*:
**Tucker McElroy**, Statistical Research Division, US Census Bureau, 4600 Silver Hill Road, Washington, DC 20233–9100, USA.



**DOI**: 10.1002/for