RESEARCH REPORT SERIES
(Statistics #2007-15)

A Nonparametric Method for Asymmetrically Extending Signal Extraction Filters

Tucker McElroy

Statistical Research Division
U.S. Census Bureau
Washington, DC 20233

Report Issued: November 13, 2007

A Nonparametric Method for Asymmetrically Extending Signal Extraction Filters

Tucker McElroy[*]
U.S. Census Bureau

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 the mean squared error optimal linear estimate of the signal given 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 a single time series and provide comparisons with X-11-ARIMA seasonal adjustments.

Keywords. ARIMA model, Nonstationary time series, Seasonal adjustment, X-11.

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 U.S. Census Bureau.

1 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

[*]Statistical Research Division, U.S. Census Bureau, 4700 Silver Hill Road, Washington, D.C. 20233-9100, tucker.s.mcelroy@census.gov

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 (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, Tiao, and Bell (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. The matrix smoother is derived according to either of two paradigms (output-matching and Wiener-Kolmogorov, or WK) discussed below. 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). Section 2 discusses this theory, with details on the output-matching and WK recasting paradigms. We also discuss how a matrix smoother and error covariance matrix can be generated from the spectra for signal and noise. In Section 3 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 Section 4 we discuss an approximation to the X-11 filter that resolves these difficulties. The output-matching X-11 matrix smoother is then constructed and illustrated on a seasonal time series. Some technical results on the estimation of autocovariance functions is left in the Appendix.

2 Recasting Paradigms

We here introduce our basic notation for the paper. 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). 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

$$U_t = \delta^S(B) S_t \qquad V_t = \delta^N(B) N_t \tag{2}$$

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 δ, and d_S and d_N are the orders of δ^S and δ^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 γ:

$$\gamma(h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) e^{i\lambda h} \, d\lambda.$$

The autocovariances can be arranged into a symmetric Toeplitz matrix Σ 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 Σ_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 $\Sigma_X = \Sigma(f)$).

2.1 The Output-Matching Paradigm

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 $\lambda \in [-\pi, \pi]$. Below we will require that

$$0 \leq \Psi^2(e^{-i\lambda}) \leq 1. \tag{3}$$

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). Let the filter output be

$$\tilde{S}_t = \Psi(B) Y_t,$$

with t any integer. Of course, the true signal S_t is an unknown stochastic component, and a component of Y_t. The output-matching paradigm operates on the assumption that the dynamics of $\tilde{S}_t$ and S_t are identical, i.e., their pseudo-spectra are equal. This is stated in the equation

$$f_S(\lambda) = f_{\tilde{S}}(\lambda) = \Psi^2(e^{-i\lambda}) f_Y(\lambda), \tag{4}$$

where the first equality is the defining property of the paradigm, and the second equality is known to be true for stationary processes. The following discussion indicates that (4) can be extended to nonstationary data, so long as it is interpreted correctly. Since $\Psi(B)$ is a signal extraction filter, at a minimum it must annihilate deterministic noise components and reduce N_t to stationarity, so we may write

$$\Psi(B) = \Gamma(B) \delta^N(B)$$

for some function $\Gamma(B)$ with bounded frequency response (i.e., no poles on the unit circle). Now in analogy with $U_t = \delta^S(B) S_t$, we define

$$\tilde{U}_t = \delta^S(B) \tilde{S}_t = \Gamma(B) W_t,$$

where the second equality trivially follows from the commutativity of polynomial operators. Hence

$$f_{\tilde{U}}(\lambda) = \left| \Gamma(e^{-i\lambda}) \right|^2 f_W(\lambda),$$

where all of these functions are bounded in λ. Now dividing this through by $|\delta^S(e^{-i\lambda})|^2$, we obtain (4) as an equality of pseudo-spectra (although these pseudo-spectra have poles at the signal frequencies, the interpretation at these frequencies is given by the above expression in terms of spectra of differenced signal).

Some comments on the output-matching paradigm are in order. Some authors have thought it desirable that UC estimates have similar dynamics to the target; Wecker (1979) – also see Ansley and Wecker (1984) – actually developed filters with this criterion in mind, called square-root WK filters. It is well-known that WK filters do not have this property (Bell and Hillmer, 1984).

Hence we have a defining equation for f_S and f_U, in terms of f_Y. Since we assume that the components are orthogonal, it follows that

$$f_N(\lambda) = \left(1 - \left(\Psi(e^{-i\lambda})\right)^2\right) f_Y(\lambda).$$

We note that this is implied by the relation $f_Y = f_S + f_N$; the above is *not* the pseudo-spectra of $\tilde{N}_t = Y_t - \tilde{S}_t$, which would be $(1 - \Psi(e^{-i\lambda}))^2 f_Y(\lambda)$. Now to generate a matrix smoother (see Section 2.3) we require a knowledge of f_U and f_V, since δ^S and δ^N are already presumed to be known. These should be bounded functions. It follows from the above discussion that

$$f_U(\lambda) = f_S(\lambda)\left|\delta^S(e^{-i\lambda})\right|^2 = \left|\Gamma(e^{-i\lambda})\right|^2 f_W(\lambda) \tag{5}$$
$$f_V(\lambda) = f_N(\lambda)\left|\delta^N(e^{-i\lambda})\right|^2 = \frac{\left(1 - \left(\Psi(e^{-i\lambda})\right)^2\right)}{|\delta^S(e^{-i\lambda})|^2} f_W(\lambda).$$

We see that in order for f_V to be bounded, it is necessary that the signal differencing operator cancel out with $1-\Psi^2(B)$, which factors into $1-\Psi(B)$ times $1+\Psi(B)$. Noting that the former term is the noise extraction filter, it should accomplish signal differencing. Of course, it is unrealistic to assume in practice that $1+\Psi(B)$ would accomplish any signal differencing. However, note that we are dividing by $\delta^S(B)$ *as well as* $\delta^S(F)$; therefore we need the following assumption to proceed:

$$1 - \Psi(B) = \Phi(B)\delta^S(B)\delta^S(F) \tag{6}$$

for some function $\Phi(B)$ with no poles on the unit circle. Then under the assumption (6), we have

$$f_V(\lambda) = \Phi(e^{-i\lambda})\left(1 + \Psi(e^{-i\lambda})\right) f_W(\lambda). \tag{7}$$

Now in both (5) and (7) we have our target spectra as the product of known (or computable) bounded functions multiplying f_W. Hence, our estimates 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 Σ_U and Σ_V, whose (j,k)th entries are given by $\gamma_U(j-k)$ and $\gamma_V(j-k)$. These are estimated as follows:

$$\widehat{\gamma}_U(h) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\Gamma(e^{-i\lambda})\right|^2 \widehat{f}_W(\lambda)e^{i\lambda h}\,d\lambda$$
$$\widehat{\gamma}_V(h) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi(e^{-i\lambda})\left(1 + \Psi(e^{-i\lambda})\right)\widehat{f}_W(\lambda)e^{i\lambda h}\,d\lambda,$$

where $\widehat{f}_W$ is an estimate of f_W. If $\widehat{f}_W$ 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. Note that no smoothing of the periodogram is needed in this case, because the integration essentially performs this automatically. If $\widehat{f}_W$ 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. Since the periodogram method gives consistency regardless of model-class assumptions, we focus our treatment on this case (though those who prefer model-based methods may easily adapt our method to their favorite model).

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, 2005)

$$\widehat{\gamma}_U(h) = \frac{1}{n-d} W' \Sigma \left(|\Gamma(e^{-i\cdot})|^2 e^{ih\cdot} \right) W$$
$$\widehat{\gamma}_V(h) = \frac{1}{n-d} W' \Sigma \left(\Phi(e^{-i\cdot}) \left(1 + \Psi(e^{-i\cdot})\right) e^{ih\cdot} \right) W,$$

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

$$\Sigma_{jk} \left(g(e^{-i\cdot}) e^{ih\cdot} \right) = g_{j-k+h}.$$

Thus the autocovariance estimates are very easy to calculate in practice. Now these autocovariance estimates fill out the entries of $\widehat{\Sigma}_U$ and $\widehat{\Sigma}_V$, which in turn can be plugged into the WK matrix smoother discussed in Section 2.3 below.

2.2 The WK Paradigm

This discussion is very similar to the approach of the previous subsection, but there are a few subtle differences. Again we let $\Psi(B)$ be our generic filter, and we still require (3). Our main assumption is that the dynamics of the output match those of a WK estimate, i.e.,

$$f_{\tilde{S}}(\lambda) = \frac{f_S^2(\lambda)}{f_Y(\lambda)}.$$

See Bell (1984) for a demonstration of this result (assuming Assumption A on the initial values), when $\tilde{S}_t$ is the WK estimate of S_t. Since $f_{\tilde{S}}(\lambda) = \Psi^2(e^{-i\lambda}) f_Y(\lambda)$ as well, after taking square roots we obtain

$$f_S(\lambda) = |\Psi(e^{-i\lambda})| f_Y(\lambda).$$

Note that this is a little more general than what we would obtain assuming that $\Psi(B)$ is a WK filter – that would require that $\Psi(e^{-i\lambda})$ be non-negative. But in our formulation, the frequency response can be negative so long as it is real and satisfies (3). The noise pseudo-spectrum is

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

since $f_S + f_N = f_Y$. It is required to determine f_U and f_V.

$$f_U(\lambda) = f_S(\lambda) |\delta^S(e^{-i\lambda})|^2 = \left(\frac{\Psi(e^{-i\lambda})\Psi(e^{i\lambda})}{\delta^N(e^{-i\lambda})^2 \delta^N(e^{i\lambda})^2} \right)^{1/2} f_W(\lambda),$$

from which we see the requirement that $\delta^N(B)\delta^N(F)$ divides $\Psi(B)$. Hence we write

$$\Psi(B) = \Gamma(B)\delta^N(B)\delta^N(F).$$

Note that this implies a different definition of $\Gamma(B)$ from that implied in the output-matching paradigm. Thus we obtain

$$f_U(\lambda) = |\Gamma(e^{-i\lambda})| f_W(\lambda). \tag{8}$$

Secondly we have

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

which is obtained by multiplying the top and bottom of $f_N(\lambda)$ by $1 + |\Psi(e^{-i\lambda})|$. Now $1 - \Psi^2(B) = (1 - \Psi(B))(1 + \Psi(B))$, so we require that $\delta^S(B)\delta^S(F)$ divides $1 - \Psi(B)$, i.e., (6). In that case, we have

$$f_V(\lambda) = \frac{1 + \Psi(e^{-i\lambda})}{1 + |\Psi(e^{-i\lambda})|}\,\Phi(e^{-i\lambda}) f_W(\lambda). \tag{9}$$

So the autocovariance functions can be estimated by

$$\begin{aligned}
\widehat{\gamma}_U(h) &= \frac{1}{n-d} W' \Sigma \left(|\Gamma(e^{-i\cdot})| e^{ih\cdot} \right) W \\
\widehat{\gamma}_V(h) &= \frac{1}{n-d} W' \Sigma \left(\frac{1 + \Psi(e^{-i\cdot})}{1 + |\Psi(e^{-i\cdot})|}\,\Phi(e^{-i\cdot})\, e^{ih\cdot} \right) W,
\end{aligned}$$

taking the nonparametric approach. One practical difficulty lies in the determination of $|\Gamma(e^{-i\lambda})|$ and $1 + |\Psi(e^{-i\lambda})|$. Even if $\Gamma(B)$ and $\Psi(B)$ are polynomials in B and F, their absolute values need not be. Hence the determination of the Toeplitz matrices will require numerical integration in general, and the simple method outlined in Section 2.1 for getting ACF estimates will not apply. Thus, the implementation for the output-matching paradigm is potentially much easier.

2.3 Matrix Smoothers

A matrix smoother is a matrix F that is applied to a vector of observed data $Y = (Y_1, Y_2, \cdots, Y_n)'$, 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

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

for some matrices G and H, with Δ^S and Δ^N as defined in McElroy (2005). These latter matrices accomplish the differencing of the polynomials δ^S and δ^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 $g_l = 0$ 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

$$W = \Delta Y \qquad U = \Delta_S S \qquad V = \Delta_N N$$

where W, U, V, S, and N are column vectors for W_t, U_t, V_t, S_t, and N_t. We also define further differencing matrices $\underline{\Delta}_N$ and $\underline{\Delta}_S$ with row entries $\delta^N_{i-j+d_N}$ and $\delta^S_{i-j+d_S}$ given by the coefficients of $\delta^N(z)$ and $\delta^S(z)$ respectively, which are $(n-d) \times (n-d_S)$ and $(n-d) \times (n-d_N)$ dimensional. It follows from Lemma 1 of McElroy and Sutcliffe (2006) that

$$\Delta = \underline{\Delta}_N \Delta_S = \underline{\Delta}_S \Delta_N. \tag{10}$$

Thus the error process is

$$\epsilon = HV - GU$$

for a general matrix smoother. We are generally interested in the error covariance matrix Σ_ϵ, whose diagonal entries are the signal extraction MSEs (i.e., they are $\mathbb{E}(\hat{S}_t - S_t)^2$):

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

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, 2005), 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 Σ_U and Σ_V from the recast method, one can easily calculate these matrices. The error covariance matrix is just Σ_ϵ. Other matrix smoothers can be found in Pollock (2000, 2002).

3 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.

3.1 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

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

where q is a smoothness parameter, which can be interpreted as a signal-noise ratio (SNR). For either of the output-matching or WK paradigms, 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.) So $\Psi(B) = HP(B)$, and

$$\begin{aligned} 1 - \Psi(B) &= \frac{(1-B)^2(1-F)^2}{q + (1-B)^2(1-F)^2} \\ \Phi(B) &= \frac{1}{q + (1-B)^2(1-F)^2}, \end{aligned}$$

which holds for both paradigms. Hence for the output-matching paradigm we have

$$\begin{aligned} f_U(\lambda) &= \left(\frac{q}{q + |1 - e^{-i\lambda}|^4} \right)^2 f_W(\lambda) \\ f_V(\lambda) &= \frac{2q + |1 - e^{-i\lambda}|^4}{\left(q + |1 - e^{-i\lambda}|^4 \right)^2} f_W(\lambda). \end{aligned}$$

For the WK paradigm we have

$$\begin{aligned} f_U(\lambda) &= \frac{q}{q + |1 - e^{-i\lambda}|^4} f_W(\lambda) \\ f_V(\lambda) &= \frac{1}{q + |1 - e^{-i\lambda}|^4} f_W(\lambda). \end{aligned}$$

3.2 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 $\delta^N(B) = 1$. Letting

$$\Phi(B) = \frac{1 - H(B)}{(1-B)^2(1-F)^2},$$

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

$$\begin{aligned}
\Phi^5(B) &= .07343 \\
\Phi^7(B) &= .05874(B+F) + .17622 \\
\Phi^9(B) &= .04072(B^2+F^2) + .17277(B+F) + .32826 \\
\Phi^{13}(B) &= .01935(B^4+F^4) + .10526(B^3+F^3) + .30495(B^2+F^2) + .60014(B+F) + .82520 \\
\Phi^{15}(B) &= .01373(B^5+F^5) + .07942(B^4+F^4) + .24943(B^3+F^3) \\
&\quad + .55209(B^2+F^2) + .93283(B+F) + 1.19115 \\
\Phi^{17}(B) &= .00996(B^6+F^6) + .06021(B^5+F^5) + .19972(B^4+F^4) \\
&\quad + .47500(B^3+F^3) + .89046(B^2+F^2) + 1.35820(B+F) + 1.64924 \\
\Phi^{23}(B) &= .00428(B^9+F^9) + .02803(B^8+F^8) + .10214(B^7+F^7) + .27202(B^6+F^6) + .58803(B^5+F^5) \\
&\quad + 1.08709(B^4+F^4) + 1.76721(B^3+F^3) + 2.55807(B^2+F^2) + 3.29197(B+F) + 3.67926
\end{aligned}$$

Thus for the output-matching paradigm, the signal and noise spectra are

$$\begin{aligned}
f_U(\lambda) &= H^2(e^{-i\lambda}) f_W(\lambda) \\
f_V(\lambda) &= \Phi^p(e^{-i\lambda}) \left(1 + H(e^{-i\lambda})\right) f_W(\lambda).
\end{aligned}$$

For the WK paradigm (noting that the frequency response of $H(B)$ is always non-negative), the spectra are

$$\begin{aligned}
f_U(\lambda) &= H(e^{-i\lambda}) f_W(\lambda) \\
f_V(\lambda) &= \Phi^p(e^{-i\lambda}) f_W(\lambda).
\end{aligned}$$

3.3 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

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

which is a bounded function of λ 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 ω and π. 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 identical for both paradigms.

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

noting that $\Psi^2(e^{-i\lambda}) = \Psi(e^{-i\lambda})$, etc. The autocovariances are found by numerical integration as follows:

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

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 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

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

3.4 Seasonal Adjustment

We consider the scenario of seasonal adjustment, where the nonstationary operators are $U(B) = 1 + B + \cdots + B^{11}$ for the seasonal (for monthly data) and $(1-B)^m$ for the trend, where $m = 1, 2$ for $I(1)$ or $I(2)$ trends. We use the output-matching paradigm, although the WK paradigm could be used for the second filter below.

SA for $I(1)$ 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 2×12 trend filter, or "crude trend" filter. Letting $\Psi(B) = \mu(B)$, we immediately see that $\Psi(e^{-i\lambda}) \leq 1$, but the function is actually negative (though greater than negative one) at some frequencies. In any case $\Psi^2(e^{-i\lambda}) \leq 1$, so that (3) is satisfied.

So $U(B)$ is the obvious candidate for noise differencing operator $\delta^N(B)$; we could include the $1+B$ factor in $\delta^N(B)$ as well, since it has unit roots, but that would imply that the frequency π seasonal unit root occurs twice in the pseudo-spectrum, whereas the other seasonal unit roots only occur once. This would seem to be a strange scenario, so it is more natural to let $\delta^N(B) = U(B)$. Thus $\Gamma(B) = (1+B)/24$. Now we have

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

so that the natural signal differencing operator is $\delta^S(B) = 1-B$, which is associated with an $I(1)$ trend. Thus we have

$$\Phi(B) = \left(F^5+4F^4+9F^3+16F^2+25F+36+25B+16B^2+9B^3+4B^4+B^5\right)/24,$$

which allows us to define f_U and f_V:

$$\begin{aligned} f_U(\lambda) &= \frac{|1+e^{-i\lambda}|^2}{576} f_W(\lambda) \\ f_V(\lambda) &= \Phi(e^{-i\lambda})\left(1+\frac{U(e^{-i\lambda})(1+e^{-i\lambda})}{24}\right) f_W(\lambda). \end{aligned}$$

From here any desired matrix smoothers can be generated.

SA for $I(2)$ trend Alternatively, consider the seasonal adjustment filter

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

So that we can encompass an $I(2)$ trend, we set $\Psi(B)$ equal to $1-\nu(B)$, so that the signal is the seasonal and the noise is the trend. Since $(1-B)^2$ divides $\Psi(B)$ (shown below), we have $\delta^N(B) = (1-B)^2$. Likewise,

$$1-\Psi(B) = \nu(B) = \delta^S(B)\delta^S(F)/144$$

with $\delta^S(B) = U(B)$. In this case, $\Phi(B) = 1/144$. In addition,

$$\begin{aligned} \frac{\Psi(B)}{(1-B)(1-F)} &= \frac{a(B)a(F)-c(B)c(F)}{144} \\ a(z) &= z^9+3z^8+6z^7+10z^6+15z^5+21z^4+28z^3+36z^2+45z+55 \\ c(z) &= z^{10}+2z^9+3z^8+4z^7+5z^6+6z^5+7z^4+8z^3+9z^2+10z+11. \end{aligned}$$

Since $1-F = -F(1-B)$, this clearly implies that $(1-B)^2$ divides $1-\nu(B)$, as claimed. Hence

$$\Gamma(B) = -\frac{B}{144}\left(a(B)a(F)-c(B)c(F)\right),$$

which has no poles on the unit circle (interpreting F as $1/B$). Then the implied spectra are

$$\begin{aligned} f_U(\lambda) &= \frac{\left(|a(e^{-i\lambda})|^2 - |c(e^{-i\lambda})|^2\right)^2}{12^4} f_W(\lambda) \\ f_V(\lambda) &= \frac{1}{12^2}\left(2-\frac{|U(e^{-i\lambda})|^2}{12^2}\right) f_W(\lambda). \end{aligned}$$

3.5 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 λ_p for the filter, following the treatment of Bell and Monsell (1992); this should not be confused with the frequency argument λ. Now $\nu_j(B)$ has all of its many roots on the unit circle; indeed, letting $Z = B^{12}$, 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 $k = 1, \cdots, 12j$, 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 kth 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):

$$\begin{aligned}
\frac{Z^3 - 1}{Z - 1} &= g_9(B)\, g_{18}(B)\, g_{36}(B) \\
\frac{Z^5 - 1}{Z - 1} &= g_5(B)\, g_{10}(B)\, g_{15}(B)\, g_{20}(B)\, g_{30}(B)\, g_{60}(B) \\
\frac{Z^7 - 1}{Z - 1} &= g_7(B)\, g_{14}(B)\, g_{21}(B)\, g_{28}(B)\, g_{42}(B)\, g_{84}(B) \\
\frac{Z^9 - 1}{Z - 1} &= g_9(B)\, g_{18}(B)\, g_{27}(B)\, g_{36}(B)\, g_{54}(B)\, g_{108}(B).
\end{aligned}$$

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. So these seasonal filters λ_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 $j = 1, \cdots, 6$) 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.

Set $\Psi(B) = \lambda_p(B)$. We first consider the output-matching paradigm, noting that $\nu_j(e^{-i\lambda})$ is always bounded above by one, but may be negative (but never less than negative one). We must determine the signal differencing operator by examining $1 - \Psi(B)$. Now

$$\begin{aligned}
&1 - \lambda_p(B) = (1 - \nu_3(B)) + \nu_3(B)(1 - \nu_p(B)) \\
&1 - \nu_{2q+1}(B) = -(1 - Z)^2 \frac{Z^{-q}}{2q + 1} \left(Z^{2q-2} + 3Z^{2q-3} + 6Z^{2q-4} + \cdots + \binom{q+1}{2} Z^{q-1} + \cdots + 3Z + 1 \right),
\end{aligned}$$

from which it is seen that $(1 - Z)^2$ divides $1 - \lambda_p(B)$. Hence $\delta^S(B) = (1 - B^{12})$, which implies that the signal is not just seasonal, but can have an $I(1)$ trend as well (if we are only interested in situations where the seasonal filter is applied to nontrending data, then we can let $\delta^S(B) = U(B)$).

Then we can compute $\Phi(B)$ as follows:

$$\Phi(B) = \frac{1}{3} + \nu_3(B) \frac{\sum_{j=2}^{q+1} \binom{j}{2} Z^{q+1-j} + \sum_{j=2}^{q} \binom{j}{2} Z^{-q+j-1}}{2q+1}.$$

Hence the implied spectra are

$$\begin{aligned} f_U(\lambda) &= |\lambda_p(e^{-i\lambda})|^2 f_W(\lambda) \\ f_V(\lambda) &= \Phi(e^{-i\lambda}) \left(1 + \lambda_p(e^{-i\lambda})\right) f_W(\lambda). \end{aligned}$$

Now considering the WK paradigm, we still have $\delta^N(B) = 1$ and $\delta^S(B) = 1 - Z$. In this case the implied spectra are

$$\begin{aligned} f_U(\lambda) &= |\lambda_p(e^{-i\lambda})| f_W(\lambda) \\ f_V(\lambda) &= \frac{1 + \lambda_p(e^{-i\lambda})}{1 + |\lambda_p(e^{-i\lambda})|} \Phi(e^{-i\lambda}) f_W(\lambda). \end{aligned}$$

4 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 -1 and 1. Nevertheless it is possible to proceed, as the various component filters (i.e., the Henderson trend, the 2×12, and the seasonal moving averages) do satisfy the necessary properties, as demonstrated in subsections 3.2, 3.4, and 3.5. We begin with the basic definition of the X-11 filters in 4.1, and how they arise from a sequential or iterative approach to signal extraction. In 4.2 we derive the recasted filters, using this iterative approach. We then obtain implied spectral densities for the seasonal, trend, and irregular components, using the output-matching paradigm. Then in 4.3 we apply the method to a seasonal time series (with regression effects removed), and explicitly construct the matrix smoothers for its components.

4.1 X-11 as an Iterative Filtration

The general philosophy behind X-11 is that we first obtain a crude trend estimate, 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). Then 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 leftover, 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, Young, and Musgrave (1967), Shiskin (1978), and Ladiray and Quenneville (2001); also see Bell and Hillmer (1984) for a historical discussion. Following the notation of Bell and Monsell (1992), we let μ denote the 2×12 "crude trend" filter, λ_p the $3 \times p$ seasonal filter, and H_q will be the Henderson trend of order q. Then the seasonal filter ω_S is defined via

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

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

$$\omega_N = 1 - \omega_S.$$

The trend and irregular filters are then given by

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

respectively, by definition. The notation for the four components is Seasonal (S), Nonseasonal (N), Trend (T), and Irregular (I). So the estimate of S is given by $\omega_S Y$, and the estimate of N is given by $\omega_N Y$.

By the output-matching approach to recasting, we should define the components by multiplying f_Y by the squared gain of ω_N, ω_S, etc. As mentioned previously, this direct approach fails because the squared gains of these X-11 filters are not bounded between zero and one. Instead, we replace the frequency response function of each *constituent* filter, i.e., μ, λ_p, and H_q, by its squared gain function. A heuristic justification for this procedure is given in the Appendix. Hence the implied pseudo-spectra for S, N, T, and I are obtained by replacing each constituent filter in the frequency response functions of the composite filters ω_S, ω_N, $H_q\omega_N$, and $(1-H_q)\omega_N$ by their squared magnitudes. That is – suppressing the frequency argument of the functions for clarity of presentation –

$$\begin{aligned}
f_S &= (1-\mu^2)\lambda_{p_2}^2 \left[1 - H_q^2 \left(1 - (1-\mu^2)\lambda_{p_1}^2 \left(1-\mu^2\right)\right)\right] f_Y \\
f_N &= \left\{1 - (1-\mu^2)\lambda_{p_2}^2 \left[1 - H_q^2 \left(1 - (1-\mu^2)\lambda_{p_1}^2 \left(1-\mu^2\right)\right)\right]\right\} f_Y \\
f_T &= H_q^2 \left\{1 - (1-\mu^2)\lambda_{p_2}^2 \left[1 - H_q^2 \left(1 - (1-\mu^2)\lambda_{p_1}^2 \left(1-\mu^2\right)\right)\right]\right\} f_Y \\
f_I &= (1-H_q^2) \left\{1 - (1-\mu^2)\lambda_{p_2}^2 \left[1 - H_q^2 \left(1 - (1-\mu^2)\lambda_{p_1}^2 \left(1-\mu^2\right)\right)\right]\right\} f_Y.
\end{aligned} \tag{11}$$

Finally, we note that all of the constituent filters μ, λ_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. So 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 2.1.

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

$$f_S = (1 - \mu^2)\lambda_{p2}^2 \left[(1 - H_q^2) + H_q^2(1 - \mu^2)^2 \lambda_{p_1}^2\right] f_Y,$$

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

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

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

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

The first and third terms each admit two factors of $1 - e^{-i12\lambda}$ (see 3.5); the second term has two factors of $U(e^{-i\lambda})$ and four factors of $1 - e^{-i\lambda}$, which come from the μ^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

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

which is a bounded function. The trend component's pseudo-spectrum is just H_q^2 times that of the nonseasonal. 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 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. Letting

$$\Gamma_I^2(e^{-i\lambda}) = \frac{1 - H_q^2(e^{-i\lambda})}{|1 - e^{-i\lambda}|^{2d}},$$

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

$$\begin{aligned} f_{U^S}(\lambda) &= \Gamma_S^2(e^{-i\lambda}) f_W(\lambda) \\ f_{U^N}(\lambda) &= \Gamma_N^2(e^{-i\lambda}) f_W(\lambda) \\ f_{U^T}(\lambda) &= H_q^2(e^{-i\lambda})\Gamma_N^2(e^{-i\lambda}) f_W(\lambda) \\ f_I(\lambda) &= \Gamma_I^2(e^{-i\lambda})\Gamma_N^2(e^{-i\lambda}) f_W(\lambda). \end{aligned} \tag{12}$$

We note that these Γ 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 form of the implied spectra for seasonal, nonseasonal, trend, and irregular. The plotted functions can be multiplied by the data's true pseudo-spectrum to obtain the spectra for the components; these are the $\Psi^2(e^{-i\lambda})$ functions found implicitly in (11). Alternatively, one obtains these functions by dividing each Γ^2 in (12) by the appropriate differencing operator, but the first view-point is more informative. 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. Figure 1 considers $p_1 = 3$, $p_2 = 5$, and $q = 17$; figure 2 considers the case $p_1 = 3$, $p_2 = 9$, and $q = 9$.

4.2 X-11 Matrix Smoothers

We now describe the details of implementing X-11 matrix smoothers. The key is to determine the functions $\Gamma^2_S(B)$, $\Gamma^2_N(B)$, and $\Gamma^2_I(B)$, which are polynomials in B and F. Based on the calculations in 4.2, we have the following formulas:

$$\begin{aligned}
\Gamma^2_S(B) &= |1-B|^{6-2d}(1+\mu(B))\Phi^{\mu}(B)\lambda^2_{p_2}(B)\left[(1+H_q(B))\Phi^q(B) + H^2_q(B)(1+\mu(B))^2(\Phi^{\mu}(B))^2\lambda^2_{p_1}(B)\right] \\
\Gamma^2_N(B) &= |U(B)|^{2-2D}(|1-B|^2\left[(1+\lambda_{p_2}(B))\Phi^{p_2}(B) + (1+\lambda_{p_1}(B))\Phi^{p_1}(B)\lambda^2_{p_2}(B)H^2_q(B)\right] \\
&\quad + 24^{-2}|1+B|^2\lambda^2_{p_2}(B)\left[1 - H^2_q(B) + \left(3 - 3\mu^2(B) + \mu^4(B)\right)\lambda^2_{p_1}(B)H^2_q(B)\right]) \\
\Gamma^2_I(B) &= |1-B|^{4-2d}\Phi^q(B)(1+H_q(B)).
\end{aligned}$$

The following notations are used in the above formulas:

$$\begin{aligned}
\Phi^{\mu}(B) &= \left(F^5 + 4F^4 + 9F^3 + 16F^2 + 25F + 36 + 25B + 16B^2 + 9B^3 + 4B^4 + B^5\right)/24 \\
\Phi^q(B) &= \frac{1 - H^q(B)}{|1-B|^2} \\
\Phi^p(B) &= \frac{1}{3} + \nu_3(B)\frac{\sum_{j=2}^{(p+1)/2}\binom{j}{2}B^{12([p+1]/2-j)} + \sum_{j=2}^{(p-1)/2}\binom{j}{2}F^{12([p+1]/2-j)}}{p}.
\end{aligned}$$

The explicit formulas for $\Phi^q(B)$ for various q are given in 3.2. In order to compute the autocovariance matrices needed for the matrix smoothers (see Section 2.3 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. In implementation, it is easiest to create a large-dimension $\Sigma(f)$ matrix, recognizing that various submatrices will then correspond to $\Sigma(fe^{ih\cdot})$. In this way all the quadratic form estimates of the autocovariances are easily obtained, and the matrix smoothers (as well as the MSE matrices) are obtained via plugging into the formulas. As an illustration of the techniques, we examine U.S. Retail Sales of Shoe Stores data from the monthly Retail Trade Survey of the Census Bureau (with regression effects such as trading day removed), from 1984 to 1998, which will be referred to as the *Shoe* series. 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 U.S. Census Bureau; see Findley, Monsell, Bell, Otto, and Chen (1998)), whereas ACF and $PACF$ plots indicate $d = 1$, $D = 1$ instead. Of course, both $d = 1$ and $d = 2$ are accommodated by the X-11 filters, in the sense that the seasonal and irregular filters contain four nonseasonal differencings (see the discussion in Section 4.2).

So in this manner we can obtain estimates of $\Sigma(f_{U^S})$, $\Sigma(f_{U^N})$, $\Sigma(f_{U^T})$, and $\Sigma(f_I)$. More properly speaking, we have estimates of their autocovariance functions, which are consistent for the true values under the assumptions of our approach. It is then a simple matter to produce the corresponding matrix smoothers, and apply these to the data. Focusing on the WK matrix smoother, we present the estimated components with MSEs in Figures 3, 4, 5, and 6 for Nonseasonal, Trend, Seasonal, and Irregular respectively. There is little visible difference between taking $d = 1$ or $d = 2$, which is reassuring. There are some differences between the first and second component estimates; "first" refers to the choice considers $p_1 = 3$, $p_2 = 5$, and $q = 17$, and "second" refers to $p_1 = 3$, $p_2 = 9$, and $q = 9$. As expected, the first trend is smoother (since a longer Henderson is used) than the second. For the other components, it is difficult to detect any visible discrepancies. The MSEs are of course time-varying, and characteristically rise at the boundaries of the sample. These plots are provided for Nonseasonal and Trend, recognizing that the MSE matrix is the same for Nonseasonal and Seasonal (the Irregular MSEs are not shown, but are easily computed like the others). The procedure was implemented in R (R Development Core Team, 2005), and runs took about a second for the shoe series.

4.3 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; thus some careful thought is needed in order to make apt comparisons. We propose two: (1) compare the frequency response functions, and (2) compare estimated components on several series.

For the first comparison, we note that the implied spectra for the components can be used to construct a WK filter, simply by taking the ratio of signal to data. In a rough sense, this is the frequency response function of the X-11 matrix smoother – though a more exact frequency response function could be determined for each time point, and would depend on the estimated spectrum of the data. In order to have a single graphical comparison, we choose this WK approach to determining a frequency response function. For Nonseasonal, Seasonal, Trend, and Irregular we have

$$\begin{aligned}
\frac{f_N(\lambda)}{f_Y(\lambda)} &= \Gamma_N^2(e^{-i\lambda}) \left| U(e^{-i\lambda}) \right|^{2D} \\
\frac{f_S(\lambda)}{f_Y(\lambda)} &= \Gamma_S^2(e^{-i\lambda}) \left| 1 - e^{-i\lambda} \right|^{2d} \\
\frac{f_T(\lambda)}{f_Y(\lambda)} &= H_q^2(e^{-i\lambda}) \Gamma_N^2(e^{-i\lambda}) \left| U(e^{-i\lambda}) \right|^{2D} \\
\frac{f_I(\lambda)}{f_Y(\lambda)} &= \left(1 - H_q^2(e^{-i\lambda})\right) \Gamma_N^2(e^{-i\lambda}) \left| U(e^{-i\lambda}) \right|^{2D}.
\end{aligned}$$

From Figures 7 and 8, we see that the match for seasonal and nonseasonal components is remarkably close for values between zero and one. For the trend and irregular there are some salient discrepancies, and it is apparent that the X-11 matrix smoother does more smoothing. This effect could be lessened by using H_q instead of H_q^2 in the formulas defining the components; since we are mainly interested in the seasonal adjustments, we will not be concerned with the discrepancy for the trend and irregular.

For the second comparison we consider the following seven time series: m00110, m00100, *Shoe*, *Emp*, *Hours*, *Order*, and *Starts*. The first two time series are from the Foreign Trade Division of the U.S. Census Bureau; the first series is Imports of Meat Products, and the second series is Imports of Dairy Products and Eggs. Both of these series are for the time period from January 1989 to December 2003. The *Shoe* series is described above. The fourth series refers to Employed Males, aged 16 to 19, covering the period January 1976 through October 2006. The *Hours* series title is "Total private: Average Weekly Hours of Production Workers, NSA, Bureau of Labor Statistics"; the *Order* series title is "Manufacturing: Nondefense Capital Goods: New Orders: Millions of Dollars: NSA, Census Bureau"; and the *Starts* series title is "US Total New Privately Owned Housing Units Started; Thousands; NSA, Census Bureau."

All of these seven series were first adjusted for trading day and outlier effects using the X-12-ARIMA program (Findley, Monsell, Bell, Otto, and Chen, 1998). The adjusted series were then run through the X-11 matrix smoother, after a log transformation if required. The output seasonal adjustments were then compared to those generated by the X-11-ARIMA method, which uses forecast and backcast extension (see Findley et. al., 1998). We used two specifications: $p_1 = 3, p_2 = 5, q = 17$ and $p_1 = 3, p_2 = 9, q = 9$. For the X-11 matrix smoother we also specified $d = 1, D = 1$ for all the series (some series could arguably have used $d = 2$, but for purposes of comparison the differencing parameters were held to be the same). Seasonal adjustments of the same specification were compared by taking the average of the squared difference between the components; these values are reported in Table 1. The last column of the table compares the X-11-ARIMA seasonal adjustments for the two specifications, which gives a rough baseline against which the other mean squares can be contrasted. In general, discrepancies between the X-11 matrix smoother and X-11-ARIMA were a bit larger than the baseline, though generally of the same order of magnitude; only for the long series *Start* was the baseline mean square actually larger. The results for these seven series indicate that in practice the seasonal adjustments coming from the X-11 matrix smoother closely approximate those of X-11-ARIMA (generally, there was more discrepancies at the boundaries of the sample, which is to be expected).

5 Conclusions

In conclusion, 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, which is a time-consuming and error-prone process. Within the context of seasonal adjustment at the U.S. Census Bureau, there is some appeal to this non-MBA method of filter extension because of the large numbers of series that must be seasonally adjusted.

Acknowledgements The author thanks David Findley for stimulating discussions on this topic, and Bill Bell for careful reading of the manuscript.

Appendix

A.1 Asymptotic Properties of Estimated Spectra

Here we discuss the asymptotic properties of the estimates $\widehat{\gamma}(h)$ discussed in Section 2. If we are using nonparametric estimates, then $\widehat{\gamma}(h)$ is given by

$$\widehat{\gamma}(h) = \frac{1}{n-d} W' \Sigma(g) W = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) \widehat{f}_W(\lambda)\, d\lambda.$$

Here g is a bounded (possibly complex) function, and $\widehat{f}_W$ denotes the continuous-argument periodogram. The model-based estimate is given by

$$\widehat{\gamma}(h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda) f_W(\lambda; \widehat{\theta})\, d\lambda.$$

We have in mind that the function $f_W(\cdot; \theta)$ belongs to a model class $\mathcal{M}_\Theta$, where the parameters θ belong to a parameter space Θ with compact closure. We give two consistency results for these estimates. For the nonparametric case, some mild conditions on the data are required for the asymptotic theory; we follow the material in Taniguchi and Kakizawa (2000, Section 3.1.1). Condition (B), due to Brillinger (1981), states that the process is strictly stationary and condition (B1) of Taniguchi and Kakizawa (2000, p. 55) holds. Condition (HT), due to Hosoya and Taniguchi (1982), states that the process has a linear representation, and conditions (H1) through (H6) of Taniguchi and Kakizawa (2000, pp. 55 – 56) hold.

Theorem 1 *(Theorem 1 of McElroy (2006).) Suppose that* $\sum_h |h||\gamma_g(h)| < \infty$*. If the third and fourth cumulants of* W_t *are zero, then the mean and variance of* $Q(W) = W'\Sigma(g)W/(n-d)$ *are given by*

$$\begin{aligned} \mathbb{E}Q(W) &= \frac{1}{n-d}\, tr(\Sigma(g)\Sigma(f_W)) \\ VarQ(W) &= \frac{2}{(n-d)^2}\, tr\left((\Sigma(g)\Sigma(f_W))^2\right) \end{aligned} \tag{A.1}$$

where tr *denotes the trace of a matrix. Moreover, the mean and variance have the following limiting behavior as* $n \to \infty$*:*

$$\begin{aligned} \mathbb{E}Q(W) &\to \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\lambda)\, f_W(\lambda)\, d\lambda \\ nVarQ(W) &\to \frac{2}{2\pi}\int_{-\pi}^{\pi} g^2(\lambda)\, f_W^2(\lambda)\, d\lambda. \end{aligned}$$

Also, if the process $\{W_t\}$ *satisfies either condition (B) or (HT), then as* $n \to \infty$*:*

$$\sqrt{n}\,\frac{(Q(W) - \mathbb{E}Q(W))}{\sqrt{nVarQ(W)}} \stackrel{\mathcal{L}}{\Longrightarrow} \mathcal{N}(0,1).$$

Remark 1 The results for the mean $\mathbb{E}Q(W)$ hold even when the cumulant condition is not satisfied. So for non-linear non-Gaussian processes, we can still have consistency so long as (B) or (HT) are satisfied.

For the model-based case, we suppose that θ_0 is the true parameter, in the sense that the differenced data have spectral density $f_W(\cdot;\theta_0)$. Given that $\widehat{\theta}$ is a consistent estimate of θ_0, then consistency of the autocovariance estimates follows from a simple continuity condition on the spectra: we suppose that $f_W(\cdot;\theta)$ is uniformly continuous for all θ in a sufficiently small neighborhood of θ_0, where θ_0 is in the interior of Θ. This will be referred to as condition (C).

Theorem 2 *Suppose that* g *is uniformly bounded and (C) holds. Then*

$$\widehat{\gamma}(h) \stackrel{P}{\longrightarrow} \gamma(h).$$

Remark 2 Asymptotic theory for parameter estimates can be found in Taniguchi and Kakizawa (2000); various conditions on the data process sufficient to guarantee parameter consistency are discussed therein. We impose that θ_0 is in the interior of Θ so that unit roots (which create a pole in f_W) are avoided.

Proof of Theorem 2. We have

$$|\widehat{\gamma}(h) - \gamma(h)| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(\lambda)| \, |f_W(\lambda; \widehat{\theta}) - f_W(\lambda; \theta_0)| \, d\lambda.$$

We can bound g by its supremum, and $\sup_\lambda |f_W(\lambda; \cdot) - f_W(\lambda; \theta_0)|$ is a continuous function. So the result follows from the consistency of $\widehat{\theta}$. $\square$

A.2 Heuristic Justification for X-11 Approximation

We present an iteration scheme for X-11 that differs somewhat from the usual approach (Bell and Kramer, 1996), since generally the order of components is T, I, S, and N. For the methodology pursued in the next subsection, it is more convenient to adopt the order S, N, T, and I. We initialize the algorithm with $T^0 = Y$.

1. $S^1 = (1-\mu)\lambda_{p_0} \left(Y - T^0 \right)$: crude seasonal
2. $N^1 = Y - S^1$: crude nonseasonal
3. $T^1 = \mu N^1$: crude trend
4. $I^1 = (1-\mu) N^1$: crude irregular
5. $S^2 = (1-\mu)\lambda_{p_1} \left(Y - T^1 \right)$: refined seasonal
6. $N^2 = Y - S^2$: refined nonseasonal
7. $T^2 = H_q N^2$: refined trend
8. $I^2 = (1 - H_q) N^2$: refined irregular
9. $S^3 = (1-\mu)\lambda_{p_2} \left(Y - T^2 \right)$: final seasonal
10. $N^3 = Y - S^3$: final nonseasonal
11. $T^3 = H_q N^3$: final trend
12. $I^3 = (1 - H_q) N^3$: final irregular

So there are twelve steps, involving three iterations of a 4-step loop, although the first two steps are trivial (but are placed here for cohesion of presentation). Moreover, steps 4 and 8 are not calculated in practice, since they are not required for subsequent computations; they are included here in order to make the iterative pattern more clear. It is easy to see that steps 1 through 4 are essentially updated in steps 5 through 8, and again in steps 9 through 12. Also we see that ω_S and ω_N are the composite filters that result from the iteration, so that $S^3 = \omega_S Y$ and $N^3 = \omega_N Y$.

This scheme also allows us to see how the iterations could be continued, if desired. Letting θ_k and ν_k denote the trend and seasonal filters used in the kth iteration, we have

$$
\begin{aligned}
S^{k+1} &= (1-\mu)\nu_k \left(Y - T^k\right) \\
N^{k+1} &= Y - S^{k+1} \\
T^{k+1} &= \theta_k \, N^{k+1} \\
I^{k+1} &= (1-\theta_k) \, N^{k+1}.
\end{aligned}
$$

Note that μ appears in the first step; the fact that $\theta_0 = \mu$ is coincidental, as it were. Since μ involves an averaging over 13 consecutive months, $1-\mu$ has the effect of recentering the "preliminary seasonal" component P so that it sums to zero over any annual period. The pre-seasonal component P is only an artifice we use in our analysis, and does not appear in the X-11 procedure explicitly. It is defined by $P^{k+1} = \nu_k \left(Y - T^k\right)$, so that $S^{k+1} = (1-\mu)P^{k+1}$. The key principle for our approach, is to view the iterative steps above as defining true components rather than estimates. More precisely, for any fixed integer k we have the following relations by assumption:

$$
Y = T^k \oplus S^k \oplus I^k = P^k \oplus (Y - P^k),
$$

where $\oplus$ indicates that the summands are orthogonal, in the sense that their pseudo-spectra add up to the pseudo-spectrum of the data process. Now we also have $N^k = T^k \oplus I^k$ by definition. In addition, there are certain estimating equations that tells us how the components are related:

$$
\begin{aligned}
\widehat{P}^{k+1} &= \nu_k \left(Y - T^k\right) \\
\widehat{N}^{k+1} &= Y - (1-\mu)P^{k+1} \\
\widehat{T}^{k+1} &= \theta_k N^{k+1}.
\end{aligned}
\tag{A.2}
$$

We refer to (A.2) as a theoretical iteration scheme, since it relies at each step on a knowledge of the true T^k, P^k, and N^k, in a cyclical fashion. Note that this iterative scheme does not involve S or I; I is not needed, and it is more convenient to work with P than S. If we were to plug in estimates for T, P, and N on the right hand side of (A.2), we obtain the real iteration scheme described in Section 4.1. So the viewpoint is that the X-11 algorithm is obtained by plugging in previous estimates into a recursive relation of components.

Now letting f_X denote the pseudo-spectra of the process X, the third equation in (A.2) implies that $f_{\widehat{T}^{k+1}} = \theta_k^2 f_{N^{k+1}}$. Here we are suppressing the frequency argument in the spectra – θ_k^2 is the magnitude squared of the frequency response function of $\theta_k(B)$. This abuse of notation will be suffered so as not to clutter the subsequent derivations. Next, $\widehat{N}^{k+1} = (Y - P^{k+1}) \oplus \mu P^{k+1}$ so that

$$
f_{\widehat{N}^{k+1}} = f_{Y-P^{k+1}} + \mu^2 f_{P^{k+1}} = f_Y - (1-\mu^2) f_{P^{k+1}}.
$$

This follows from the assumption that $Y = P^{k+1} \oplus (Y - P^{k+1})$, so that $f_Y = f_{P^{k+1}} + f_{Y-P^{k+1}}$. Finally,

$$f_{\widehat{P}^{k+1}} = \nu_k^2 f_{Y-T^k} = \nu_k^2 \left(f_Y - f_{T^k}\right),$$

which uses $Y = T^k \oplus (Y - T^k)$. We emphasize that these relations hold for the idealized iteration scheme. Now applying the output-matching recasting method of 2.1 yields

$$\begin{aligned}
f_{P^{k+1}} &= f_{\widehat{P}^{k+1}} = \nu_k^2 \left(f_Y - f_{T^k}\right) \\
f_{N^{k+1}} &= f_{\widehat{N}^{k+1}} = f_Y - (1-\mu^2) f_{P^{k+1}} \\
f_{T^{k+1}} &= f_{\widehat{T}^{k+1}} = \theta_k^2 f_{N^{k+1}},
\end{aligned}$$

where the first equalities are really definitions. That is, the pseudo-spectra for the $(k+1)$th iterates of P, N, and T are defined to be the same as the pseudo-spectra of the corresponding component estimates in the theoretical iteration scheme. Now we can iteratively determine the implied pseudo-spectra for the components, with interest focusing on $k = 2$. We initialize with $f_{T^0} = f_Y$ as in the X-11 algorithm, so we recursively obtain

$$\begin{aligned}
f_{P^1} &= 0 \\
f_{N^1} &= f_Y \\
f_{T^1} &= \theta_0^2 f_Y \\
f_{P^2} &= \nu_1^2 \left(1-\theta_0^2\right) f_Y \\
f_{N^2} &= \left[1 - (1-\mu^2)\nu_1^2 \left(1-\theta_0^2\right)\right] f_Y \\
f_{T^2} &= \theta_1^2 \left[1 - (1-\mu^2)\nu_1^2 \left(1-\theta_0^2\right)\right] f_Y \\
f_{P^3} &= \nu_2^2 \left[1 - \theta_1^2 \left(1 - (1-\mu^2)\nu_1^2 \left(1-\theta_0^2\right)\right)\right] f_Y \\
f_{N^3} &= \left\{1 - (1-\mu^2)\nu_2^2 \left[1 - \theta_1^2 \left(1 - (1-\mu^2)\nu_1^2 \left(1-\theta_0^2\right)\right)\right]\right\} f_Y \\
f_{T^3} &= \theta_2^2 \left\{1 - (1-\mu^2)\nu_2^2 \left[1 - \theta_1^2 \left(1 - (1-\mu^2)\nu_1^2 \left(1-\theta_0^2\right)\right)\right]\right\} f_Y.
\end{aligned}$$

Moreover, since $Y = N^3 \oplus S^3$ and $N^3 = T^3 \oplus I^3$, we obtain

$$\begin{aligned}
f_{S^3} &= (1-\mu^2)\nu_2^2 \left[1 - \theta_1^2 \left(1 - (1-\mu^2)\nu_1^2 \left(1-\theta_0^2\right)\right)\right] f_Y \\
f_{I^3} &= (1-\theta_2^2) \left\{1 - (1-\mu^2)\nu_2^2 \left[1 - \theta_1^2 \left(1 - (1-\mu^2)\nu_1^2 \left(1-\theta_0^2\right)\right)\right]\right\} f_Y.
\end{aligned}$$

Now in the X-11 procedure, we have $\nu_1 = \lambda_{p_1}$, $\nu_2 = \lambda_{p_2}$, $\theta_0 = \mu$, $\theta_1 = \theta_2 = H_q$. Hence the implied pseudo-spectra for S^3, N^3, T^3, and I^3 are obtained by replacing each constituent filter in the frequency response functions of the composite filters ω_S, ω_N, $H_q\omega_N$, and $(1-H_q)\omega_N$ by their squared magnitudes. Making this substitution, and writing S for S^3, T for T^3, etc., we obtain (11).

References

[1] Ansley, C. and Wecker, W. (1984) On Dips in the Spectrum of a Seasonally Adjusted Time Series. Comment to "Issues involved with the seasonal adjustment of economic time series" by Bell and Hillmer. *Journal of Business and Economics Statistics* **2**, 323–324.

[2] Bell, W. (1984) Signal Extraction for Nonstationary Time Series. *The Annals of Statistics* **12**, 646 –664.

[3] Bell, W. (2005) Some Consideration of Seasonal Adjustment Variances. *Proceedings of the American Statistical Association, Section on Survey Research Methods.* [CD-ROM], Alexandria, VA: American Statistical Association, 2747–2758.

[4] Bell, W. and Hillmer, S. (1984) Issues involved with the Seasonal Adjustment of Economic Time Series. (With discussion) *Journal of Business and Economic Statistics* **2**, 291–320.

[5] Bell, W. and Kramer, M. (1996) Toward Variances for X-11 Seasonal Adjustments. SRD Research Report No. $RR - 96/07$, U.S. Census Bureau, available at http://www.census.gov/srd/www/byname.html.

[6] Bell, W. and Monsell, B. (1992) X-11 Symmetric Linear Filters and their Transfer Functions. SRD Research Report No. $RR - 92/15$, U.S. Census Bureau, available at http://www.census.gov/srd/www/byname.html.

[7] Brillinger, D. (1981) *Time Series Data Analysis and Theory.* San Francisco: Holden-Day.

[8] Burridge, P. and Wallis, K. (1985) Calculating the variance of seasonally adjusted series. *Journal of the American Statistical Association* **80**, 541–552.

[9] Chu, Y., Tiao, G., and Bell, W. (2007) A mean squared error criterion for comparing X-12-ARIMA and model-based seasonal adjustment filters. SRD Research Report No. $RRS2007/10$, U.S. Census Bureau, available at http://www.census.gov/srd/www/byname.html.

[10] Cleveland, W. and 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.

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

[12] Depoutot, R. and Planas, C. (1998) Comparing seasonal adjustment and trend extraction filters with application to a model-based selection of X11 linear filters, *Eurostat Working Paper* no9/1998/A/9.

[13] Durbin, J. and Koopman, S. (2001) *Time Series Analysis by State Space Methods.* Oxford University Press, Oxford.

[14] Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C. and Chen, B. C. (1998) New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program. *Journal of Business and Economic Statistics* **16**, 127–177 (with discussion).

[15] Hosoya, Y., and Taniguchi, M. (1982) A central limit theorem for stationary processes and the parameter estimation of linear processes. *Annals of Statistics* **10**: 132–153.

[16] Hungerford, T. (1974) *Algebra.* Springer-Verlag, New York.

[17] Kaiser, R. and Maravall, A. (2005) Combining Filter Design with Model-based Filtering: An Application to Business-cycle Estimation. *International Journal of Forecasting* **21**, 691–710.

[18] Ladiray, D. and Quenneville, B. (2001) *Seasonal Adjustment with the X-11 Method.* Springer-Verlag, New York.

[19] McElroy, T. (2005) Matrix Formulas for Nonstationary Signal Extraction. SRD Research Report No. $RRS2005/04$, U.S. Census Bureau.
http://www.census.gov/srd/www/byname.html.

[20] McElroy, T. (2006) Statistical Properties of Model-Based Signal Extraction Diagnostic Tests. SRD Research Report No. $RRS2006-06$, U.S. Census Bureau.
http://www.census.gov/srd/www/byname.html.

[21] McElroy, T. and Holan, S. (2005) A Nonparametric Test for Assessing Spectral Peaks. SRD Research Report No. $RRS2005-10$, U.S. Census Bureau, available at http://www.census.gov/srd/www/byname.html.

[22] McElroy, T. and Sutcliffe, A. (2006) An Iterated Parametric Approach to Nonstationary Signal Extraction. *Computational Statistics and Data Analysis, Special Issue on Signal Extraction* **50**, 2206–2231.

[23] Pfeffermann, D. (1994) A General Method for Estimating the Variances of X-11 Seasonally Adjusted Estimators. *Journal of Time Series Analysis* **15**, 85–116.

[24] Pierce, D. (1980) Data revision with moving average seasonal adjustment procedures. *Journal of Econometrics* **14**, 95–114.

[25] Pollock, D. (2000) Trend estimation and de-trending via rational square-wave filters. *Journal of Econometrics* **99**, 317–334.

[26] Pollock, D. (2002) Circulant Matrices and Time-Series Analysis. *The International Journal of Mathematical Education in Science and Technology* **33**, 213–230.

[27] President's Committee to Appraise Employment and Unemployment Statistics (1962). Measuring Employment and Unemployment, U.S. Government Printing Office, Washington, D.C.

[28] R Development Core Team (2005) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

[29] Shiskin, J., Young, A., and Musgrave, J. (1967) The X-11 Variant of the Census Method II Seasonal Adjustment Program. Technical Paper No. 15, U.S. Department of Commerce, Bureau of Economic Analysis.

[30] Shiskin, J. (1978) "Seasonal adjustment of sensitive indicators," in *Seasonal Analysis of Economic Time Series*, ed. A. Zellner, Washington, D.C.: U.S. Department of Commerce, Bureau of the Census, pp. 97–103.

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

[32] Wecker, W. (1979) A New Approach to Seasonal Adjustment. *Proceedings of the American Statistical Association, Business and Economic Statistics Section.* 322–323.

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

Table 1. Comparisons of Seasonal Adjustments

	MS Comparisons		
Series	Spec 1	Spec 2	Baseline
m00100	.000976	.001207	.000208
m00110	.000647	.000608	.000375
emp	.0000637	.0000602	.0000219
hours	.00434	.00391	.00047
order	.000114	.000118	.000050
start	3.296	1.128	5.110
shoe	.0000625	.0000779	.0000574

Table 1: Empirical mean square differences between the seasonal adjustments. Spec 1 compares the X-11 matrix smoother with the X-11-ARIMA method, using $p_1 = 3$, $p_2 = 5$, and $q = 17$. Spec 2 compares the X-11 matrix smoother with the X-11-ARIMA method, using $p_1 = 3$, $p_2 = 9$, and $q = 9$. Baseline compares the X-11-ARIMA method to itself with the two different specifications.

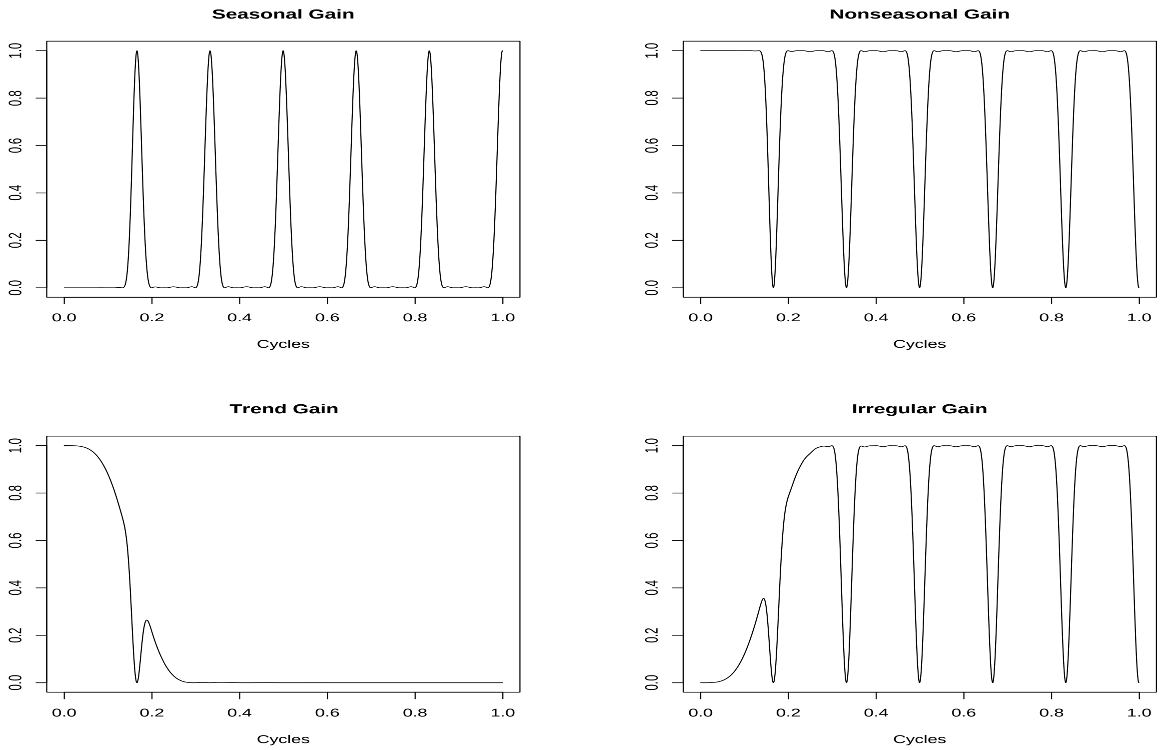

Figure 1: Squared gain functions defining the seasonal, nonseasonal, trend, and irregular components for the X-11 method, with $p_1 = 3$, $p_2 = 5$, and $q = 17$.

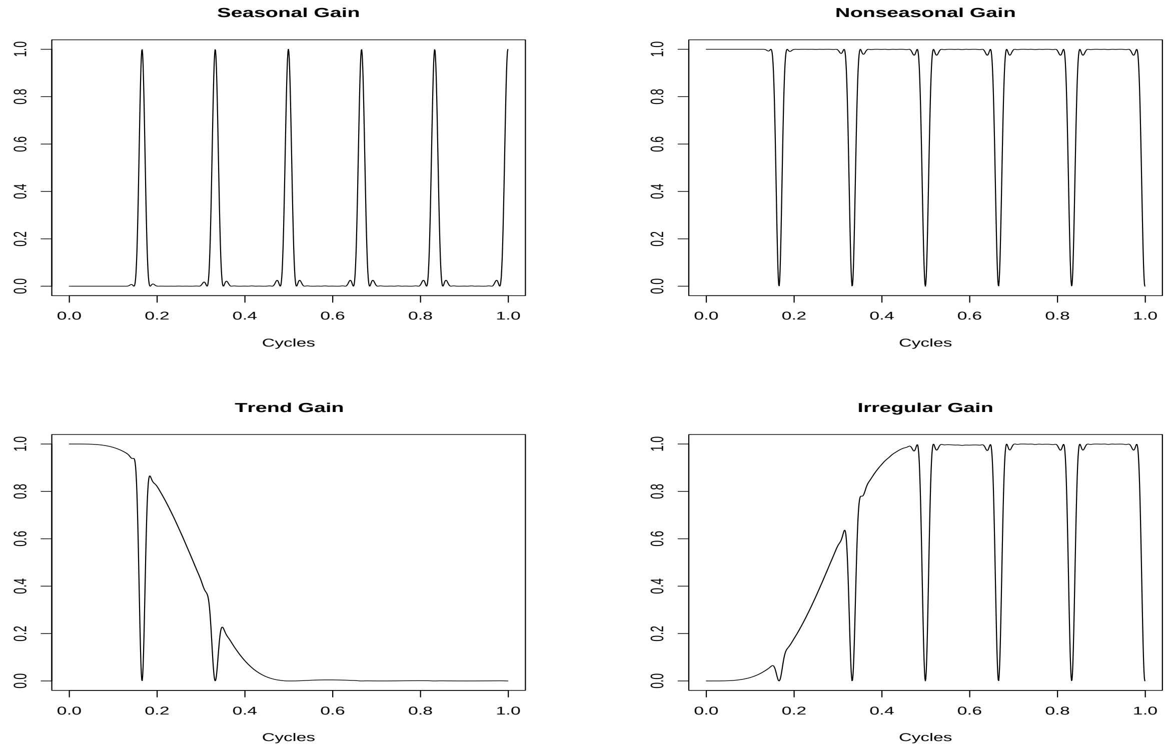

Figure 2: Squared gain functions defining the seasonal, nonseasonal, trend, and irregular components for the X-11 method, with $p_1 = 3$, $p_2 = 9$, and $q = 9$.

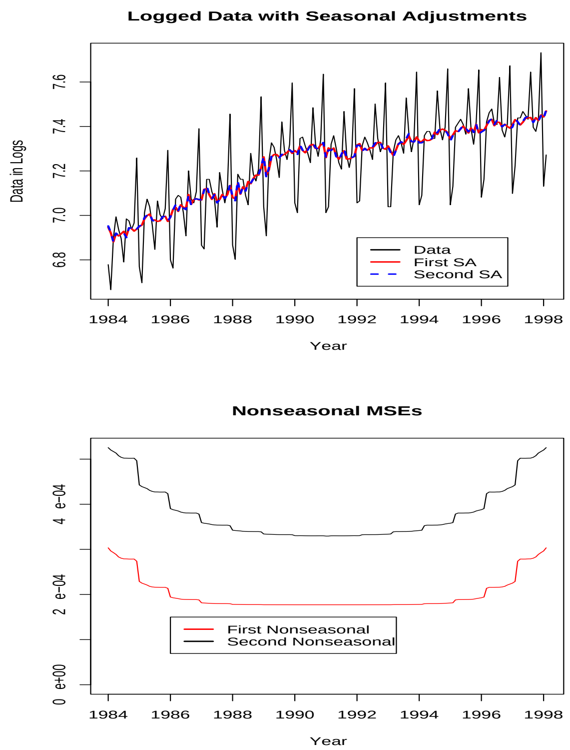

Figure 3: Estimated nonseasonal (SA) components, plotted with the logged data. Left panels are with $d = 1$, right panels with $d = 2$. Upper panels are for component estimates, bottom panels for corresponding MSEs. First SA refers to X-11 method with $p_1 = 3$, $p_2 = 5$, and $q = 17$; second SA refers to X-11 method with $p_1 = 3$, $p_2 = 9$, and $q = 9$.

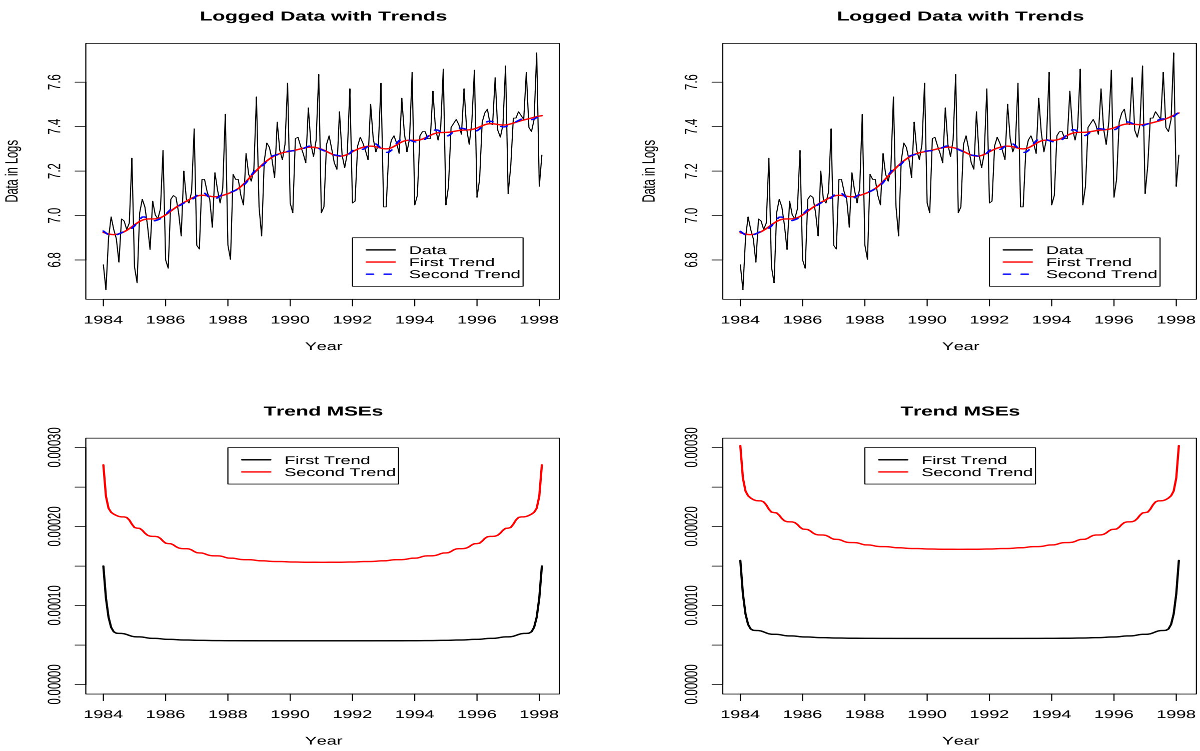

Figure 4: Estimated trend components, plotted with the logged data. Left panels are with $d = 1$, right panels with $d = 2$. Upper panels are for component estimates, bottom panels for corresponding MSEs. First trend refers to X-11 method with $p_1 = 3$, $p_2 = 5$, and $q = 17$; second trend refers to X-11 method with $p_1 = 3$, $p_2 = 9$, and $q = 9$.

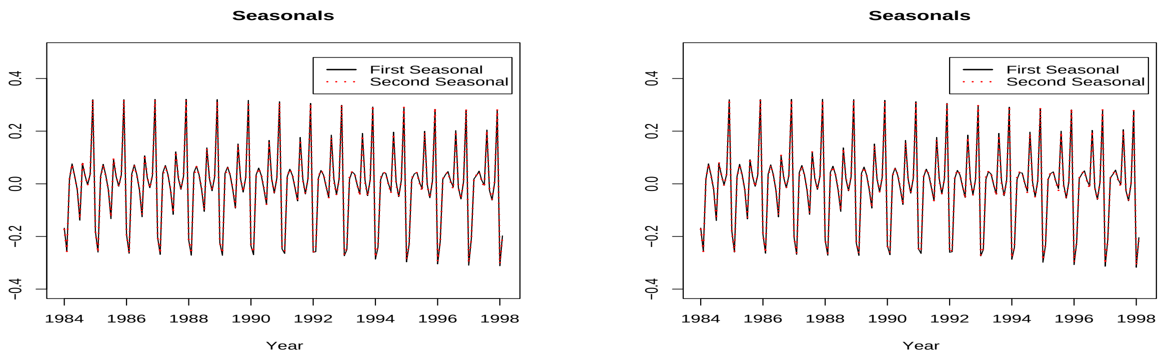

Figure 5: Estimated seasonal components. Left panels are with $d = 1$, right panels with $d = 2$. (MSEs are the same as for the SA.) First seasonal refers to X-11 method with $p_1 = 3$, $p_2 = 5$, and $q = 17$; second seasonal refers to X-11 method with $p_1 = 3$, $p_2 = 9$, and $q = 9$.

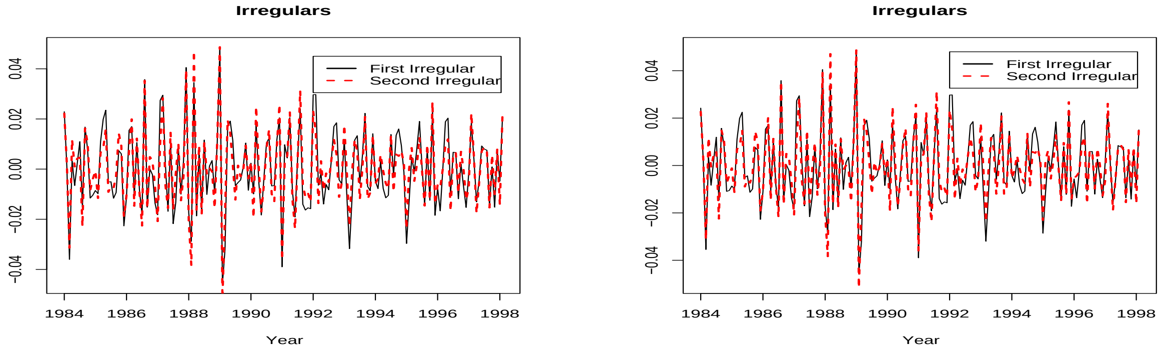

Figure 6: Estimated irregular components. Left panels are with $d = 1$, right panels with $d = 2$. First irregular refers to X-11 method with $p_1 = 3$, $p_2 = 5$, and $q = 17$; second irregular refers to X-11 method with $p_1 = 3$, $p_2 = 9$, and $q = 9$.

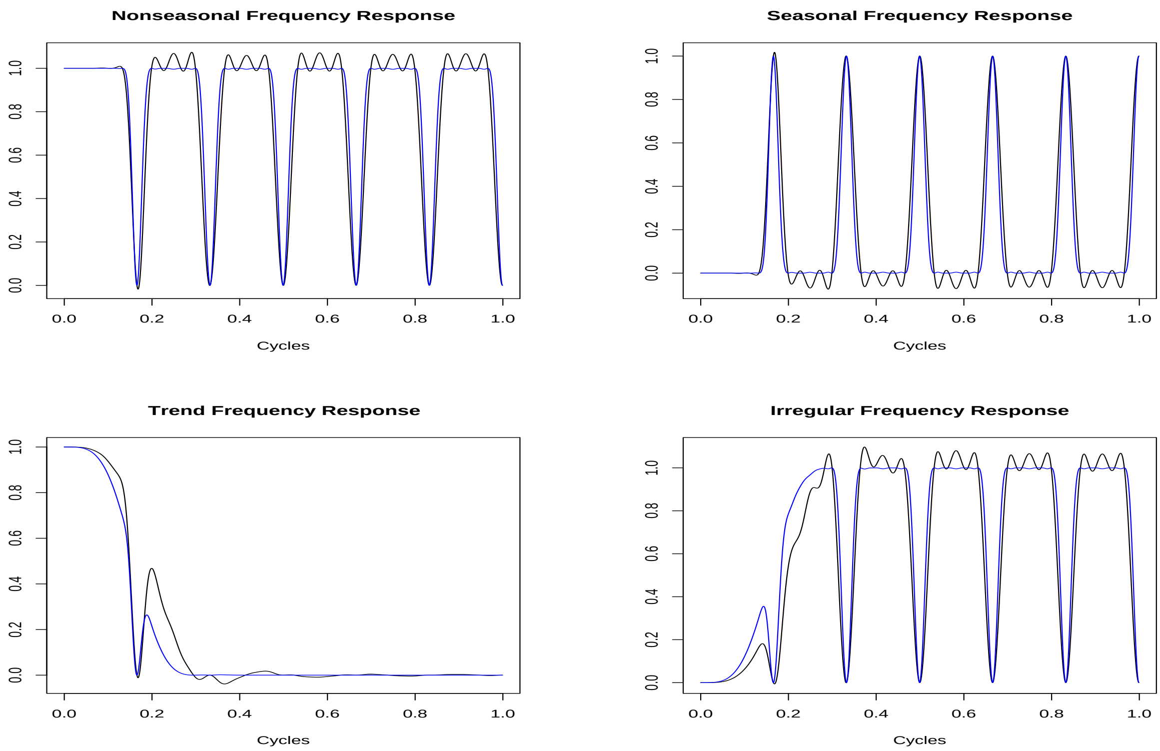

Figure 7: Frequency response functions for nonseasonal, seasonal, trend, and irregular components. In black is f.r.f for X-11 method, and in blue is f.r.f for WK approximation of X-11 matrix smoother (see section 4.4). Used $p_1 = 3$, $p_2 = 5$, and $q = 17$.

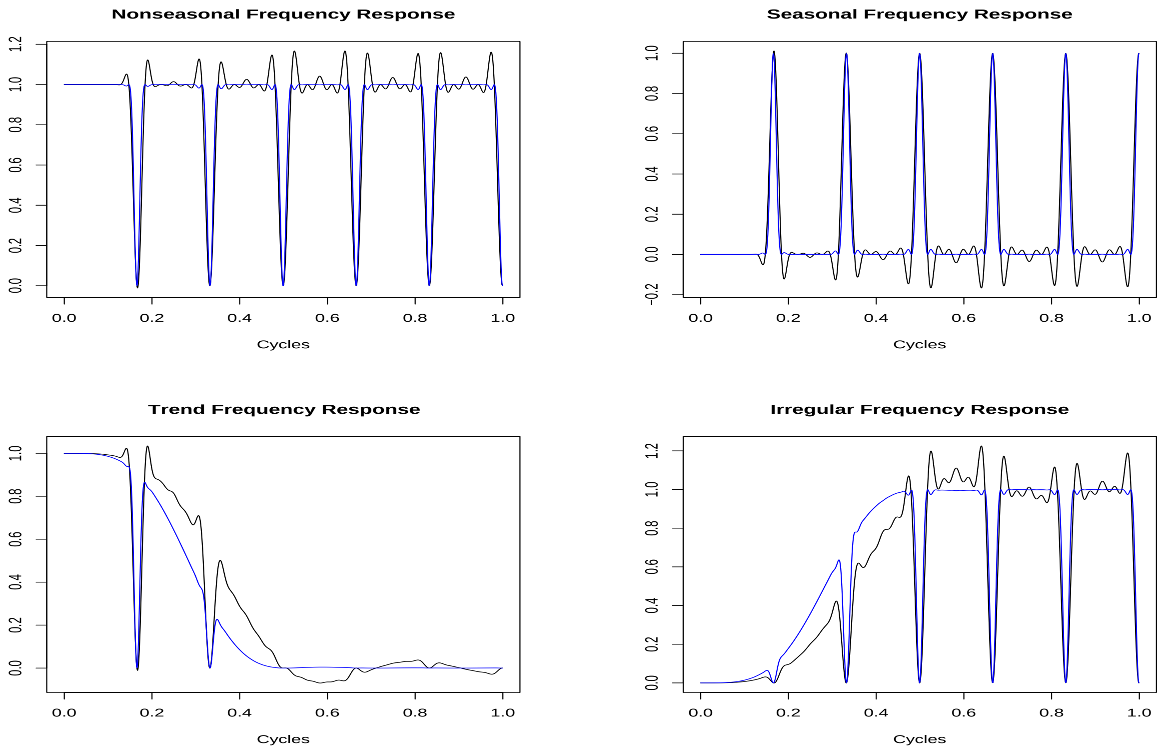

Figure 8: Frequency response functions for nonseasonal, seasonal, trend, and irregular components. In black is f.r.f for X-11 method, and in blue is f.r.f for WK approximation of X-11 matrix smoother (see section 4.4). Used $p_1 = 3$, $p_2 = 9$, and $q = 9$.