Philippine Statistician, 2004
Vol. 53, Nos. 1-4, pp. 21-34

Seasonal Adjustment by Signal Extraction Through Wavelets[1]

Daniel C. Bonzo[2]

ABSTRACT

Seasonal adjustment is undertaken in order to excise out the short-term effects of seasonal factors on a series. This results in an adjusted series that is more amenable to policy and statistical analysis. This procedure is typically carried out by the use of programs such as the US Bureau of the Census' X-11 or one of its variants such as Statistics Canada's X-11 ARIMA. These methods, however, are ad hoc in nature in the sense that they fail to exhibit optimal properties. In this technical report we are proposing an alternative procedure based on filters defined by wavelet coefficients. Wavelets are used to extract an estimate of the ARIMA parameters through an approximate maximum likelihood procedure. The estimate is then used to design a filter that will extract the seasonal component out of the original series. This filter is optimal in the sense defined by Whittle (1963).

1. INTRODUCTION TO SEASONAL ADJUSTMENT

Why seasonally adjust a series?

The economic impetus for development, in general, and concerns about sustainable growth require constant monitoring of socio-economic data series. Through careful assessment and analysis of trends in such series, policy makers and advisers would be able to formulate and adopt proper interventions that would secure long-term growth prospects for a country's economy. However, such an analysis requires the isolation of seasonal fluctuations in order to rid the series of the short-term effects of seasonal factors. Such a procedure is called deseasonalization or seasonal adjustment.

General Procedure

Seasonal adjustment of a time series consists of a series of procedure that seeks the removal of series oscillations due to seasonal factors. Removal of such variations is important in policy formulation since it seeks to eliminate short-term changes that are usually traced to factors such as calendar effects, institutional factors, weather or climate, among others. It is logical to think that the removal of such fluctuations will bring about a smoothed series that is more amenable to statistical and policy analyses.

Seasonal adjustment was first introduced by the US Bureau of Census in 1957 under the direction of Julius Shiskin. The general adjustment process requires the expression of any time series $\{X_t\}$ into one of two model forms: multiplicative or additive. The additive model is given by

$$X_t = T_t + C_t + S_t + I_t \qquad (1)$$

[1] This paper was supported by the Statistical Research and Training Center under its Research Fellowship Program for the year 2000.
[2] Former Associate Professor at the School of Statistics of the University of the Philippines and is currently a consultant at the Biostatistics and Statistical Programming, Purdue Pharma LP in Standford, Conneticut, USA. Email Address: Daniel.Bonzo@purduepharma.com

where

T_t = trend term;
C_t = cycle term;
S_t = seasonal term; and
I_t = irregular term.

On the other hand, the multiplicative model is given by

$$X_t = T_t C_t S_t I_t \ . \qquad (2)$$

The centerpiece of the process is the estimation of the seasonal component S_t. Once this has been achieved a seasonally adjusted series $\{X_t^a\}$ can be produced using the formulas,

$$X_t^a = X_t - S_t \text{ (for additive models)}$$

and

$$X_t^a = \frac{X_t}{S_t} \text{ (for multiplicative models).}$$

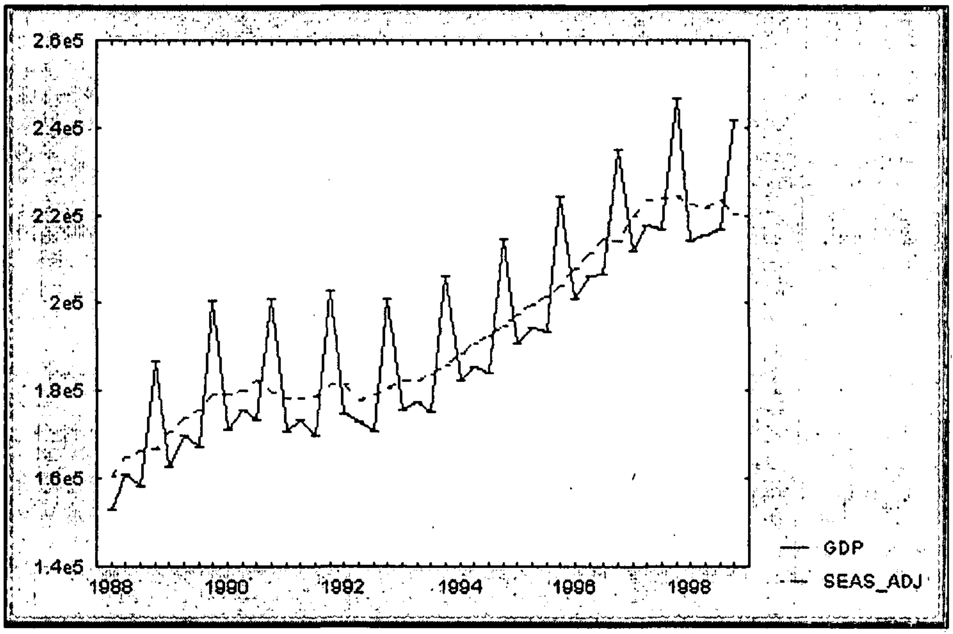
The techniques for estimating S_t, as contained in Technical Paper No. 15, provide the use of moving average weights or filters to smoothen the series. The general approach requires the detrending of the series first to center it about 0 (additive) or 1 (multiplicative) and then averaged by month (all Januarys, all Februarys, and so on) to get the seasonal deviation from the center. The process is repeated with moving averages (Henderson's) of different widths and with down-weighting of selected observations based on the size of irregular estimates. A rudimentary explanation on the major steps involved in the decomposition process is given in Bersales and Sarte (2000).

It should be noted that the use of X-11 on a series will result to a reduced deseasonalized series. The reduction is equivalent to approximately one year of data concentrated at the beginning and end of the adjusted series. Thus, the ensuing variants, specifically X-11 ARIMA, sought to address this problem brought about by the asymmetry in the filtering process by incorporating one year of forecasts at the end of the original series and one year of backcasts at the beginning of the original series. The X-11 method is then applied on the augmented series. Such an augmentation process seeks to improve adjustments near the end of the series and obtain balanced seasonal projections and full seasonal adjusted series.

The Philippine Experience

The X-11 method and its most famous variant, the X-11 ARIMA, is widely used by government agencies and private businesses in many countries, including the Philippines. The use of the procedure to adjust Philippine time series data was facilitated by a grant from the Canadian government in 1992. The grant was funneled through the National Statistical Coordination Board (NSCB) and mandated the implementation of seasonal adjustment on important Philippine time series, particularly those used of policy analysis.

Table 1. Original and Seasonally Adjusted GDP, 1988 – 1998



The seasonal adjustment project called for the creation of a technical committee consisting of agencies in the Philippine Statistical System (PSS) as member agencies. Under the guidance of the technical committee, the Technical Working Group on Seasonal Adjustment (TC/TWG-SAPTS) as created and was responsible for the implementation of the project's mandate. The seasonal adjustment procedure used was Statistics Canada's X-11 ARIMA. To date the practice of seasonal adjustment for policy analysis has become a common practice in government agencies responsible for the generation of economic data.

In this technical paper, we discuss some of the raging issues involved in seasonal adjustment. We tackle the issue of optimality by using the signal extraction approach. Here we devise a filtering process that will extract the seasonal component out in an optimal sense. The filter is designed keeping in mind a multiplicative ARIMA process governing the series of interest. Estimation of the ARIMA parameters is done through the application of an approximate wavelet maximum likelihood estimation procedure.

The paper is constructed as follows. We discuss the fundamental issues involved in seasonal adjustment in section 2. We next discuss the signal extraction procedure in section 3, following largely the discussion of Burman (1980). Section 4 discusses rudimentary concepts in wavelets and how wavelet coefficients are used in characterizing time series. Finally, the suggested approach to seasonal adjustment is discussed in section 5.

2. ISSUES IN SEASONAL ADJUSTMENT

Usual Criticisms

The complication involved in seasonal adjustment procedures based on X-11 can be traced o the decomposition method given in (1) and (2). Mathematically, there exists an infinite number of ways of implementing such decompositions and the use of moving averages/linear filters to effect such decompositions is just one small sub-class of such possibilities. In the presence of denumerable choices, it is best to apply statistical ideas in order to identify an optimal decomposition process.

The presence of so many competing methods of decomposition, no matter how appealing they maybe, is naturally a source of large-scale confusion. Each has its unique selling point and technical consideration. As Burman (1965) emphasized, none of these methods can be shown to have optimal properties and are therefore considered ad hoc in nature. This predicament arises mainly from the contention that the time domain is incapable of providing precise definition of what a decomposition method is.

On the use of X-11 and its variants alone one can already identify a number of issues to grope with. These include:

1. The method of decomposition(addition or multiplicative) to be used;
2. Handling of extreme values and prior adjustments;
3. Treatment of Easter and trading day effects;
4. Treatment of missing observations (the variants don't allow data gaps);
5. Raking/benchmarking/streamlining;
6. Determination of the lowest level of disaggregation;
7. Concurrent adjustment; and
8. Incorporating structural change in the original series.

Add to these the guidelines to be used in assessing the adequacy of the seasonal adjustment applied. Even X-11 ARIMA, which was developed to produce more 'accurate' estimates of seasonally adjusted series when seasonality changes rapidly in a stochastic manner can not be spared from these technicalities. For instance it has been noted (see Bersales and Sarte, 2000) that X-11 ARIMA can not cope with abrupt changes in the structure of the time series.

Optimality

Notable attempts in the literature have been made to add the optimality dimension in the decomposition process. Cleveland and Tiao (1976) for instance provided an additive model with stochastic trend, seasonal and noise components to approximate the linear filter version of the X-11 decomposition process. Here, the parameters of the model are extracted using least squares, and hence the incorporation of the optimality dimension.

A more rigorous approach, which will be followed in his paper, was given by Burman (1980). In this approach, called signal extraction, the decomposition process is treated in the spectral/frequency domain framework. In this domain, the seasonal component of a time series is usually exhibited as spectrum peaks at the seasonal frequency and its multiples, while the trend is represented by a broad peak at low frequencies. Thus, seasonal adjustment under this framework can be defined as the operation of removing peaks while leaving the rest of the spectrum undisturbed. This extraction procedure can be designed optimally through the application of an appropriate filter. Filter design can be carried out using nonparametric methods to avoid model misspecification. Hence, the use of wavelets in signal extraction was conceived in this paper.

3. SIGNAL EXTRACTION FORMULATION

We begin the discussion of the extraction formulation by discussing filters. Here we consider a signal X_t in $L^2(\Re)$ as the input of the filtering process.

What is a filter?

Consider a linear process $\{Y_t\}$ obtained from the signal $\{X_t\}$ through the application of a linear filter $C = \{c_{t,k},\ t, k = 0, \pm 1, \pm 2, \ldots\}$, i.e.,

$$Y_t = \sum_{k=-\infty}^{\infty} c_{t,k} X_k\,, \quad t = 0, \pm 1, \pm 2, \ldots$$

This linear system has the capability of attenuating some frequencies, e.g. noise, of X_t while passing others. In general, we are interested in time invariant filters ($c_{t,t-k} = h_k$), so that

$$\begin{aligned} Y_t &= \sum_{k=-\infty}^{\infty} h_k X_{t-k} \\ &= H * X. \end{aligned}$$

In addition, we require that h_j, $j < 0$ so that Y_t is expressed in terms of X_s, $s \leq t$.

Given a signal X_t in $L^2(\Re)$ and a causal, time-invariant linear filter (CTLF) H, the filtered output Y_t will have a spectral density given by

$$\begin{aligned} f_Y(w) &= |H(e^{-iw})|^2 f_X(w) \\ &= H(e^{-iw})H(e^{iw}) f_X(w) \end{aligned} \tag{3}$$

where

$f_X(w) = \dfrac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_X(j) e^{-iwj}$ (spectral density of X_t) ;

$\gamma_X(j)$ = autocovariance function of X_t; and

$H(e^{-iw}) = \sum_{j=0}^{\infty} h_j e^{-ijw}$.

The function $H(e^{-iw})$ is called the transfer function of the filter and the squared modulus $|H(e^{-iw})|^2$ is referred to as the power function of the filter.

In Fourier transform representation,

$$Y(w) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} Y_t e^{-iwt}\,, \qquad -\pi \leq w \leq \pi$$

$$= \frac{1}{2\pi}\sum_{t=-\infty}^{\infty}\sum_{j=0}^{\infty} h_j X_{t-j} e^{-iwt}$$
$$= \mathrm{H(w)X(w)}\,.$$

We call H(w) a low-pass filter if H(w) = 0 for $\pi/2 \le w \le \pi$; a high-pass filter if it satisfies H(w) = 0 for $-\pi \le w \le \pi/2$; and a band-pass filter if H(w) = 0 for $|w| \le \pi/2$.

In terms of the backshift operator B,

$$H(B) = \sum_{j=0}^{\infty} h_j B^j$$

becomes the generator of Y_t or the transfer function polynomial.

The use of filters will be taken in the context of ARIMA models. We assume that the signal takes the multiplicative ARIMA form

$$\Phi(B^S)\phi(B)(1-B)^D(1-B)^d X_t = \Theta(B^S)\theta(B)a_t \tag{4}$$

where S is the seasonality of the series; a_t is white noise; D and d are the seasonal and non-seasonal differencing parameters, respectively; $\Phi(B^S)$ and $\phi(B)$ are the seasonal and non-seasonal autoregressive polynomials, respectively; and $\Theta(B^S)$ and $\theta(B)$ are the seasonal and non-seasonal moving average polynomials, respectively. (4) can be expressed alternatively as

$$X_t = H(B)a_t = \frac{\theta(B)}{\varphi_M(B)\varphi_S(B)}a_t \tag{5}$$

where $\psi_M(B)$ and $\psi_S(B)$ have no common factors and represent the trend-cycle and seasonal components, respectively. Following (3), the spectrum of X_t is given by

$$f_Y(w) = |H(e^{-iw})|^2 \frac{\sigma_a^2}{2\pi} \tag{6}$$
$$= \left|\frac{\theta(e^{iw})}{\psi_M(e^{iw})\psi_S(e^{iw})}\right| \frac{\sigma_a^2}{2\pi}.$$

Decomposition Using Filters

Following Burman (1980), we define the trend-cycle and seasonal components as those responsible for the spectral peaks at the origin and at the seasonal frequencies, respectively, and the irregular component covers the transient characteristics, i.e., white noise or a low order MA process.

Now, let

$$X_t = M_t + S_t + I_t \tag{7}$$

where

$M_t = H_M(B)b_t$ (trend-cycle component);

$S_t = H_S(B)c_t$ (seasonal component);

$I_t = H_I(B)d_t$ (irregular component)

with H_M, H_S and H_I being the transfer functions of the components of X_t and b_t, c_t, and d_t are independent white noise processes. With the spectrum of the components defined similarly as in (6), X_t's spectrum can be expressed as

$$f_X(w) = f_M(w) + f_S(w) + f_I(w). \quad (8)$$

Assuming for convenience that $\sigma_b^2 = \sigma_c^2 = \sigma_d^2$, (8) can be expressed as

$$g_X(w) = g_M(w) + g_S(w) + g_I(w) \quad (9)$$

where g_M, g_S and g_I are rational functions of x = cos w. This shows the decomposition of X_t via the spectrum, which obviously is not unique.

Minimum Signal Extraction Method

For implementation purposes we apply the result of Whittle (1963), who showed that the best minimum square error (MSE) linear estimator of a component M_t given X_t is

$$\hat{M}_t = \frac{H_M(B)H_M(F)}{H(B)H(F)} X_t$$

where F is the forward operator.

Denote by p and q the degree of the polynomial in the denominator and numerator in (5), respectively. Let ε_M and ε_S be the minimum of f_M and f_S, respectively and replace f_S by $f_S{}^* = f_S - \varepsilon_S$, f_M by $f_M{}^* = f_M - \varepsilon_M$, and f_I by $f_I{}^* = f_I + \varepsilon_S + \varepsilon_M$. Let $f_S{}^* = g_S{}^*(x) = G_S(B,F)$ where $x = \frac{1}{2}(e^{iw} + e^{-iw})$ is replaced by $\frac{1}{2}(B+F)$. Then the minimum signal extraction filter for the seasonal component is

$$\begin{aligned}\delta_S(B,F) &= \frac{g_S{}^*(x)}{g_X(x)} = \frac{G_S(B,F)}{\psi_S(B)\psi_S(F)} \frac{\psi_S(B)\psi_S(F)\psi_M(B)\psi_M(F)}{\theta(B)\theta(F)} \\ &= \frac{G_S(B,F)\psi_M(B)\psi_M(F)}{\theta(B)\theta(F)} \\ &= \frac{C_S(B,F)}{\theta(B)\theta(F)} \quad \text{(say)} \end{aligned} \quad (10)$$

where $C_S(B,F)$ is a symmetric polynomial in B and F of degree p. Similarly, the minimum trend removal filter is given by

$$\begin{aligned}\delta_M(B,F) &= \frac{g_M{}^*(x)}{g_X(x)} \\ &= \frac{C_M(B,F)}{\theta(B)\theta(F)} \quad \text{(say)}\end{aligned} \quad (11)$$

where $C_M(B,F)$ is also symmetric and of degree p.

In order to avoid convergence problems, (10) and (11) are expressed in the following rational form:

$$\frac{C(B,F)}{\theta(B)\theta(F)} = \frac{K(B)}{\theta(B)} + \frac{K(F)}{\theta(F)} \qquad (12)$$

where K(.) is a polynomial of degree r = max(p,q). K is then used to filter X_t as

$$X_{1t} = \frac{K(F)}{\theta(F)} X_t = H_1(F) X_t .$$

Define $\Phi^*(B) = \frac{\phi(B)\Phi(B^S)}{\Theta(B^S)}$. Forecast X_t as

$$\Phi^*(B) X_t = \theta(B) a_t , \quad t = N+1, N+2, \ldots, N+q+r$$

where N is the length of the series.

We then construct an intermediate series

$$A_t = K(F) X_t , \quad 1 \le t \le N+q .$$

Now,

$$\begin{aligned} \Phi^*(B) X_{1t} &= \Phi^*(B) H_1(F) X_t \\ &= H_1(F) \Phi^*(B) X_t \\ &= 0, \qquad \text{for } t \ge N+q+1 . \end{aligned}$$

From the above formulation we can define a system of p + q equations in X_{1t} (t = N+q-p+1, . . . , N+2q) given by

$$\begin{aligned} \theta(F) X_{1t} &= A_t , t = N+q-p+1, \ldots, N+q \qquad (13) \\ \Phi^*(B) X_{1t} &= 0, \qquad t = N+q+1, \ldots, N+2q . \end{aligned}$$

The remaining X_{1t} can be found recursively from the relation in the first part of (13), working backwards to t = 1. The mirror image of these steps is applied to the backcast of X_t to give X_{2t}. Finally, the filtered component is the sum of X_{1t} and X_{2t}. The whole process is applied with K_M(.) for the trend and K_S(.) for the seasonal component, respectively. This entire methodology is known as the Minimum Signal Extraction Method, otherwise known as MSX.

4. WAVELETS

Definition and Some Concepts

A wavelet is defined as nay function W(t) in $L^2(\Re)$ satisfying the admissibility condition

$$\int_{-\infty}^{\infty} \frac{dw}{w} |\hat{W}(w)|^2 < \infty \qquad (14)$$

where $\hat{W}$ denotes the Fourier transform of W. It is usually expressed in the parametric form $W_{a,b}(t) = \dfrac{W\left(\frac{t-b}{a}\right)}{\sqrt{|a|}}$ where a, b are in $\Re$ ($a \neq 0$) and are known as the dilation and translation parameters, respectively. The admissibility condition in (14) is required so that the wavelet transform becomes invertible. It turns out that this is ensured whenever $\hat{W}$(w) has sufficient fast decay, i.e., $\hat{W}(w) \to 0$ as $|w| \to \infty$ and

$$\hat{W}(0) = \int_{-\infty}^{\infty} W(t)dt = 0 .$$

Wavelet Transforms

The continuous wavelet transform (CWT) of X_t in $L^2(\Re)$ at the time-scale location parameter (b,a) is defined by the inner product

$$< X, W_{a,b} >= \int X_t W_{a,b}(t)dt .$$

The above transform satisfies the property

$$\int |< X, W_{a,b} >|^2 = \int | X_t |^2 dt .$$

Hence, CWT's completely characterize X_t in the L^2 sense. Moreover, X_t maybe reconstructed by the inverse transform given by

$$X_t = C_W^{-1} \int\int a^{-2} < X, W_{a,b} > W_{a,b} dadb$$

where $C_W^{-1} = 2\pi \int | W(w) |^2 | w |^{-1} dw < \infty$. The admissibility condition $\int_{-\infty}^{\infty} W(t)dt = 0$ is implied by $C_W^{-1} < \infty$ if W(t) has sufficient decay.

The discrete wavelet transform (DWT) of X(t) in $L^2(\Re)$ is the doubly-indexed sequence $\{d_{j,k} : j, k \in Z\}$ such that

$$d_{j,k} = 2^{j/2} < X, W(2^j(t - k/2^j)) > .$$

Note that $d_{j,k}$ is just the value of the CWT of X_t at the time-scale location $(k/2^j, 1/2^j)$ or at the time-frequency location $(k/2^j, c/2^j)$ where $c > 0$ is a constant that depends on the choice of W(t).

If the time interval is normalized to the unit interval, the support of the wavelet becomes $[(n-1)2^{-(m-1)}, n2^{-(m-1)}]$ so that the wavelet covers the entire time series. Hence, for a scaling parameter m, the translation parameter has value n = 1, 2, . . . , 2^{m-1}. Thus, for a time series of length $N = 2^l$, the DWT (wavelet coefficients) consists of

$$\{d_{m,n} : m \in (1,2,.....,r), n(m) \in (1,2,......,2^{m-1})\} .$$

Daubechies (1988) has shown that $\{d_{m,n}\}$ is a complete orthonormal basis of $L^2(\Re)$ so that any X_t in the said space can be represented as

$$X_t = \sum_m \sum_{n(m)} d_{m,n} W_{m,n}(t)$$

where

$$d_{j,k} = < X, W_{m,n} > \tag{15}$$

is the wavelet coefficient.

Filter Banks

Since X_t is only known on a discrete set of points, $d_{m,n}$ is calculated by a two-channel filter bank. A two-channel filter bank representation of the wavelet transform consists of a low-pass filter

$$Z(t) = \sqrt{2}\sum_{k=0}^{2M-1} r_k Z(2t-k)$$

where $\{r_k\}$ are non-zero filter coefficients and a high-bandpass filter

$$W(t) = \sum_{k=0}^{2M-1} s_k Z(2t-k)$$

where $s_k = (-1)^k r_{2M-1-k}$. The function Z(t) is referred to as a scaling function.

The low-pass filter coefficients $\{r_k\}$ are moving average filters that smooths the high frequency traits (jumps, cusps, singularities) of a series. On the other hand, the high-bandpass filter coefficients $\{s_k\}$ act as a differencing operators that capture the details filtered out by the low-pass filter.

Defining the dilations and translations of Z as

$$Z_{m,n} = 2^{m/2} Z(2^m t - n)$$

where $m,n \in Z$, the filter bank definition of $W_{m,n}$ can be written as

$$2^{m/2} W(2^m t - n) = 2^{m/2} \sum_{k=0}^{2M-1} s_k Z(2^{m-1} t - 2n - k).$$

Using the high-bandpass filter definition of $W_{m,n}$ it follows that

$$d_{m,n} = 2^{m/2} \sum_{k=0}^{2M-1} s_k < X, W(2^{m-1} t - 2n - k) >$$

$$= \sqrt{2} \sum_{k=0}^{2M-1} s_k u_{m-1,2n+k} \tag{16}$$

where $u_{m,n} = 2^{m/2} < X, W(2^m t - n) >$ is the scaling coefficient.

Calculation of $u_{m,n}$ can be performed by writing $Z_{m,n}$ in terms of the low-pass filter as

$$2^{m/2} Z(2^m t - n) = 2^{(m-1)/2} \sum_{k=0}^{2M-1} r_k Z(2^{m-1} t - 2n - k).$$

Convoluting X_t with the above equation we find that $u_{m,n}$ can be expressed as

$$u_{m,n} = 2^{(m-1)/2} \sum_{k=0}^{2M-1} r_k < X, Z(2^{m-1}t - 2n - k) >$$
$$= \sum_{k=0}^{2M-1} r_k u_{m-1,2n+k} \quad (17)$$

Thus, both $u_{m,n}$ and $d_{m,n}$ are calculated recursively from the smallest to the largest scale with the simple multiplication and addition operators of a two-channel filtered bank.

5. SUGGESTED APPROACH

Consider an additive model for X_t such as the one given in (7) [If model is multiplicative consider $X_t' = \log X_t$, $M_t' = \log M_t$, $S_t' = \log S_t$, and $I_t' = \log I_t$.]. Define $g_Z(x)$, $g_M(x)$, $g_S(x)$ and $g_I(x)$ in (9) as follows:

$$g_Z(x) = \frac{U(x)}{V_M(x)V_S(x)} = Q(x) + \frac{R(x)}{V_M(x)V_S(x)}$$
$$= Q(x) + \frac{R_M(x)}{V_M(x)} + \frac{R_S(x)}{V_S(x)}$$
$$= g_I(x) + g_M(x) + g_S(x)$$

where

$U(x) = \theta(e^{iw})\theta(e^{-iw})$;
$V_M(x) = \varphi_M(e^{iw})\varphi_M(e^{-iw})$; and
$V_S(x) = \varphi_S(e^{iw})\varphi_S(e^{-iw})$.

The deseasonalized series is obtained as the sum of X_{1t} and X_{2t} where X_{1t} is obtained from (13) using forecasts of X_t, X_{2t} using backcasts of X_t and $K_S(.)$ is as defined in (12).

In deriving the ARIMA coefficients to be used in the filtering process, we employ an approximate wavelet maximum likelihood estimator (AWMLE) as given in Jensen (2000). Jensen introduced this type of estimators in estimating long and short memory parameters in ARIMA models. The procedure extracts an approximate MLE by defining a likelihood function based on the wavelet coefficients instead of the original observations $\{X_t\}$. Jensen's simulation experiments showed that the AWMLE fairs well in comparison to the usual techniques of estimating short and long memory parameters in ARIMA models.

From (15), the covariance between wavelet coefficients of different dilations $m \neq m'$ and translations $n, n' \in Z$ is given by

$$\text{cov}(d_{m,n}, d_{m',n'}) = E(d_{m,n}d_{m',n'})$$
$$= \int\int W_{m,n}(t)\gamma_X(|t-s|)W_{m',n'}(s)dtds$$
$$= 2^{(m+m')/2} \int f_X(w)e^{-i(2^m n - 2^{m'} n')w} \overline{\hat{W}(2^m w)\hat{W}(2^{m'} w)} dw \quad (18)$$

Using an ideal high-bandpass wavelet W(t) with Fourier transform

$$\hat{W}(w) = \begin{cases} 1, & |w| \in (\pi, 2\pi) \\ 0, & ow \end{cases},$$

one can show that the expression in (18) is zero. This result allows us to define an approximate log-likelihood function for $d_{m,n}$ based on the multivariate normal distribution.

Consequently, we have the following procedure for the estimation of the ARIMA parameters which are needed in implementing the optimal seasonal adjustment procedure discussed in section 3.

The procedure consists of the following steps:

1. Consider observations of X_t for $t = 1, 2, \ldots, 2^{max}$.
2. Define $u_{0,n} = X(n)$ and compute $d_{m,n}$ and $u_{m,n}$ recursively from (16) and (17), respectively. This will generate observations $\mathbf{d_m} = (d_{m,1}, d_{m,2}, \ldots, d_{m,2^{max-m}})'$, $m = 1, 2, \ldots, max$.
3. Letting $\mathbf{d} = (\mathbf{d_1}', \mathbf{d_2}', \ldots, \mathbf{d_{max}}')'$, it can be shown that $\mathbf{d}$ has a multivariate normal distribution with mean vector zero and covariance matrix $\Sigma = \mathrm{Diag}(\sigma_1 I_1, \sigma_2 I_2, \ldots, \sigma_{max} I_{max})$ where I_m is a $2^{max-m} \times 2^{max-m}$ identity matrix. Thus, the ARIMA parameters $\mu = (\Phi, \phi; \Theta, \theta)$ can be estimated through the maximization of the approximate log-likelihood function given by
$$L(\mu; w) = -\frac{1}{2}\sum_{m=1}^{max} 2^{max-m} \log|\hat{\sigma}^2 \sigma_m'| + \frac{d_m' d_m}{\hat{\sigma}^2 \sigma_m'}$$
where $\hat{\sigma} = \frac{1}{2^{max}-1}\sum_{m=1}^{max} \frac{d_m' d_m}{\hat{\sigma}^2 \sigma_m'}$. This function can be numerically maximized over the parameter space of μ to yield the AWMLE.

The above procedure can be implemented empirically using R's Wavethresh or Mathlab's Wavelab.

6. CONCLUDING REMARKS

It can be seen that the suggested approach in this paper consists of two major procedures. The first procedure deals with the estimation of the ARIMA parameters in (4). Once the parameters have been estimated using the AWMLE procedure they are used in the filtering procedure. This filtering procedure (MSX) consists of deriving $\delta_S(B,F)$ in (10) and applying it to X_t. This is carried out by applying (13) to the forecasts and backcasts of X_t using $K_S(.)$ defined in (12).

ACKNOWLEDGMENT

The author wishes to thank the Statistical Research and Training Center for supporting this research under its Research Fellowship Program for the year 2000, and the anonymous referee for the constructive comments.

References

ABRY, P., GONCALVES, P. and FLANDRIN, P. (1995). *Wavelets, Spectrum Analysis and 1/f Processes, In Wavelet and Statistics*. Springer-Verlag: New York

BERSALES, L.G.S. and SARTE, G. (2000). "Seasonal Adjustment of Time Series with Applications to Philippine Data". UP School of Statistics Teaching Monograph.

BOX, G.E.P., PIERCE, D.A., and NEWBOLD, P. (1987). "Estimating Trend and Growth Rates in Seasonal Time Series", *Journal of American Statistical Association*, 2, pp. 276-282.

BROCKWELL, P.J. and DAVIS, R.A. (1987). *Time Series: Theory and Methods*. Springer-Verlag: New York.

____________________ (1995). Introduction to Time Series and Forecasting. Springer-Verlag: New York.

BURMAN, J.P. (1965). "Moving Seasonal Adjustment of Economic Time Series", *J.R. Statist. Soc. A*, 128, pp. 534-558.

__________ (1980). "Seasonal Adjustment by Signal Extraction", *J.R. Statist. Soc. A*, 143, pp. 321-337.

CARLIN, J.B. and DEMPSTER, A.P. (1989). "Sensitivity Analysis of Seasonal Adjustment: Empirical Case Studies", *Journal of American Statistical Association.*, 84, pp. 6-32.

CHUI, C. K. (1997). Wavelets: "A Mathematical Tool for Signal Extraction". *SIAM Monographs*: Philadelphia.

CLEVELAND, W.P. and TIAO, G.C. (1976). "Decomposition of Seasonal Time Series: A Model for the Census X-11 Program", *J. Amer. Statist. Ass.*, 71, pp. 581-587.

DAUBECHIES, I. (1992). "Ten lectures on Wavelets". *SIAM Monographs*: Philadelphia.

DELYON, B. and JUDITSKY, A. (1995). "Estimating Wavelet Coefficients", Working Paper, IRISA-INRIA, Campus de Beaulieu, France.

GONZAGA, A. C. (2001). "Some Alternative Approaches in the Analysis of ARFIMA(p,d,q) Process". UP School of Statistics Dissertation.

JENSEN, M.J. (2000). "An Approximate Wavelet MLE of Short and Long Memory Parameters", Working Paper, Department of Economics, University of Missouri-Columbia.

LOUIS, A.K., MAAβ, P., and RIEDER, A. (1997). "Wavelets: Theory and Applications". J. Wiley and Sons:Chichester.

NASON, G.P. and SILVERMAN, B.W. (1995). "The Stationary Wavelet Transform and Some Statistical Applications", *In Lect. Notes in Statist.*, 103, pp. 281-289. Springer-Verlag: New York.

"Philippine Statistical Yearbook (2000)". National Statistical Coordination Board Publication.

SHISKIN, J., YOUNG, A., and MUSGRAVE, J.C. (1965). "The X-11 Variant of Census Method II Seasonal Adjustment Program". Technical Paper 15, US Department of Commerce, Bureau of the Census, Washington DC.

VON SACHS, R., NASON, G.P., and KROISANDT, G. (1996). "Spectral Representation and Estimation for Locally Stationary Wavelet Processes", Working Paper, Department of Mathematics, University of Bristol, UK.

WEI, W.W.S. (1990). "Time Series Analysis: Univariate and Multivariate Methods". Addison-Wesley: New York.

WHITTLE, P. (1963). "Prediction and Regulation by Linear and Least-Square Methods". D. Van Nostrarid: Princeton.

"X-11 Method". Encyclopedia of the Statistical Sciences.