Statistics Canada / Statistique Canada

Methodology Branch

Time Series Research and Analysis Division

Direction de la méthodologie

Division de la recherche et de l'analyse des chroniques

Canada

[redacted]

Working paper TSRA-87-018E
Time Series Research And Analysis Division
Methodology Branch
Statistics Canada

Sources of Variance in Seasonally Adjusted Data

by

Benoit Quenneville
Time Series Research and Analysis Division
Statistics Canada

November 1987

RESUME

L'objectif de ce rapport est de stimuler la discussion sur ce que devrait être la définition de la variance pour les données corrigées de leur variations saisonnières par le projiciel X-11-ARMMI dans une agence statistique tel que Statistique Canada. Cinq sources de variance sont présentées avec leur méthodes d'estimation. Il sagit de la variance du signal aléatoire, de la variance d'échantillonnage, de la variance provenant d'une méthode sous-optimale pour corriger des variations saisonnières, de la variance de révision et finallement, de la variance provenant de l'estimation des paramètres nécessaires à la méthode de correction pour les variations saisonnières.

Il est coutimier de décomposer une série chronologique Y_t comme la somme d'une composante saisonnière S_t et non-saisonnière N_t. La variance du signal aléatoire est la variance des estimateurs des composantes saisonnière et non-saisonnière sous les distributions aléatoires sous-jacentes à S_t et N_t. Etant donné les modèles aléatoires pour les distributions de la série chronologique et ses composantes saisonnière et non-saisonnière, le filtrage de Kalman est un outil permettant d'évaluer la variance des signaux aléatoires. Du point de vue pratique, cette variance est nécessaire pour construire des intervalles de confiance pour les estimés des signaux aléatoires et de leurs projections temporelles.

L'erreur d'échantillonnage est présente dans les données lorsqu'un échantillon aléatoire de la population est sélectionné pour produire un estimé y_t de Y_t. Si Y_t est considéré comme une constante fixe ou bien une variable aléatoire l'erreur d'échantillonnage est la même. Les covariances d'échantillonnage pour les données corrigées de leur variations saisonnières peuvent se calculer de deux manières. La première utilise une approximation linéaires des filtres de la méthode X-11 et la seconde utilise les méthodes de ré-échantillonnage. La variance d'échantillonnage des données corrigées pour leur variations saisonnières permet une comparaison avec les données non-corrigées.

La variance provenant de l'utilisation d'une méthode sous-optimale pour corrigés des variations saisonnières est introduite en remplaçant la méthode optimale par la méthode X-11-ARMMI. Cette variance donne une

mesure de qualité de la correction apportée par X-11-ARMMI.

X-11-ARMMI produit de meilleur estimés en utilisant des filtres de plus en plus symmétriques lorsque des données futures sont disponibles. La variance introduite par un changement de filtre saisonnier fait partie de la variance de révision. Cette variance donne une mesure de qualité de X-11-ARMMI et peut être produite independamment des données. L'autre partie de la variance de révision dépend des données. Elle mesure la réduction dans les variances des signaux aléatoires provenant de l'utilisation des données futures. Cette variance se calcule ent utilisant le lissage de Kalman.

Finallement, la variance provenant de l'estimation des paramètres nécessaires à la méthode de correction pour les variations saisonnières donne une mesure de qualité sur la méthode d'estimation des paramètres et l'effet de l'utilisation de paramètres estimés (à l'opposé des vrais paramètres) sur la variance totale.

EXECUTIVE SUMMARY

This report aims at precisely defining the variance of seasonally adjusted data produced by the X-11-ARIMA seasonal adjustment in a statistical agency such as Statistics Canada. Five sources of variances in seasonally adjusted data are presented along with their estimation methods. The sources are the signal variance, the measurement error variance, the increase in variance from the use of a suboptimal seasonal adjustment procedure, the revision variance and the increase in variance from the parameter estimation of a seasonal adjustment procedure.

It is a common practice in time series to decompose a time series Y_t into seasonal S_t and nonseasonal N_t components such that:

$$Y_t = N_t + S_t.$$

The signal variance is the variance of the estimator of the nonseasonal component under the stochastic model underlying the distribution of N_t. Given stochastic models for the series and its components, the Kalman filter provides a mean of computing the signal variance. The signal variance is needed to construct confidence intervals around the nonseasonal component and eventual forecasts.

The measurement error variance is restricted to the sampling variance which arises in the data because a sample of the population is selected instead of the whole population to produce an estimate y_t of Y_t. Whether or not there is a stochastic model on the time series Y_t, the sampling variance is the same. Sampling covariances of seasonally adjusted data can be computed in two ways. The first one uses a linear approximation of the X-11 filters and the second one uses resampling methods. The sampling variance of the seasonally adjusted data allows a comparison with the unadjusted figure.

The variance from the use of a suboptimal adjustment procedure is defined as the increase in the variance of the estimators of the components introduced by replacing an "optimal" method by the X-11-ARIMA method. This variance provides a measure of the quality of the seasonal adjustment performed by X-11-ARIMA.

As more data are available, X-11-ARIMA produces "better" estimates by using more symmetric linear filters. The variances introduced by a change

in the filters is part of the revision variance. This part of the revision variance provides a measure of quality of X-11-ARIMA and can be produced independently of the data. The other part of the revision variance is dependent on the data. It measures the reduction in the signal variance from using future observations. Smoothing theory provides a way of computing this variance. In practice it can be obtained as a by-product of the Kalman filter and smoother.

Finally, the variance from the estimation of the parameters of a seasonal adjustment procedure is the contribution to the variance (signal or sampling) arising from the use of estimated (as opposed to true) parameters (e.g.: ARIMA parameters in X-11-ARIMA). This variance estimation provides a quality measure on how "well" the parameter estimation was done and the resulting effect on the total variance.

1. INTRODUCTION.

1.1 Subject and purpose.

This report aims at precisely defining the variance of seasonally adjusted data produced by the X-11-ARIMA seasonal adjustment method in a statistical agency such as Statistics Canada. It is intended to stimulate the discussion on the topic to help decide which definition should be used in an option of the X-11-ARIMA package.

1.2 Content of the report.

Hausman and Watson (1985) have identified five sources of variance in seasonally adjusted data. The sources are the signal variance, the measurement error variance, the increase in variance from the use of a suboptimal seasonal adjustment procedure, the revision variance and finally, the increase in variance from the estimation of the parameters of a seasonal adjustment procedure. This report presents these sources along with their estimation methods. Section 2 introduces some preliminary concepts used throughout the report. Section 3 discusses the sources of variance. Section 4 gives the conclusions of the study.

2. CONCEPTS AND NOTATION.

2.1 Seasonal Adjustment Concepts and Methods.

The additive variant of seasonal adjustment assumes that the true series Y_t (read capital Y_t) can be decomposed into the trend-cycle C_t, the seasonal S_t, the trading day D_t and the irregular variations I_t:

$$Y_t = C_t + S_t + D_t + I_t. \qquad (2.1)$$

It is sometimes useful to also write Y_t as the sum of a nonseasonal N_t and a seasonal S_t components:

$$Y_t = N_t + S_t. \qquad (2.2)$$

The seasonally adjusted series is obtained by removing from the observed time series y_t (read little y_t) an estimate s_t and d_t of the seasonal S_t and trading-day D_t components. For an adjustment performed by using information available at time t+k the adjusted value at time t is:

$$y^a_{t,t+k} = y_t - s_{t,t+k} - d_{t,t+k} \qquad (2.3)$$

2.1.1 The X-11 and X-11-ARIMA methods.

The X-11 seasonal adjustment method developed by Shiskin et al.(1967) provides estimates of $y^a_{t,t+k}$. The method is based on moving averages or linear smoothing filters and moving averages assume that the time series components are functions of time which cannot be closely approximated by simple functions of time over the entire series. Implicit assumptions are that the trend, the cycle and seasonal components are stochastic and not deterministic.

If the procedures for outlier identification and modification and the procedures for trading-day adjustment are ignored, then the X-11 seasonal adjustment in its additive form can be represented by a single set of moving averages. For an observation sufficiently far from the end points of the serie ($m+1 <= t <= T-m$) the seasonally adjusted value is obtained with a symmetric filter $a_m(L)$:

$$y^a_{t,t+m} = a_m(L) y_t = \sum_{j=-m}^{m} a_{m,j} y_{t-j} \qquad (2.4)$$

where L is the lag operator ($Lx_t = x_{t-1}$) and $a_{m,j} = a_{m,-j}$. For current and recent data ($t-m+1 < t <= T$) asymmetric filters $a_i(L)$ have to be used leading to:

$$y^a_{t,t+i} = a_i(L)y_t = \sum_{j=-i}^{m} a_{i,j}y_{t-j} \qquad i=0,\ldots,m-1 \qquad (2.5)$$

Details can be found in Young (1968) and Wallis (1982)

The X-11-ARIMA method, developed by Dagum (1980), is an extension of X-11. The method consists of extending the original series at each end with extrapolated values from seasonal AutoRregressive Integrated Moving Average (ARIMA) models of the Box and Jenkins (1970) type and then of seasonally adjusting the extended series with the X-11 filters.

If z_t follows an ARIMA model the process will be written in the form:

$$\phi_z(L)z_t = \theta_z(L)a_t$$

where $\phi_z(L)$ is the autoregressive polynomial that include the difference operator $(1-L)^d(1-L^s)^D$ (needed to make the process stationary), $\theta_z(L)$ is the moving average polynomial and a_t is a white noise sequence.

2.1.2 The model-based approaches

During the last decade a great effort has been made by researchers in the development of seasonal adjustment methods based on the decomposition of univariate time series models. Two approaches have been followed for the development of "model-based" seasonal adjustment methods. In the first approach the observed series is assumed to follow an ARIMA model. From that model, resulting ARIMA models are derived for the components. In the second approach, each of the observed components is assumed to follow a stochastic model. The first model-based approach is called "reduced-form" and the second, "structual". Major properties and operational limitations of these two approaches are discussed by Dagum (1987). In the remaining part of this sub-section the two approaches are explained.

2.1.2.1 The reduced-form approach.

In the reduced-form approach the trading-day component and any other kind of variations of a deterministic character, e.g. holiday effects, has to be removed from the original series. One way of doing it is discussed in Bell and Hillmer (1983). The input series is thus composed of a trend-cycle, a seasonal and an irregular component.

Hillmer and Tiao (1982) have proposed the general procedure for the

ARIMA modeling of the unobserved components. They assume that each of the components of Y_t follows an ARIMA model:

$$\phi_S(L)S_t = \theta_S(L)b_t$$
$$\phi_C(L)C_t = \theta_C(L)e_t$$
$$\phi_I(L)I_t = \theta_I(L)d_t$$

where each of the pairs of the autoregressive polynomials have their zeros lying on or outside the unit circle and have no common zeros, and b_t, e_t and d_t are three independent white noise processes, identically and independently distributed as $N(0,\sigma_b^2)$, $N(0,\sigma_e^2)$ and $N(0,\sigma_d^2)$, respectively. Given a known model of Y_t:

$$\phi_y(L)Y_t = \theta_Y(L)a_t$$

and the restrictions:

i) $\phi_y(L) = (1-L)^d(1+L+...+L^{11})\phi_I(L)$,

ii) the order of θ_S is less than or equal to 11,

iii) $\sigma_d^2 = \text{var}(d_t)$ is maximized,

Hillmer and Tiao (1982) derived what they called the "canonical" decomposition of Y_t. They prove in particular that the canonical decomposition is unique and minimizes the innovation variances σ_b^2 and σ_e^2. In practice a model for Y_t is identified from the data, and this model and its corresponding estimated parameter values are used as if they were true.

2.1.2.2 The structural approach

When structural models are used for the unobserved components the trading-day component can either be removed from the original series as in the reduced-form approach or modelled with the other components. For simplicity of the discussion we shall assume that there is no trading-day component.

The use of structural models is mainly discussed by Harvey and Todd (1983) and Harvey (1984): " The basic structural model has the form:

$$Y_t = \mu_t + \gamma_t + \epsilon_t \quad , t=1,...,T$$

where μ_t, γ_t and ϵ_t are the trend ,seasonal and irregular components respectively.

The process generating the trend is of the form:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t,$$

and

$$\beta_t = \beta_{t-1} + \zeta_t,$$

where η_t and ζ_t are normally distributed independent white noise processes with zero means and variance σ_η^2 and σ_ζ^2 respectively. The essential feature of this model is that it is a local approximation to a linear trend. The level and slope both change slowly over time according to a random walk mechanism.

The process generating the seasonal component is:

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t$$

where ω_t is distributed as NID$(0,\sigma_\omega^2)$ and s is the number of "seasons" in the year. The seasonal pattern is thus slowly changing but by a mechanism that ensures that the sum of the seasonal components over any s consecutive time periods has an expected value of zero and a variance that remains constant over time...

The disturbances η_t, ζ_t and ω_t are independent of each other and of the irregular component that is a normally distributed white noise process, that is, ϵ_t is distributed as NID$(0,\sigma^2)$."

2.2 Survey Errors.

2.2.1 Sampling and non-sampling errors.

The general problem of sampling theory is to find an optimal combination of a design and an estimator $y_t(*)$ of Y_t. The design is the strategy that selects a sample from the population. The estimator is a function that combines the observations from the selected units (and possibly external information) to produce an estimate of the characteristic of interest, namely Y_t.

Typically the Mean Square Error (MSE) of $y_t(*)$, defined as $E[y_t(*)-Y_t]^2$ measures the uncertainty or the error in $y_t(*)$. Samplers have identified two major sources of errors in surveys. They are the sampling and the non-sampling errors.

The sampling error arises from setecting a sample instead of the whole population. For example, in a census the sampling error is zero. The MSE of $y_t(*)$ is :

$$E[y_t(*)-Y_t]^2 = \sum_{s \epsilon S} p(s)[y_t(s)-Y_t]^2 \qquad (2.6)$$

where S is the set of all possible samples s, p(s) is the probability of selecting the sample s and $y_t(s)$ is the estimate of Y_t, the true value, based on the sample s. Under this notation $y_t(s)$ is the estimator $y_t(*)$ evaluated from sample s.

The MSE of $y_t(*)$ can be further decomposed as:

$$E[y_t(*)-Y_t]^2 = E[y_t(*)-E(y_t(*))]^2 + [E(y_t(*)-Y_t)]^2 \qquad (2.7)$$

where the first term on the right hand side of (2.7) is the sampling variance of $y_t(*)$ and the second term is the square bias of $y_t(*)$. In practice the bias is usually assumed to be equal to zero by choosing an unbiased or nearly unbiased estimator. It is impossible to compute the sampling variance from (2.6) since there is only one selected sample. The sampling variance is therefore estimated from the sample by means of a variance estimator. In practice the process stops there and there is no further estimation of the variance of the variance estimator.

The non-sampling errors are all the other possible errors. They include undercoverage or overcoverage, nonresponse and response errors, and processing errors. More details on these errors as well as their estimation can be found in the Quality Guidelines published by Statistics Canada

(1985). In practice measures of non-sampling errors are computed for some specific type of errors but are rarely published.

In the context of this report the discussion will be restricted to the sampling variance assuming that the estimator is unbiased with respect to the design. Thus we assume:

$$y_t(*) = Y_t + u_t \qquad (2.8)$$

where u_t is a random variable representing the sampling error.

2.2.2 Sampling covariances.

The estimation of the sampling variance of $y_t(*)$ depends on the design used. Many surveys produce an estimate of the sampling variance of $y_t(*)$ as part of their survey operations. If the samples for the survey at different time periods are drawn independently of each other then the sampling covariances between $y_{t1}(*)$ and $y_{t2}(*)$ for t1 different than t2 is zero. But in practice it is seldom the case and correlation exists between the two estimates. From equation (2.8) we obtain the following definition for the sampling covariance between y_{t1} and y_{t2}:

$$E[y_{t1}(*)-Y_{t1}][y_{t2}(*)-Y_{t2}] = E[u_{t1}u_{t2}] \qquad (2.9)$$

When there is no change in the population composition being surveyed on two occasions, Tam (1984) gives the necessary results to compute the covariances for different sampling plan under simple random sampling. Laniel (1987) extends the results when there are births and deaths in the population between two occasions. The extension to a more complex design is not obvious.

Another approach to the estimation of the sampling covariance is to assume that the sampling error follows a stochastic model. In a panel survey where a unit stays in the survey for a fixed number of occasions, say q, $\{u_t\}$ can be assumed to follow a MA(q) process. Inded the sampling covariances between $y_{t1}(*)$ and $y_{t2}(*)$ for $|t1-t2|>q$ are zero which defines a MA(q) process. In rotating surveys it might be reasonable to assume that u_t is correlation stationary. Thus we can assume that $\rho(u_t,u_{t+k}) = \rho(k)$. This approach has been used by Wolter and Monsour (1981). The special case of $\rho(k) = \exp(-\lambda k)$ has been successfully used by Quenneville and Srinath (1984) in the context of the Survey of Employment, Payroll and Hours.

It has to be recognized that the problem of the estimation of the sampling variances for complex surveys is not straightforward. When non-linear statistics such as the ratio estimator are used the problem is even more complex, see Kovar (1985,1987). The extension to the sampling covariances is not solved yet even thought it could be straightforward for some cases.

3 SOURCES OF VARIANCE IN SEASONALLY ADJUSTED DATA.

This section now defines the five sources of variance in the seasonally adjusted data.

3.1 Signal variance.

3.1.1 Definition.

Assume that the time series Y_t is observed without sampling errors. From (2.8) and (2.2):

$$y_t = Y_t = N_t + S_t \ .$$

The signal variance is defined as $E_m[n_t - N_t]^2$ where E_m denotes the expectation under the model (or stochastic process) generating N_t and n_t is an unbiased estimator of N_t. This means that under repeated realisation of the stochastic process N_t, the expectation of n_t is equal to the expectation of N_t.

Unfortuneately, there are no explicit stochastic models in X-11 (or X-11-ARIMA) but under certain assumptions implicit models have been derived by Cleveland and Tiao (1976) and Burridge and Wallis (1984). There models are now given.

According to Cleveland and Tiao (1976), the default options of the additive version of X-11 can be approximate by a model with stochastic trend-cycle and seasonal components. Very specifically:

$$Y_t = C_t + S_t + I_t$$
$$(1-L)^2 C_t = (1+.49L-.49L^2) b_{1t}$$
$$(1-L)^{12} S_t = (1+.64L^{12}+.83L^{24}) b_{2t}$$
$$I_t = e_t \text{ (white noise)}$$
$$\sigma^2_{b1}/\sigma^2_{b2} = 1.3 \text{ and } \sigma_e^2/\sigma^2_{b1} = 14.4.$$

The resulting overall model for Y_t is:

$$(1-L)(1-L^{12})Y_t = (1-.337L+.141L^2+.141L^3+.139L^4+.136L^5$$
$$+.131L^6+.125L^7+.117L^8+.106L^9+.093L^{10}$$
$$+.077L^{11}-.417L^{12}+.232L^{13}-.001L^{20}-.003L^{21}$$
$$-.004L^{22}-.006L^{23}+.035L^{24}-.021L^{25})c_t.$$

Burridge and Wallis (1984) extend Cleveland and Tiao's analysis to the concurrent and first-year revised asymmetric filters $a_0(L)$ and $a_{12}(L)$. They use a technique different from Cleveland and Tiao. In their model for the final symmetric filter $a_{84}(L)$ the coefficients of the polymomial $\theta(L)$ are different. The models of the components S_t and N_t for $a_0(L)$, $a_{12}(L)$ and $a_{84}(L)$ are:

$$a_0(L): \quad (1+L+\ldots+L^{11})S_t = (1+1.00L^{24})a_t$$
$$(1-L)^2N_t = (1-1.43L+.70L^2)b_t$$
$$\sigma_a{}^2/\sigma_b{}^2 = .018$$
$$a_{12}(L): \quad (1+L+\ldots+L^{11})S_t = (1+.33L^{12}+.99L^{24})a_t$$
$$(1-L)^2N_t = (1-1.55L+.82L^2)b_t$$
$$\sigma_a{}^2/\sigma_b{}^2 = .026$$
$$a_{84}(L): \quad (1+L+\ldots+L^{11})S_t = (1+.71L^{12}+1.00L^{24})a_t$$
$$(1-L)^2N_t = (1-1.59L+.86L^2)b_t$$
$$\sigma_a{}^2/\sigma_b{}^2 = .017\ .$$

The overall model for Y_t is:

$$(1-L)(1-L^{12})Y_t = \beta(L)e_t$$

where the moving average polynomial $\beta(L)$ is of degree 26 and the coefficients are given in Table 1 for each of the three cases.

TABLE 1

Coefficients of the Composite Moving Average Operator

$\beta(L) = 1 + \beta_1 L + \ldots + \beta_{26} L^{26}$

L	$a_0(L)$	$a_{12}(L)$	$a_{84}(L)$
1	-.53	-.64	-.67
2	.28	.30	.29
3	.23	.23	.23
4	.23	.22	.22
5	.22	.22	.21
6	.22	.21	.20
7	.21	.20	.20
8	.21	.20	.20
9	.20	.19	.19
10	.19	.18	.18
11	.14	.13	.12
12	-.43	-.33	-.33
13	.45	.44	.46
14	*	.02	.02
15	*	*	*
16	*	*	*
17	*	*	*
18	*	*	*
19	*	*	*
20	*	*	*
21	*	*	*
22	*	*	*
23	-.01	-.01	-.01
24	.05	.05	.04
25	-.04	-.05	-.04
26	.01	.02	.01
σ_a^2/σ_e^2	.012	.016	.011

*: indicates a coefficient less than .01 in absolute value.

3.1.2 Estimation.

For series where the true models for the components are those given in the previous sections, the X-11 (or the X-11-ARIMA) program could be used to estimate the unobserved trend and seasonal components in an optimal way. Here, optimal means that the signal variance is minimum. The X-11 (or X-11-ARIMA) seasonal adjustment method does not provide an estimate of the signal variance. One way of doing it is now described.

Given the models on the seasonal S_t and the trend-cycle C_t, the optimal estimators s_t and c_t are derived along with their mean square errors $E(s_t-S_t)^2$ and $E(c_t-C_t)^2$ using a state space representation and the Kalman filter. We shall assume for simplicity that the irregular component is a white noise process. Details of the discussion are given in Burridge and Wallis (1985) but the main ideas are presented here.

The models for S_t, C_t, I_t and Y_t are written in a state-space form which consists of a state transition equation and a measurement equation:

$$X_{t+1} = FX_t + GW_{t+1} \qquad (3.1)$$

$$Y_t = H'X_t + I_t. \qquad (3.2)$$

In general X_t and Y_t respectively denote the state vector and the output vector. W_t and I_t are independent serially uncorrelated normal random vectors with means zero and covariance matrix Q and R. Here Y_t is a scalar. Denoting the degree of the lag polynomials of S_t and C_t in the models:

$$\phi_S(L)S_t = \theta_S(L)w_{1t} \qquad (3.3)$$

$$\phi_C(L)C_t = \theta_C(L)w_{2t} \qquad (3.4)$$

by m,n,p,q, respectively, a convenient state-space representation of the unobserved-component model is obtained through the following definitions and equivalences:

$x_t = (x'_{1t}\ x'_{2t})'$

$x_{1t} = (S_t, S_{t-1}, \ldots, S_{t-m+1}, w_{1,t}, w_{1,t-1}, \ldots, w_{1,t-n+1})'$

$x_{2t} = (C_t, C_{t-1}, \ldots, C_{t-p+1}, w_{2,t}, w_{2,t-1}, \ldots, w_{2,t-q+1})'$

F = block diagonal $[F_1, F_2]$

$$F_1 = \begin{bmatrix} \phi_{s,1} & \phi_{s,2} \cdots & \phi_{s,m-1} & \phi_{s,m} & -\theta_{s,1} & -\theta_{s,2} \cdots & -\theta_{s,n-1} & -\theta_{s,n} \\ 1 & 0 \ \ldots & 0 & 0 & & & & \\ 0 & 1 \ \ldots & 0 & 0 & & 0 & & \\ \cdot & & \cdot & & & & & \\ \cdot & & \cdot & & & & & \\ \cdot & & \cdot & & & & & \\ 0 & & 1 & 0 & & & & \\ & & & & 0 & 0 \ \ldots & & 0 \\ & & & & 1 & 0 \ \ldots & & 0 \\ & 0 & & & 0 & 1 \ \ldots & & 0 \\ & & & & \cdot & & & \cdot \\ & & & & \cdot & & & \cdot \\ & & & & \cdot & & & \cdot \\ & & & & 0 & 0 \ \ldots & & 0 \end{bmatrix}$$

F_2 is similarly defined by matching coefficients in the model for C_t to elements of x_{2t}

$$G' = \begin{bmatrix} 1\ 0\ \ldots\ 0\ 1\ 0\ \ldots\ 0\ 0\ 0\ \ldots\ 0\ 0\ 0\ \ldots\ 0 \\ 0\ 0\ \ldots\ 0\ 0\ 0\ \ldots\ 0\ 1\ 0\ \ldots\ 0\ 1\ 0\ \ldots\ 0 \end{bmatrix}$$

$$H' = [1\ 0\ \ldots \qquad 0\ 1\ 0\ \ldots \qquad 0]$$

$w_t = (w_{1t}, w_{2t})'$ $\quad Q = \mathrm{diag}(\sigma^2_{w1}, \sigma^2_{w2})$

$R = \sigma^2_v$ (scalar).

The specification is completed by the knowledge of the initial state vector X_0 assumed to be normally distributed with a known mean x_0 and variance matrix P_0.

Let us assume that there are observations available up to time T, then the minimum mean square estimate x_T of X_T is given by:

$$x_T = x_{T/T-1} + K_T\, y_T^i \tag{3.5}$$

$$x_{T+1/T} = F x_T \tag{3.6}$$

with covariance matrix $P_T = E[x_T - X_T]^2$ given by:

$$P_T = (I - K_T H')P_{T/T-1} \tag{3.7}$$

$$P_{T+1/T} = F P_T F' + GQG' \tag{3.8}$$

where

$$y_T^i = Y_T - H' x_{T/T-1} \tag{3.9}$$

$$f_T = H' P_{T/T-1} H + R \tag{3.10}$$

$$K_T = P_{T/T-1} H / f_T . \tag{3.11}$$

The most common structual models used the state-space representation obtained through the following definitions and equivalences, where it is assumed quarterly observations:

$$X_t = [\ \mu_t\ \beta_t\ \gamma_t\ \gamma_{t-1}\ \gamma_{t-2}]'$$

$$F = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$G = I$$

$$W_t = [\eta_t\ \zeta_t\ \omega_t\ 0\ 0]'$$

$$H' = [1\ 0\ 1\ 0\ 0]$$

$$Q = \mathrm{diag}[\sigma_\eta^2\ \sigma_\zeta^2\ \sigma_\omega^2\ 0\ 0]$$

$$R = \sigma^2.$$

Kalman filter theory as described by equations (3.5) to (3.11) can be applied to derive the minimum mean square estimate x_t of X_t along with its MSE matrix but it requires knowledge of the variances σ_η^2, σ_ζ^2, σ_ω^2 and σ^2.

When the variances σ_η^2, σ_ζ^2, and σ_ω^2 are expressed relative to σ^2 then the likelihood function can be written in the form:

$$-2\mathrm{Log}\ L(\sigma_\eta^2,\sigma_\zeta^2,\sigma_\omega^2,\sigma^2) = T\log(2\pi) + T\log(\sigma^2) + T\sum_{t=1}^{T}\log(f_t) + \sigma^{-2}\sum_{t=1}^{T}(y_t^{i})^2/f_t \qquad (3.12)$$

where Q has been redefined as being equal to $\mathrm{diag}(\sigma_\eta^2/\sigma^2,\sigma_\zeta^2/\sigma^2,\sigma_\omega^2/\sigma^2,0,0)$ and R is equal to 1. In this case the MSE matrix of x_t becomes $\sigma^2 P_t$.

Differentiation of (3.12) with respect to σ leads to the maximum likelihood estimator of σ^2:

$$s^2 = T^{-1}\sum_{t=1}^{T} v_t^2/f_t$$

It is now possible to take out σ^2 out of (3.12), leaving the concentrated log likelihood function:

$$-2L_c = T\log(2\pi) + T + T\log(s^2) + \sum_{t=1}^{T}\log(f_t). \qquad (3.13)$$

Numerical optimization has to be carried out with respect to the remaining three parameters. For example the Davidon-Fletcher-Powell algorithm can be used.

3.2 Measurement error variance.

3.2.1 Definition.

This section discusses the estimation of the nonseasonal component of a time series Y_t, when the latter is observed through y_t, subject to sampling errors, and the estimation of the sampling variance of $y^a_{t,t+k}$.

To reformulate the problem, we assume:

$$Y_t = N_t + S_t \quad [2.2]$$

and it is observed through:

$$y_t = Y_t + u_t \quad [2.8]$$

or equivalently:

$$y_t = N_t + S_t + u_t. \quad (3.14)$$

Seasonal adjustment remove S_t from y_t whereas the estimation of N_t requires that both S_t and u_t be purged from the observed series y_t.

Since the irregular component I_t is often modelled as a white-noise process, the estimation of N_t from y_t or from Y_t would yield similar results if u_t is white noise. The latter assumption is valid when the surveys are independent of each other and the same design has been used. However, most surveys conducted by Statistics Canada are overlapping sample designs. The Labour Force Survey, Wholesale and Retail Trade Surveys and the Survey of Employment, Payroll and Hours are examples. Such designs generally produce sampling errors with some form of serial correlation. Questions then arise: i) what effect does this serially correlated measurement error have on the estimation of the non-seasonal component N_t, and ii) what is the increase in the variance of the estimator of N_t?

From X-11 :

$$Y^a_{t,t+i} = a_i(L)Y_t = a_i(L)(y_t - u_t) = a_i(L)y_t - a_i(L)u_t.$$

If $y^a_{t,t+i}$ denotes $a_i(L)y_t$, the published adjusted figure, then:

$$y^a_{t,t+i} = Y^a_{t,t+i} + a_i(L)u_t. \quad (3.15)$$

Now since u_t has mean zero under both the design and the model, $y^a_{t,t+i}$ has the same expectation as $Y^a_{t,t+i}$. Turning to variance and conditioning on the sample realisation we obtain:

$$V_m(y^a_{t,t+i}) = E_m V_s(y^a_{t,t+i}) + V_m E_s(y^a_{t,t+i}))$$

which reduces to:

$$V_m(y^a_{t,t+i}) = E_m V_s(a_i(L)u_t) + V_m(Y^a_{t,t+i}). \quad (3.16)$$

where E_m and V_m denote the expectation and variance under the stochastic models.

Equations (3.15) and (3.16) are the key equations for the discussion of the estimation of the sampling variance of the seasonally adjusted data.

For classical survey samplers Y_t $t=1,\ldots,T$ are considered as fixed parameters. Therefore, from (3.15) the sampling variance of the seasonally adjusted data is given by:

$$V_s(y^a_{t,t+i}) = V_s(a_i(L)u_t). \qquad (3.17)$$

But for time series analysts, there is a stochastic process for Y_t, so that the sampling variance is a component of the model variance (3.16). If the design is unbiased for all t, then Y_t and u_t are uncorrelated time series (Bell and Hillmer (1987)). In this case the total variance of $y^a_{t,t+i}$ is the sum of the sampling variance of $y^a_{t,t+i}$ and the signal variance of $Y^a_{t,t+i}$.

The next questions to ask are: i) is there an estimator of N_t with a smaller variance than $y^a_{t,t+i}$ given we usually have a knowledge of the sampling design and ii) how to estimate the sampling variance of the seasonally adjusted data produced by X-11 or X-11-ARIMA?

Those questions are now answered.

3.2.2 Estimation.

3.2.2.1 Optimal estimator of N_t.

The answer to the first question is yes. In fact it requires a generalisation of the optimal filter presented in section 3.1.2 where the measurement error is assumed to be autocorrelated instead of purely random. Here autocorrealted error means that the sampling error u_t is itself the output of a linear system of the form:

$$\chi_{t+1} = A\chi_t + B\lambda_{t+1}$$
$$u_t = C'\chi_t.$$

For instance, if u_t is assumed to be an AR(1) process with autoregressive parameter ϕ, then we let $C'=1$, $\chi_t=u_t$, $A=\phi$, $B=1$ and $\{\lambda_t\}$ is assumed to be a white noise sequence. If u_t is assumed to be a MA(1) process with moving average parameter θ, then we have $u_t=v_t+\theta v_{t-1}$, $\chi_t'=[u_t,\theta v_t]$, $C'_t=[1\ 0]$, $B'=[1\ \theta]$, $\lambda_t=v_t$ and

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

An optimal filter, knowm as the augmented Kalman filter is obtained by redefining the state vector X_t to include χ_t. Details are omitted (ref. Anderson and Moore(1979)).

3.2.2.2 Sampling covariances of seasonally adjusted data, Approach 1.

To facilitate the discussion let $y = (y_1,\ldots,y_T)'$ and $y^* = (y^*_{-11},\ldots,y^*_0,y,y^*_{T+1},\ldots,y^*_{T+12})'$ where y^*_j $j=-11,\ldots 0,T+1,\ldots,T+12$ denote the ARIMA extrapolated values obtained from X-11-ARIMA.

Also let A be the (TxT) matrix where the t^{th} row represents the moving average weights used to produce the seasonaly adjusted figure (ref: equation (2.5)) and A^* be the (T+24 x T+24) matrix when ARIMA extrapolated values are used.

Denote by y^a the vector of the seasonaly adjusted data and by y^{*a} the vector of seasonally adjusted data when ARIMA extrapolations are used. Then we have $y^a = Ay$ and $y^{*a} = A^*y^*$.

If we denote by $V(y)$ the sampling covariance matrix of y then we have the sampling covariance matrix of y^a : $V(y^a) = AV(y)A'$.

Denoting by $V(y^*)$ the sampling covariance matrix of y^* we obtain the sampling covariance matrix of y^{*a} : $V(y^{*a}) = A^{*}V(y^{*})A^{*\prime}$.

Note that the first and last twelve rows, as well as the first and last twelve elements of each row of $V(y^{*a})$ represent the sampling variance and/or covariance of the ARIMA extrapolations with either an ARIMA extrapolation or a true value. Now since the ARIMA extrapolations are linear functions of the vector y there should be no problem in evaluating their sampling covariances given the parameters of the ARIMA model are knowm.

The problem of the estimation of the sampling covariance of the seasonally adjusted data is therefore reduce to the problem of the estimation of the sampling covariance matrix of y as discussed in section 2.2.2.

3.2.2.3 Sampling covariances of seasonally adjusted data, Approach 2

In the approach discussed in section 3.2.2 the X-11-ARIMA and the X-11 filters are approximated by the filters $a_0(L),\ldots,a_m(L)$. Moreover the sampling covariance matrix of y is not straightforward unless simple hypotheses are assumed.

This has lead to estimate the sampling variability of seasonally adjusted data directly using resampling methods.

If a resampling method such as the random groups, balanced fractional samples, jacknife or the bootstrap are used for the estimation of the sampling variance of y_t then the sampling covariances of the seasonally adjusted data can be estimated. This is achieved by first seasonally adjusting the series corresponding to each replicate and then computing the covariance between the deseasonalized replicate values.

This approach has been used by Wolter and Monsour(1981) and Armstrong and Gray(1986).

3.3 Suboptimal seasonal adjustment procedure variance.

3.3.1 Definition.

We saw in section 3.1.2 and 3.2.2.1 how to obtain the optimal estimate of the seasonal and nonseasonal components given an ARIMA or a structural model on Y_t. In this section we would like to know when X-11 is optimal? That is, what is the class of stochastic models for which X-11 or X-11-ARIMA are optimal and if X-11 is not optimal for the given models, what is the increase in the signal variance?

These questions do not seem to have been answered yet. However, the literature provides at least ont type of ARIMA models behind X-11, Cleveland and Tiao (1976) and Burridge and Wallis (1984) (ref: section 3.1.1).

If a given series follows these models, then X-11 could be used to estimate the unobserved trend and seasonal components. A natural question that arises is if the X-11 procedure robust to departure from this model?

By analysing two sets of data Cleveland and Tiao (1976) suspect that a series obeying $(1-L)(1-L^{12})y_t = \theta(L)c_t$ for some polynomial $\theta(L)$ might be fairly accurately analysed by X-11 but this robustness does not extend to models which differ markedly from the overall model they described.

They do not provide, however, a measure of increase in the signal variance by using X-11 instead of the optimal method described in sections 3.1.2 and 3.2.2.1. Hence, given models on Y_t , its components S_t and N_t, and on the sampling variance u_t the question that remains to be answered is: what is the increase in $E[s_t-S_t]^2$ and $E[n_t-N_t]^2$ introduced by replacing the optimal method by the X-11 method?

3.3.2 Estimation.

There is no general answer to the last question since any answer depends on the assumed models.

One specific study, of interest for a statistical agency, has been made by Hausman and Watson (1985). They consider the effect of ignoring the measurement error (sampling error) on the "optimal" procedure and on the X-11 method by studying two American series: the civilian unemployment rate and the teenage unemployment rate. The second series exhibits more dramatic seasonal behavior and is subject to more severe measurement error than the first one.

They show that for the overall unemployment rate that this effect is negligible whereas it is considerable for the teenage unemployment rate. They conclude that there may be large gains from the use of a model-based rather than X-11 seasonal adjustment methods for the teenage unemployment rate.

As a related issue, they consider the optimal sample design. If the objective of the survey is the estimation of the nonseasonal component and if X-11 is to be used, they conclude that the rotation scheme should be chosen to make the measurement error as seasonal as possible (see equation (3.15).

3.4 Revision variance.

3.4.1 Definition.

Future values of the observed series usually contain information on the value of the current nonseasonal component. This statement is valid when both X-11-ARIMA or a model-based procedure are used.

With X-11-ARIMA, the revision between the final and the first annonced seasonally adjusted figures is given by:

$$r_t(0,84) = y^a_{t,t+84} - y^a_{t,t}. \qquad (3.18)$$

Users are generally aware of this type of revision. This revision reflects (a) the innovations introduced by the new observations and (b) the difference between the two filters $a_0(L)$ and $a_{84}(L)$. Assuming that $y^a_{t,t}$ and $y^a_{t,t+84}$ are unbiased estimators of N_t the revision variance can be defined as:

$$E[y^a_{t,t}-y^a_{t,t+84}]^2 = E[(a_0(L)-a_{84}(L))y_t]^2. \qquad (3.19)$$

In the model-based procedures the components N_t and S_t are usually correlated and hence, future values of the observed series can also be used to estimate them. The revision variance is thus defined as the difference between the variances of the two estimators of N_t, namely the first one using the observations up to time t, and the second one using all the observations up to time t+k, for some $k>0$.

3.4.2 Estimation.

3.4.2.1 X-11-ARIMA.

An unbiased estimator of (3.19) is obviously given by:

$$\sigma^2_r(n) = (n-1)^{-1} \sum_{t=1}^{n} [r_t(0,84)-r(0,84)]^2$$

where $\sigma^2_r(n)$ is the sample variance of the revisions, $r(0,84)$ is the sample mean of the revisions and $n>1$. Since $y^a_{t,t}$ and $y^a_{t,t+84}$ are assumed to be two unbiased estimators of N_t one would expect that the mean of the revisions be equal to zero. In this case $\sigma^2_r(n)$ reduces to:

$$\sigma^2_r(n) = n^{-1} \sum_{t=1}^{n} r_t(0,84)^2.$$

The subject of the revision variance induced by a change in the filters $a_i(L)$ $i=0,1,...,m$ used to seasonally adjust the data has been thoroughly studied by Dagum (1982a,1982b) and Dagum and Laniel (1987). In fact, equations (2.5) represent a linear system where $y^a_{t,t+i}$ is the convolution of the input y_t and a sequence of weights $a_{i,j}$ called the impulse response function of the filter. The properties of the filter $a_i(L)$ can be described by its Fourier transform called the frequency response function:

$$\Gamma^{(i)}(\omega) = \sum_{j=-i}^{m} a_{i,j}\exp(-i2\pi\omega j) \qquad 0<=\omega<=.5,$$

where ω is the frequency in cycle per time period. $\Gamma(\omega)$ fully describes the effect of the linear filter on the given input. In general, the frequency response function may be expressed in polar form by:

$$\Gamma(\omega) = A(\omega) + iB(\omega) = G(\omega)\exp(i\phi(\omega))$$

where $G(\omega) = (A^2(\omega) + B^2(\omega))^{1/2}$ is called the gain of the filter and $\phi(\omega)=\arctan(B(\omega)/A(\omega))$ is called the phase shift of the filter and is expressed in radians.

One way of computing a robust measure of variance is to assume that the input series is white noise. In this case, it can be shown that the total revision variance is given by:

$$R^{(84,0)} = [\ 2 \int_0^{1/2} ||\Gamma^{(84)}(\omega) - \Gamma^{(0)}(\omega)||\ d\omega\ ]^{1/2}$$

which is the average distance of the asymmetric filter $a_0(L)$ to $a_{84}(L)$ over all frequencies.

3.4.2.2 Model-based methods.

In section 3.1.2 the estimate x_t of the state vector X_t is made based on the noisy measurement set $\{y_1,\ldots,y_T\}$. There is no need for a delay between the receipt of the last measurement y_{T-1} or y_T and production of the estimate $x_{T/T-1}$ and x_T. However, one can allow a delay of $N>0$ time units, during which $y_{T+1},\ldots,y_{T+N}$ appear, and use all the measurements available to produce the estimate $x_{T/T+N}$ of X_T by: $x_{T/T+N} = E(X_T|y_1,\ldots,y_{T+N})$.

The estimate $x_{T/T+N}$ is called a smoothed estimate. Any estimator producing a smoothed estimate is called a smoother.

Because more measurements are used in producing $x_{T/T+N}$ one expects the estimate to be more accurate. Further, the greater the delay, the greater the increase in complexity. Thus it is important to examine the trade-offs between the delay, improvement in performance and estimator complexity.

Clearly, it is unnecessary to construct estimators which make available the estimates $x_{t/t+k}$ for all t and k. Three types of smoothing problems exist. Fixed-point smoothing is concerned with obtaining $x_{t/t+k}$ for fixed t and all k. Fixed-lag smoothing is concerned whith obtaining $x_{t/t+N}$ for fixed N and all t. And finally, fixed-interval smoothing is concerned whith obtaining $x_{t/M}$ for fixed M and all t in the interval $0<=t<=M$.

In the context of seasonal adjustment it is clear that fixed-point smoothing is irrelevant. Either fixed-lag or fixed-interval smoothing should be considered, the choice of one over the other being dictated by operational problems. For example, in fixed-lag smoothing the lag could be fixed to 1,3,12 or 36 months so that every month the estimate 36 months ago would become final and last year, 3 months ago and previous month estimates would be revised; in fixed-interval smoothing at the end of each year, all the estimates in a specific year would be revised.

The fixed-lag smoothing equations are given by:

$$x_{t/t+N} = x_{t/t+N-1} + K(N+1)_t y_t{}^i$$

where $x_{0/-1}$ is equal to the initial state vector and $x_{0/N-1}=0$. The gain matrices are given by:

$$K(N+1)_t = P(N)_{t/t-1}H[H'P_{t/t-1}H + R]^{-1}$$

where

$$P(N+1)_{t+1/t} = P(N)_{t/t-1}[F-K_tH']'$$

initiated by $P(0)_{t/t-1} = P_{t/t-1}$. The covariance matrix turn out to be:

$$P_{t/t+N} = P_{t/t+N-1} - P(N)_{t/t-1}H[K(N+1)_t]'.$$

The fixed-interval smoothing equations are given by:

$$x_{t/M} = x_t + P_t^*[x_{t+1/M}-x_{t+1/t}]$$

where

$$P_t^* = P_tF'P_{t+1/t}^{-1}$$

with covariance matrix:

$$P_{t/M} = P_t + P_t^*[P_{t+1/M}- P_{t+1/t}]P_t^{*\prime}$$

Now, since the revision is independent of the error in the final estimate we can express the revision variance as the variance of the filetred value (section 3.1.2) minus the variance of the final value obtained through smoothing.

Fortunately, there is no need to evaluate the revision variance in practice since it is already included in P_t at the production of the filtered estimate.

3.5 Parameters estimation variance.

3.5.1 Definition.

The optimal procedures described in sections 3.1.2 and 3.2.2 to estimate the signal and sampling variances assume that the parameters are known. In practice, however, the parameters are just estimated from the observations and then assumed to be known.

In general unbiased estimators are used. This results in unbiased estimators of the state vector through the well known formula $E_m=E_mE_p$ where E_m denotes the expectation under the model and E_p denotes the expectation under the parameters estimation. However all the variances are underestimated since $V_m=V_mE_p + E_mV_p$. The second term E_mV_p is the contribution to the variance arising from the fact that the parameters are not known . This is positive since it is the expectation of a variance which is always positive.

In X-11-ARIMA there are two sets of parameters. The first set of parameters are those of the ARIMA model used to extend the series on both ends and the second set of parameters are those used in the trading-day adjustment if the latter adjustment is needed. X-11 has no ARIMA parameters.

3.5.2 Estimation.

3.5.2.1 ARIMA parameters.

The ARIMA component of X-11-ARIMA consist of fitting an ARIMA model to the series and using this model extending the series one year on both ends. This is done to reduce the revisions of the seasonally adjusted data by using an asymmetric filter closer to the symmetric filter (ref 3.3). Even if the revision variance is reduced, there is an additionnal source of uncertainty that comes from the fact that the unknown ARIMA parameters are estimated.

A valuable question is whether the ARIMA model should be on the estimated values y_t or the true population values Y_t. The modelling of y_t assumes that there is a stochastic model on $y_t = Y_t+u_t$. This could be an acceptable hypothesis for the classical survey samplers, the idea being to forecast the future estimates at times T+1,...,T+12 and the previous estimates at times -11,...,0. Let E_s and V_s denote the sampling expectation and variance and E_p and V_p the expectation and variance under the estimation method for the ARIMA parameters, then:

$$V_s(y^{\star}_{T+i}) = E_s[V_p(y^{\star}_{T+i})] + V_s[E_p(y^{\star}_{T+i})] \qquad (3.20)$$
$$= E_s[V_p(y^{\star}_{T+i})] + V_s[y_{T+i}]$$

where $y^{\star}_{T+i}$ is the forecast value of y_{T+i} $i=1,...,12$. Similar results apply for the previous values. Equations (3.20) say that the sampling variance of the forecasted values are greater than the sampling variance of y_{T+i}. This ensures that the variance coming from the ARIMA parameters estimation is taken into account in the estimation of the sampling variance of the seasonally adjusted data.

The modeling of Y_t instead of y_t assumes that Y_t is a stochastic process. For classical survey samplers such hypothesis is usually unacceptable. In fact, samplers are usually concerned with y_t -Y_t at a single point in time, in which case there is no reason to assume a stochastic model on Y_t. However, models are used as a basis for composite estimation and in the model-based school. In the latter case, the increase in variance is given by E_mV_p as previously mentionned.

3.5.2.2 Trading day parameters.

In X-11-ARIMA the trading component is treated as a deterministic component. If we assume that the original series is affected by trading day then the process of removing the component can be summarized as follows: i) a preliminary estimation of the trading day component is done on the original series from which a preliminary estimation of the trend-cycle and seasonality has been removed; and ii) the preliminary estimation of the trading day component is removed from the original series and a final estimation of the trading day component is performed in the same manner as the preliminary estimation.

In the estimation of the trading day component it is assumed that the input series, z_t, is composed of a trading day component, say $x_t'\beta$, and an irregular component i_t leading to the model:

$$z_t = x_t'\beta + i_t$$

where x_t' is the vector of regressors, β is the vector of trading-day coefficients and i_t is assumed to be $NID(0,\sigma^2)$. Such an assumption is obviously false, nevertheless it is a good approximation and it was the only feasible assumption when the adjustment was developed, see Young (1968). The estimation of the parameters β is obtained by using regression analysis and is not discussed here, see Dagum(1980).

If β^* is the estimate of β, then the final estimation of the seasonally adjusted data is done on the transformed series $w_t = y_t - x_t'\beta^*$, where the sampling variance of w_t is given by:

$$\begin{aligned} V_s(w_t) &= E_s[V_p(w_t)] + V_s[E_p(w_t)] \\ &= E_s[x_t'V_p(\beta^*)x_t'] + V_s[y_t] \end{aligned} \qquad (3.21)$$

Equations (3.21) shows that the sampling variance of the series adjusted for trading variation is increased by the first term of the right hand side. Similar results hold true when the sampling variance is replaced by the model variance.

4. CONCLUSION.

This report presented various sources of uncertainty in seasonal adjustment, with the estimation method to evaluate them.

It is hoped that a definition of the variance of the seasonally adjusted data can be found. The signal variance is needed to construct confidence intervals around the nonseasonal component and eventual forecasts. The sampling variance would allow a comparison with the unadjusted figure. The suboptimal seasonal adjustment procedure variance concerns more the method used than the published figure. Similarly for the revison variance associated with X-11-ARIMA. In the latter two cases, the variances provide measures of quality of X-11-ARIMA. As shown in this report (section 3.4.2.2) the revision variance for the model-based procedures is easily obtained. Finally, the parameters estimation variance is needed to give an indication on the increase in the variance (either signal or sampling) arising from the fact that the parameters are not known.

BIBLIOGRAPHY

Anderson,B.D.O. and Moore,J.B.,(1975), Optimal Filtering, Prentice-Hall, Englewood Cliffs, N.J.

Armstrong,J.B. and Gray,B.G.,(1986), "Variance Estimation For Seasonally Adjusted Survey Data", Proceedings of the ASA.

Bell,W.R. and Hillmer,S.C.,(1983), "Modeling Time Series With Calendar Variation", JASA, 78, 526-534.

Bell,W.R. and Hillmer,S.C.,(1987), "Time Series Methods For Survey Estimation", SRD Research Report Number: Census/SRD/RR-87/20, Bureau of the Census.

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

Burridge,P. and Wallis, K.F.,(1984), "Unobserved-Components Models For Seasonal Adjustment Filters", JBES, 2, 350-359.

Burridge,P. and Wallis, K.F.,(1985), "Calculating The Variance Of Seasonally Adjusted Series", JASA, 80, 541-552.

Cleveland,W.P. and Tiao,G.C.,(1976), "Decompositon Of Seasonal Time Series: A Model For The Census X-11 Program", JASA, 71, 581-587.

Dagum,E.B.,(1980), "The X-11-ARIMA Seasonal Adjustment Method", Statistics Canada, Catalogue No. 12-564E.

Dagum,E.B.,(1982a), "Revision Of Time Varying Seasonal Filters", Journal of Forecasting, 1, 173-187.

Dagum,E.B.,(1982b), "The Effects Of Asymmetric Filters on Seasonal Factor Revisions", JASA, 77, 732-738.

Dagum,E.B.,(1987), "Structural and Reduced-Form Approaches of ARIMA Model Based Seasonal Adjustment Methods", Working Paper No. TSRA-87-007E, Statistics Canada.

Dagum,E.B. and Laniel,N.,(1987), "Revisions of Trend-Cycle Estimators Of Moving Average Seasonal Adjustment Methods", JBES, 5, 177-189.

Harvey,A.C. and Todd,P.H.J.,(1983), "Forcasting Economic Time Series With Structural And Box-Jenkins Models: A Case Study", JBES, 1, 299-315.

Harvey,A.C.,(1984), "A Unified View Of Statistical Forcasting Procedures", Journal of Forcasting, 3, 245-275.

Hausman, J.A. and Watson, M.W., (1985), "Errors In Variables And Seasonal Adjustment Procedures", JASA, 80, 531-540.

Hillmer, S.C. and Tiao, G.C., (1982), "An ARIMA-Model-Based Approach To Seasonal Adjustment", JASA, 77, 63-70.

Kovar, J., (1985), "Variance Estimation Of Nonlinear Statistics In Stratified Samples", Working Paper No. BSMD-85-052E, Statistics Canada.

Kovar, J., (1987), "Variance Estimation Of Medians In Stratified Samples", Working Paper No. BSMD-87-004E, Statistics Canada.

Laniel, N., (1987), "Variance For a Rotating Samples From a Changing Population", Proceedings of the ASA, (to appear).

Quenneville, B. and Srinath, K.P., (1984), "Estimation Of Variances Of Averages Based On Overlapping Samples In Repeated Surveys", Proceedings of the ASA.

Shiskin, J., Young, A.H. and Musgrave, J.C., (1967), "The X-11 Variant Of The Census Method 11 Seasonal Adjustment Program", Technical Paper No. 15, U.S. Department of Commerce, Bureau of Economic Analysis.

Statistics Canada, (1985), Quality Guidelines, Ottawa.

Tam, S.M., (1984), "On Covariances From Overlapping Samples", The American Statistician, 38, 288-289.

Wallis, K.F., (1982), "Seasonal Adjustment And Revision Of Current Data: Linear Filters For The X-11-Method", JRSS-A, 145, 74-85.

Wolter, K.M. and Monsour, N.J., (1981), "On The Problem Of Variance Estimation For a Deseasonalized Series", Current Topics In Survey Sampling, Academic Press, 367-403.

Young, A.H., (1968), "Linear Approximations to the Census and BLS Seasonal Adjustment Methods", JASA, 63, 445-471.