Time Series Research and Analysis Division

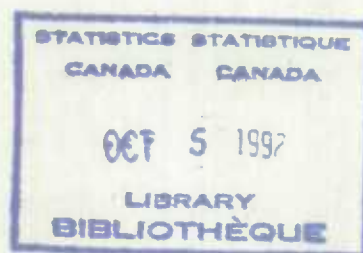# Division de Recherche et Analyse des Chroniques
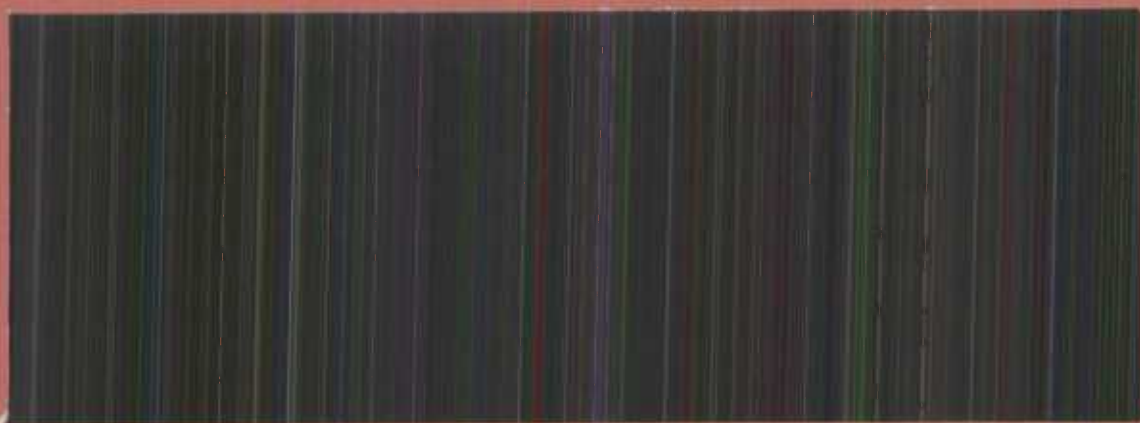
Canadä

SEASONALITY*
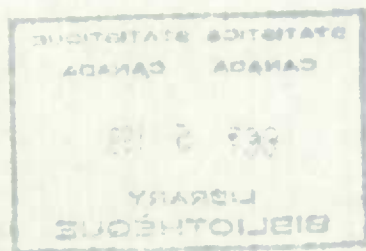
by

Estela Bee Dagum

*Entry to be published in the Encyclopedia of Statistical
Sciences (S. Kotz and N. Johnson, Editors) Vol. 6.

Estela Bee Dagum
Statistics Canada

Seasonality

6

1.

1 - Causes and Characteristics of Seasonality

A great deal of information on socio-economic activity occurs in the form of time series where observations are dependent and the nature of this dependence is of interest in itself. A time series is a sequence of observations ordered in time, say X_1, X_2, ..., X_t, ..., X_T; the interval between dates t and t+1 being fixed and constant throughout. The observations are generally compiled for consecutive periods, such as days, weeks, months, quarters or years.

The analysis of time series has long distinguished different types of evolutions, which may possibly be combined, namely: (1) the trend (2) the cycle, (3) the seasonal variations and (4) the irregular fluctuations.

The trend is a slow variation over a long period of years. It is generally associated with the structural causes of the phenomenon in question. In some cases, the trend shows a steady growth, in others, it may move downward as well as upward.

The cycle is a quasi-periodic oscillation characterized by alternating periods of expansion and contraction. In most cases is related to fluctuations in economic activity.

The seasonal variations represent the effect of climatic and institutional events that repeat more or less regularly each year.

The irregular fluctuations represent unforeseeable

movements related to events of all kinds. They have a stable random appearance but, in some cases, extreme values may be present. These extreme values, or outliers, have identifiable causes, such as strikes, or floods, and, therefore, can be distinguished from the much smaller irregular variations.

The decomposition of an observed time series in several evolutionary processes is basic for time series analysts, the study of the cycle and economic growth, and the study of seasonality. The feasibility of time series decomposition was proved by Wold's famous theorem [25] which states that if a time series X_t is stationary to the second order (i.e. its mean does not depend on time and its autocovariance function depends only on the time lag) then it can be uniquely represented as the sum of two mutually uncorrelated processes, one an infinite moving average n_t and the other, a deterministic process ξ_t, where the future evolution of the realization can be determined completely provided that all its previous values are known. Hence,

$$X_t = n_t + \xi_t \tag{1}$$

where,

$$n_t = \sum_{j=0}^{\infty} \beta_j U_{t-j}; \quad \Sigma \beta_j^2 < \infty \tag{2}$$

$$E(U_t \, U_s) = \begin{cases} \sigma^2 & \text{if } t=s \\ o & \text{if } t \neq s \end{cases} \tag{3}$$

and

3.

$$\xi_t = \sum_{j=1}^{\infty} \alpha_j \ \xi_{t-j} \qquad (3)$$

$$E(\xi_t \ \eta_s) = 0 \text{ for all } t, s.$$

There are several decomposition models that have been as-
sumed for time series analysis. The more often applied are
the additive and the multiplicative models. That is,

$$X_t = C_t + S_t + U_t \quad \text{(additive model)} \qquad (4)$$

$$X_t = C_t \ S_t \ U_t \qquad \text{(multiplicative model)} \quad (5)$$

where C_t is the trend-cycle, S_t is the seasonal component
and U_t is the irregular component (in the model (5), S_t and
U_t are expressed as percentages). The choice of decomposi-
tion model for a given series depends on whether the compo-
nents are assumed to be <u>independent</u> (additive model) or
dependent (multiplicative model) of one another. There are
few cases, however, where a mixed decomposition model in
which components are multiplicatively and additively related
may be more adequate as, for example, the one discussed in
[10],

$$X_t = C_t \ (1+S_t) + U_t \qquad (6)$$

Among the several types of fluctuations, those due to sea-
sonality has long been recognized. The organization of so-
ciety, means of production and communication; social and
religious events have been strongly conditioned by both

4.

climatic and conventional seasons. Seasonal variations in
agriculture, the low level of winter construction and high
pre-Christmas retail sales are all well known.

It is generally accepted that there are four causes of
seasonality in economic time series; namely, the calendar,
timing decisions, the weather and expectations. Most of the
seasonal movements in exports and imports are caused by
variations in the number of working days in a month. Deci-
sions on the timing of such events as school variations,
payment of company dividends or the ending of the tax year
cause seasonal variations. Changes in temperature, rainfall
and other climatic factors directly affect agricultural
production, transportation, construction and indirectly,
many other activities. The expectations of seasonality in
one economic activity will lead to actual seasonality in
another; for example, toy production will increase in expec-
tation of the Christmas sales peak. While the causes of
seasonality are generally exogeneous to the economic system,
human intervention can modify their extent and nature. For
example, seasonal variations in the automobile industry are
affected by manufacturers' decisions regarding the extent of
model changeover each year.

Another main feature of seasonality is that the pheno-
menon repeats with certain regularity every year but it may
evolve. Many reasons can produce changes in seasonal

5.

patterns. A decline in the importance of the primary sector
in the gross national product modifies seasonal patterns in
the economy as a whole, as does a change in the geographical
distribution of industry in a country extending over several
climatic zones. Changes in technology alter the importance
of climatic factors. For most economic series, an evolving
seasonality is more the rule than the exception. The assump-
tion of stable seasonality, that is, of seasonality that
repeats exactly every year is good for a few series only.
Depending on the causes of seasonal variations, their
patterns can change slowly or rapidly, gradually or abruptly,
in a deterministic manner or in a stochastic manner.

2 - Seasonal Models

The simplest and often studied seasonal model assumes
that the generating process of seasonality can be represen-
ted by strictly periodic functions of period $\underline{n}$ (e.g. 12 for
monthly data, 4 for quarterly data).

For monthly series, the problem is to estimate 12 con-
stants, one for each month, that sum to zero. That is,

$$S_t \begin{cases} a_k & \text{for } t=k \text{ or } t-k \text{ divisible by } 12 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

$$\sum_{k=1}^{12} a_k = 0$$

The model (7) is more often shown unders its equivalent form

6.

in the frequency domain as follows,

$$S_t = \sum_{j=1}^{6} (\alpha_j \cos 2\pi\lambda_j t + \beta_j \sin 2\pi\lambda_j t), \quad \lambda_j = j/12 \qquad (8)$$

Where the λ_j for j=1,2,3,...,6 are the seasonal frequencies
which correspond to cyclical periods of 12,6,4,3,2.4 and 2
months, respectively.

The seasonal model (8) is said to be _deterministic_ if
α_j and β_j are constants, and stochastic if α_j and β_j are
purely random variables with zero mean and

$$E(\alpha_j\alpha_k) = E(\beta_j\beta_k) = \begin{cases} \sigma_j^2 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \qquad (9)$$

$$E(\alpha_j\beta_k) = 0 \qquad \text{for all } j \text{ and } k.$$

Series that follow the model (8) have spectra which are
zero except for six spectral lines of height α_j^2 at each λ_j.
All the spectral power is concentrated at the seasonal
frequencies. Although this never occurs with real data,
series that show very narrow seasonal peaks can be well ap-
proximated by this kind of model.

For most economic time series, however, seasonality is
not stable but it changes gradually. In such cases, a more
general seasonal model may be expressed as follows,

$$S_t = \sum_{j=1}^{6} (\alpha_{jt} \cos 2\pi\lambda_j t + \beta_{jt} \sin 2\pi\lambda_{jt}); \quad \lambda_j = j/12 \quad (10)$$

Where α_{jt} and β_{jt} are time-varying parameters. They can
either be low degree polynomials or stochastic processes

7.

with spectra dominated by low frequency components. For
example, if the seasonal amplitudes follow a stationary
autoregressive process of order one, then,

$$\alpha_{jt} = \rho\alpha_{jt-1} + U_t \qquad (11)$$

$$\beta_{jt} = \rho\beta_{jt-1} + V_t$$

where $|\rho|<1$ and U_t, V_t are purely random mutually uncorrela-
ted variables with zero mean and constant variance.

If a series follows the seasonal model (10) where α_{jt}
and β_{jt} are assumed to be stationary stochastic processes,
its spectrum will be characterized by peaks at the seasonal
frequency bands $\lambda_s(\delta)$ of width δ and where the value of δ
depends on the rate at which α_{jt} and β_{jt} change.

The broader the bands, the less regular will be the
seasonal variations. The set of seasonal frequency bands
may be defined by,

$$\lambda_s(\delta)=\{\lambda_s|\lambda_s \in (\lambda_j-\delta, \lambda_j+\delta),j=1,2,3,4,5,6; \text{where } \lambda_6+\delta =\pi\} \quad (12)$$

3 - Seasonal Adjustment Methods

The fact that the causes of seasonality are generally
exogeneous to the economic system is the main reason for
the removal of seasonal variations from an observed series
to produce a seasonally adjusted series. The adjusted
series thus reflects only variations attributed to the trend,
the cycle and the irregulars. The removal of seasonality

8.

from a time series, however, does not indicate how the series would have evolved had there been no seasonal variations; rather, it shows more clearly the trend-cycle abstracted from seasonality.

The estimation of seasonal variations has always posed a serious problem to statisticians because the phenomenon varies and is not directly observable.

The majority of the seasonal adjustment methods developed thus far are based on univariate time series models where the estimation of the seasonal variations is made in a simple and mechanical manner and not based on a causal explanation of the phenomenon in question.

The majority of the methods for the seasonal adjustment of economic time series fall into two broad categories. The first derives from general regressions and linear estimation theory, the second depends mainly on the application of moving averages or linear smoothing filters. Very few exceptions do not fall into this broad classification, being probably the most known, the S.A.B.L. method [4] which uses a combination of both moving means and medians.

3.1 Regression Methods

Most work on regression methods for seasonal adjustment is based on the assumption that the systematic part of a time series can be approximated closely by simple functions

of time over the _entire_ span of the series. In general, two
types of functions of time are considered. One is a poly-
nomial of fairly low degree that fulfills the assumption
that the economic phenomenon moves slowly, smoothly and
progressively through time (the trend). The other is a
linear combination of periodic functions representing oscil-
lations that affect the total variation of the series (the
cycle and seasonality). For example, a simple regression
model that assumes a cubic trend and a stable deterministic
seasonality is,

$$X_t = C_t + S_t + U_t \qquad\qquad (13)$$

where

$$C_t = \sum_{i=0}^{3} \alpha_i \, t^i \qquad\qquad (14)$$

$$S_t = \sum_{j=1}^{12} \beta_j D_{jt}; \quad \sum_{j=1}^{12} \beta_j = 0 \qquad\qquad (15)$$

The α_i and β_j are the unknown parameters and D_{jt} are seasonal
dummy variables which take the value 1 when the _t_th observa-
tion concerns the jth months and zero in all other cases.
If U_t is assumed to be purely random with zero mean and
finite variance the estimation of the parameters can then
be made by least squares.

Many variations are possible concerning the specifica-
tion of the components. Thus, in the representation (15) it

10.

may be assumed that the seasonal component changes through time. The coefficient β_j is then replaced by $\beta_j + \gamma_j t$, the γ_j being restricted by the condition that their sum is zero. In other cases, a multiplicative decomposition model may be found to be more appropriate than the additive model (13), in which case, the determination of the trend and the seasonal component can be carried out as before by means of a regression on the logarithms of the X_t.

Major contributions to the development of regression models for seasonal adjustment were made in the sixties, particularly in [13], [14], [16] and [18]. To overcome the limitation of using a global representation of the trend-cycle, [11] and [22] used local polynomials (spline functions) for successive short segment of series. These regression models, however, still imply a deterministic behaviour of the time series components. More recent studies have considered the possibility of a stochastic behaviour of the components by developing mixed models [20] or regression models with time varying parameters [12].

Regression methods have been seldom used by statistical agencies. The main reasons for this are: (1) that the seasonally adjusted series are wholly revised when new observations are added; and (2) that all the regression models developed until very recently, implied a deterministic

11.

behaviour of the components.

3.2   Moving Average Methods

The majority of the seasonal adjustment methods applied
by statistical bureaus belong to the class of moving average
methods which make the assumption that although the systema-
tic part of a time series is a smooth function, it cannot
be approximated well by simple mathematical functions over
the entire range.

Since moving averages are linear transformations (see
entry in this Encyclopedia) they possess the properties of
scale preservation and additivity. Furthermore, they have
the time invariance property which is not shared by the re-
gression methods. The time invariance property means that
if two inputs X_t and $X_{t+\tau}$ to the moving average are the same
except for the time displacement τ, then the outputs are
also the same except for their time displacement. In other
words, the moving average or linear filter responds always
in the same manner.

Methods based on moving averages techniques assume that
the trend-cycle and seasonal components change through time
in a stochastic manner. The majority of the methods that
belong to this class are mainly descriptive non-parametric
procedures in the sense that they lack explicit parametric
models for each unobserved component.

In most recent years, however, several attempts have been made to develop model-based procedures where univariate statistical models are explicitly assumed for each component. The explicit models mainly belong to the Gaussian ARIMA (autoregressive integrated moving average) type developed by [2] or to variations of it as developed in [3], [15] and [19]. Other types of models which are not ARIMA have also been studied, for example, in [1]. All these new approaches are still in a developmental stage and the majority of the moving average procedures officially adopted by statistical bureaus belong to the non-parametric type (see [17]). Among the latter, the Method II-X-11 variant [21] and the X-11-ARIMA [5] are the most widely applied. The X-11-ARIMA was developed to produce more accurate estimates of current seasonally adjusted series when seasonality changes rapidly and in a stochastic manner; characteristics often found in main economic indicators.

These two methods follow an iterative estimation procedure where the trend-cycle is estimated first, the seasonal component next and the irregular is derived as a residual. The properties of the combined linear filters applied to obtain a seasonal estimate have been analysed in [23], [24] and [26] for the X-11-variant and in [6] for the X-11-ARIMA.

It is inherent to all linear smoothing procedures that the end observations cannot be smoothed with the same set of

13.

symmetric filters as applied to central observations.

Because of this, the estimates for current observations must

be revised as more data are incorporated into the series.

In the context of the X-11 and X-11-ARIMA, this means that

the first and last three and a half years of a new series

will be revised because their symmetric filters require

seven years of data to produce a central estimate. The

revisions due to differences in the moving averages applied

to the same observations as it changes its positions rela-

tive to the end of the series have been studied extensively

for both methods in [7], [8] and [9].

REFERENCES

[1] Akaike, H. and Ishiguro, M. (1980). BAYSEA, A Bayesian
     Seasonal Adjustment Procedure. Monograph 13, Inst.
     Statist. Math. Tokyo.

[2] Box, G.E.P. and Jenkins, G.M. (1970). Time Series
     Analysis: Forecasting and Control, Holden Day, San
     Francisco.

[3] Burman, J.P. (1980). J.R. Statist. Soc. A, 143, 321-27.

[4] Cleveland, W.S., Dunn, D.M. and Terpenning, I. (1978).
     In Seasonal Analysis of Economic Time Series A. Zellner,
     ed. U.S. Gov. Print. Off. Washington, D.C. 201-231.

[5] Dagum, E.B. (1980). The X-11-ARIMA Seasonal Adjustment
     Method, Statistics Canada, Ottawa.

[6] Dagum, E.B. (1982). Canad. J. Statist. (forthcoming).

[7] Dagum, E.B. (1982). J. Forecasting, 1, 173-187.

[8] Dagum, E.B. (1982). J. Amer. Statist. Ass. 77, 732-738.

14.

[9]   Dagum, E.B. (1982) Proc. Bus. Ec. Sect. Amer. Statist.
      Ass.

[10] Durbin, J. and Kenny, P.B. In Seasonal Analysis of
      Economic Time Series, A. Zellner, ed. U.S. Gov. Print.
      Off. Washington, D.C., 173-187.

[11] Duvall, R.M. (1966). J. Amer. Statist. Ass. 61,
      152-165.

[12] Gersovitz, M. and MacKinnon, J.G. (1978). J. Amer.
      Statist. Ass. 73, 264-273.

[13] Hannan, E.J. (1960). Aust. J. Statist. 2, 1-15.

[14] Henshaw, R.C. (1966). Econometrica, 34, 381-395.

[15] Hillmer, S.G. and Tiao, G.C. (1982). J. Amer. Statist.
      Ass. 77, 63-70.

[16] Jorgenson, D.W. (1964). J. Amer. Statist. Ass. 59,
      681-723.

[17] Kuiper, J. (1978). In Seasonal Analysis of Economic
      Time Series. A. Zellner, ed. U.S. Gov. Print. Off.
      Washington, D.C., 59-76.

[18] Lovell, M.C. (1963). J. Amer. Statist. Ass. 58, 993-
      1010.

[19] Nerlove, M., Grether, D.M. and Carvalho, J.L. (1979).
      Analysis of Economic Time Series. Academic Press,
      New York.

[20] Pierce, D. (1978). In Seasonal Analysis of Economic
      Time Series. A. Zellner, ed. U.S. Gov. Print. Off.
      Washington, D.C. 242-269.

[21] Shiskin, J. Young, A.H. and Musgrave, J.C. (1967).
      The X-11-Variant of Census Method II Seasonal Adjust-
      ment Program, U.S. Bureau of Census, Washington, D.C.

[22] Stephenson, J.A. and Farr, H.T. (1972). J. Amer.
      Statist. Ass. 67, 37-45.

[23] Wallis, K.F. (1974). J. Amer. Statist. Ass. 69, 18-31.

[24] Wallis, K.F. (1982). J.R. Statist. Soc. A, 145, 74-85.

15.

[25] Wold, H. (1938). A Study in the Analysis of
      Stationary Time Series, Almquist and Wiksell, Stock-
      holm (2nd ed. 1954).

[26] Young, A.H. (1968). J. Amer. Statist. Ass. 63, 445-
      457.