Lecture Notes in Statistics

158

Dominique Ladiray
Benoît Quenneville

Seasonal Adjustment with the X-11 Method

Springer

Lecture Notes in Statistics **158**

Edited by P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg,
I. Olkin, N. Wermuth, S. Zeger

Springer Science+Business Media, LLC

Dominique Ladiray
Benoît Quenneville

Seasonal Adjustment with the X-11 Method

Springer

Dominique Ladiray
EUROSTAT, BECH E4/813
Bâtiment Jean Monnet
Rue Alcide de Gasperi
Luxemborg, L-2920
Belgium

Benoît Quenneville
Statistics Canada
120 Parkdale Avenue
3G-RHC-BSMD
Ottawa, Ontario K1A 0T6
Canada

Library of Congress Cataloging-in-Publication Data

Ladiray, Dominique.
Seasonal adjustment with the X-11 method / Dominique Ladiray, Benoît Quenneville.
p. cm. — (Lecture notes in statistics; 158)
Includes bibliographical references and index.
ISBN 978-0-387-95171-3 ISBN 978-1-4613-0175-2 (eBook)
DOI 10.1007/978-1-4613-0175-2
1.Seasonal variations (Economics)—Statistical methods. I. Quenneville, Benoît.
II. Title. III. Lecture notes in statistics (Springer-Verlag); v. 158.

HB137 .L33 2001
330'.01'5195—dc21 00-053769

Printed on acid-free paper.

Camera-ready copy provided by the authors.

9 8 7 6 5 4 3 2 1

ISBN 978-0-387-95171-3 SPIN 10788799

To Julius Shiskin, Allan Young, and John Musgrave

Preface

The authors, Dominique Ladiray and Benoît Quenneville, provide a unique and comprehensive review of the X-11 Method of seasonal adjustment. They review the original X-11 Method developed at the US Bureau of the Census in the mid-1960's, the X-11 core of the X-11-ARIMA Method developed at Statistics Canada in the 1970's, and the X-11 module in the X-12-ARIMA Method developed more recently at the Bureau of the Census. The review will prove extremely useful to anyone working in the field of seasonal adjustment who wants to understand the X-11 Method and how it fits into the broader picture of seasonal adjustment.

What the authors designate as the X-11 Method was originally designated the X-11 Variant of the Census Method II Seasonal Adjustment Program. It was the culmination of the pioneering work undertaken at the Bureau of the Census by Julius Shiskin in the 1950's. Shiskin introduced the Census Method I Seasonal Adjustment Program in 1954 and soon followed it with the introduction of Method II in 1957.

In Method I, seasonal factors were estimated by taking a moving average of the seasonal-irregular ratios for each month. The seasonal-irregular ratios were ratios of the original series to a centered 12-month moving average that was assumed to represent the trend-cycle component. An alternative seasonal adjustment was also calculated by substituting a five-month moving average of the first seasonally adjusted series for the centered 12-month moving average. The most important improvements in Method II were the following.

1. The moving averages were extended to the ends of the series by means of asymmetrical sets of end weights;

2. The five-month moving average used in the second round of Method I was replaced with a weighted 15-month moving average that provided a more flexible and smooth curve for the trend-cycle component. In Method I the seasonal factors in the second round were considered alternatives that might prove useful for some series. In Method II, there was no question; the second round seasonal factors were considered superior;

3. Extreme values in the seasonal-irregular ratios for each month were identified and modified before seasonal factors were calculated.

Method II was one of the first uses of the electronic computer for large scale processing of economic data. The program was run on the two recently installed Univac computers at the Bureau of the Census. Processing time for a series of 8 to 10 years of monthly data was about seven minutes.

In designing Method II, Shiskin was familiar with the graphical method of seasonal adjustment developed before World War II at the Federal Reserve Board of Governors. This method started with a 12-term moving average of the original data, modified where necessary to better represent the trend-cycle component. Then, to estimate the seasonal factors, free-hand curves were drawn through the seasonal-irregular ratios for each month. In drawing these curves the analyst would discount the effect of values considered to be extreme based on knowledge of the industry or process, coupled with the appearance of the data. Method II may be viewed as an effort to instruct a computer to duplicate the procedure that previously had been applied by skilled professionals.

I once asked a long time employee of the National Income Division of the US Bureau of Economic Analysis how the division managed each year to update and revise seasonal adjustment factors for the GNP data before electronic computers were available. His reply was "we kept the statistical clerks all night." These "all nighters" came at the end of a two or three month period of additional work on weekends and evenings. The annual revisions of the national accounts by the BEA are still major undertakings but staff-wide "all nighters" are no longer necessary. Today they are computed in seconds on desktop computers that are rarely used beyond regular daytime hours.

Part of the explanation for the success of Method II was that it was easy to use. It included graphical output of the several components of the original series that helped users understand the procedure and assess the results. The procedure was relatively robust and did not require a strong background in mathematical statistics.

There are two uses for seasonal adjustment. One is to estimate seasonal variation for such purposes as production planning and inventory control.

The other is to estimate seasonal variation so as to be able to remove it from the total variation and thereby expose the other types of variation more clearly, particularly that reflecting the business cycle.

Shiskin was motivated by the need for government policy-makers to be provided promptly with assessments of current economic conditions and short-term forecasts. As long as they were largely independent, it made good sense to separate the seasonal variation from the cyclical variation for this purpose. Method II provided a forecast of the seasonal factors for the year ahead so that the seasonal variation could be separated quickly from other types of variation in the estimate for the current month. (It would be some years before concurrent seasonal adjustments would become practical and would be demonstrated as better than once-a-year forecasts of upcoming seasonal factors.) Early on, it was apparent that the revisions in year-ahead seasonal factors provided by Method II were often uncomfortably large when an additional year of data was incorporated.

I joined Shiskin at the Bureau in 1958 and John Musgrave came in 1959. We soon began to focus on improving the original Method II as well as assisting Shiskin in the development of business cycle indicators. Successive experimental variants designated X-1 through X-11 were developed. The program for testing rocket-propelled planes inspired the "X" designation. In 1947 Chuck Yeager broke the sound barrier in the X-1, the first of successively numbered experimental planes. In the mid-1960's the X-15, the last of the rocket-propelled planes, was flying at several times the speed of sound.

Two of the experimental variants II, X-3 and X-9, each in turn, replaced the original Method II as the official version at the Census Bureau and some other government agencies. They each included new end weights for the moving averages used in estimating the seasonal factors and provided somewhat smaller revisions.

The X-10 variant was developed in cooperation with Stephen Marris who spent a year at the Census Bureau on leave from the Organization for Economic Cooperation and Development. The X-10 included a family of moving averages from which the most appropriate was selected for each month. This approach further reduced revisions in many series. It became the official OECD method. The approach was not included in the default version of X-11, but was included as an option. A similar approach was included in the default version for the trend-cycle component.

Two other persons deserve mention as well. Following the introduction of Method II by Shiskin, Duane Evans developed a ratio-to-moving average method of seasonal adjustment at the Bureau of Labor Statistics. One of the innovations in this program may be viewed as the forerunner of the expanded treatment of extreme values introduced in the X-11 variant. In addition, shortly after the introduction of Method II, the Census Bureau undertook the development of a parametric approach to seasonal adjustment under the direction of Harry M. Rosenblatt. Neither of these approaches proved a successful competitor to the X-11 variant, partly because revisions of seasonal factors tended to be larger.

In 1965, we certainly did not expect that the X-11 variant of Method II would still be playing an important role in seasonal adjustment 35 years later. It also seemed likely that an approach to measuring the standard errors in the seasonal factors in X-11 would be developed. The authors are to be congratulated for an up-to-date review of the X-11 Method. Perhaps this review will not only help further extend the life of X-11, but will help in the development and widespread use of new methods that bear little resemblance to those of the past.

Allan H. Young Heathsville, Virginia
April 2000

Contents

List of Figures

List of Tables

Introduction

When it comes to seasonal adjustment, the most widely used statistical method is without a doubt that implemented in the Census X-11 software. Developed at the US Bureau of the Census in the 1950's and 1960's, this computer program has undergone numerous modifications and improvements, leading especially to the X-11-ARIMA software packages in 1975 and 1988 (Dagum [19, 20]) and X-12-ARIMA (the first beta version of which is dated 1998, Findley et al. [23]). While these software packages integrate, to varying degrees, parametric methods, and especially the ARIMA models popularized by Box and Jenkins [9], they remain in essence very close to the initial X-11 method, and it is this "core" that will interest us here.

The detractors of Census X-11 have often advanced the "black box" aspect of this software. The lack of an explicit model, the multiplicity of options and output tables, and the lack of easily accessible documentation doubtless largely account for this assessment. Census X-11 is, however, a modern statistical method: it is a robust non-parametric method, applying iterative estimation. Census X-11 stands as an early application of the computer. Its principles are fairly easy to explain. It is true, though, that it was difficult, if not impossible, for even a very well-informed user to reconstruct and explain each Census X-11 output table: minor programming errors and imprecision in the documentation made this task almost insurmountable. These minor programming errors have, for the most part, been eliminated from the new versions of the software, and we have decided to complete this unprecedented work: an in-depth look at an example of seasonal adjustment using the X-11 method.

To this end, we programmed the X-11 method in *Mathematica*© and in *SAS*©, thus verifying each step of the seasonal adjustment and validating step-by-step the results of X-11-ARIMA and X-12-ARIMA. Following the correction of a few errors found in each software, all these programs produce the same results[1].

This work presents a documentation for the seasonal adjustment method implemented in the X-11-based softwares. The monograph serves as a reference work for government agencies and for macroeconomists or other serious users of economic data. Anyone who becomes responsible for seasonal adjustment would benefit from this book. After some historical notes, you will find a presentation of the X-11 methodology and a complete description of the moving averages used in the programs. The following chapter is devoted to the study of moving averages with an emphasis on those used by X-11. Readers will also find a complete example of seasonal adjustment, and have a detailed picture of all the calculations. The linear regression models used for trading-day effects and the process of detecting and correcting extreme values are studied in the example. The estimation of the Easter effect will be dealt with in a separate chapter insofar as the models used in X-11-ARIMA and X-12-ARIMA are appreciably different.

The focus here will be on the X-11 part of the current software, i.e. without reference to *a priori* ARIMA modelling of the series to be seasonally adjusted. The generally minor differences between X-11-ARIMA and X-12-ARIMA relating to the operation of this "central core" will be specified where applicable.

Subsequent references to X-11 refer to the seasonal adjustment method and not to the Census X-11 software.

Acknowledgments

Most of this work was done during Dominique Ladiray's visit to Statistics Canada Time Series Research and Analysis Center from January 1998 to September 1999 as part of the exchange program between Statistics Canada and the French Institut National de la Statistique et des Études Économiques (INSEE). We thank both agencies for giving us the opportunity and the resources to work on this project.

Working on this project has involved digging into the X-11 Fortran computer code. We wish to thank Paul Wong from Statistics Canada and Brian Monsell from the US Bureau of the Census for providing answers to our numerous questions, and for fixing a few small programming errors.

We thank Ketty Attal from INSEE for her contribution to Chapters 1 and 3.

[1] X-11-ARIMA version 2000 and X-12-ARIMA version 0.2.7.

Many people spend time revising preliminary versions of this document. Special thanks are due to Guy Huot and Bernard Lefrançois from Statistics Canada for their thorough review of the initial French version. We wish to thank anonymous translators from Statistics Canada for the English translation. Thanks are due to Marietta Morry, Norma Chhab, Helen Fung and John Higginson from Statistics Canada for reviewing preliminary English versions.

Special thanks are also due to David Findley and Brian Monsell from the US Bureau of the Census, Micheal Baxter from the UK Central Statistical Office and Andrew Sutcliffe from the Australian Bureau of Statistics for stimulating email discussions and their review of the English version.

We are grateful to Allan Young for having written the Preface. We owe him our highest respect for his works.

We wish to thank John Kimmel from Springer-Verlag for having accepted our proposal to publish this work as a book.

Finally, we are obliged to Bernard Lefrançois from Statistics Canada, Marius Ooms from the Free University of Amsterdam, and Stuart Scott from the US Bureau of Labor Statistics for reviewing the final manuscript for Springer-Verlag.

The views expressed in this book are those of the authors and are not necessarily those of Statistics Canada or the Institut National de la Statistique et des Études Économiques. Any remaining errors are our own.

D.Ladiray [2] Paris
B. Quenneville [3] Ottawa
July 2000

[2] Dominique Ladiray is an Administrateur at the INSEE (France), employed with EUROSTAT, Bâtiment Jean Monnet, Rue Alcide de Gasperi, L-2920 Luxembourg. Tel: +352-4301-33339, E-mail: dladiray@hotmail.com.

[3] Benoît Quenneville is a methodologist with the Time Series Research and Analysis Centre, 3G-RHC-BSMD, Statistics Canada, Ottawa, Ontario, Canada, K1A 0T6. Tel: +1-613-951-1605, Fax: +1-613-951-5711, E-mail: quenne@statcan.ca.

1

Brief History of Seasonal Adjustment

It is common today to decompose an observed time series X_t into several components, themselves unobserved, according to a model such as:

$$X_t = T_t + C_t + S_t + I_t,$$

where T_t, C_t, S_t and I_t designate, respectively, the **trend**, the **cycle**, the **seasonality** and the **irregular** components. This is an old idea, and it is doubtless to astronomy that one should turn to find its origin[1].

The more accurate measurements of planetary motion in the 17th century seemed to invalidate Kepler's laws, and the idea that they gave an approximation of the position of a planet rather than its precise position was gradually accepted (Nerlove, Grether and Carvalho [57]). The observed position was then considered to be the sum of the "theoretical" position and of irregular fluctuations. Later, it was noticed that the orbits of the planets changed imperceptibly, and a distinction was made between secular and periodic motions. The unobservable components model was born. The explanation of these periodic or irregular motions excited many mathematicians of the late 18th century and early 19th century, among them Euler, Lagrange and Laplace.

This idea of decomposing a time series subsequently appeared in the works of economists, some of whom did not hesitate to acknowledge that

[1] This section is inspired largely by: Armatte [2], Bell and Hillmer [6], Hylleberg [38], Nerlove, Grether and Carvalho [57].

it came to them directly from astronomy[2] or meteorology[3]. At the same time, developments in mathematics gave researchers means of going beyond mere graphic visualization in analysing time series. Among the important works on the subject should be mentioned, obviously, those of Jean-Baptiste Fourier [25] on the decomposition of a mathematical function into a sum of trigonometric functions, works which would give rise to harmonic analysis and, later, once the notion of stochastic process had been defined, to spectral analysis. For at that time, in economics as in the other sciences, determinism and the search for precise laws to explain all physical, economic, demographic, biological and other phenomena were the order of the day.

The aim of many studies was, then, to reveal "cycles", the study and analysis of which might make it possible to explain and predict economic crises (Armatte [2]). In these conditions, short-term periodic components were of little interest and it was expedient to eliminate them:

> *Every kind of periodic fluctuations, whether daily, weekly, quarterly, or yearly, must be detected and exhibited not only as a subject of study in itself, but because we must ascertain and eliminate such periodic variations before we can correctly exhibit those which are irregular or non-periodic and probably of more interest and importance. (Jevons [39])*

The late 19th and early 20th centuries abounded with publications on the decomposition of economic series, component estimation techniques and definitional elements of components[4]. But it is undoubtedly W.M. Persons [59] who deserves credit for proposing, in 1919, in a single work a "complete" method of decomposition including an attempt to define and formalize unobservable components, a decomposition model and an estimation method. According to Persons, a time series is decomposed into the four types of fluctuations that are still familiar today:

1. *A long-time tendency or secular trend; in many series such as bank clearings or production commodities, this may be termed the growth element;*

[2] Nerlove, Grether, Carvalho [57] cite several examples, including those of Cournot [18] and Jevons [39]. In 1801, the British astronomer William Herschel [33] compared the observed periodicities of Sun's spots with those of the price of wheat.

[3] One can hardly fail to cite the works of the meteorologist Buys-Ballot [11] who, in 1847, studied periodic temperature variations by modelling the trend by a polynomial, seasonality by indicators and implicitly relying on linear regression techniques to estimate the parameters.

[4] For example, in 1905, Lucien March [52] distinguished *"annual changes, polyannual changes (decennial, for example), secular changes, to say nothing of periods of less than a year"* [translation] (cited by Yule [68] and Bell and Hillmer [6]). We can also cite the precursory works of the sisters Maballée [50] attempting to isolate the turning points of a series using the correlogram.

2. *A wavelike or cyclical movement superimposed upon the secular trend; these curves appear to reach their crests during the periods of industrial prosperity and their troughs during periods of industrial depression, their rise and fall constituting the business cycle;*

3. *A seasonal movement within the year with a characteristic shape for each series;*

4. *Residual variation due to developments which affect individual series, or to momentous occurrences such as wars or national catastrophes, which affect a number of series simultaneously.*

These components are then combined according to well-known additive or multiplicative decomposition models:

Additive model: $X_t = T_t + C_t + S_t + I_t$,

Multiplicative model: $X_t = T_t \times C_t \times S_t \times I_t$,

or even: $X_t = T_t \times (1 + C_t) \times (1 + S_t) \times (1 + I_t)$.

Most publications of the day accepted these models and definitions without much discussion, the emphasis being put instead on "seasonal adjustment" or "cycle extraction" techniques per se. Similarly, other concepts were studied and accepted: the idea that seasonality varies over time; the need to take into account all components simultaneously when estimating the seasonal portion; the impossibility of describing trends and cycles as simple deterministic mathematical functions of time; the need to process outlying points—all these ideas giving rise to different estimation methods (Menderhausen [53]).

The works of the time were, however, essentially inspired by two main methods. Here is a brief description of each in the context of a multiplicative model (Armatte [2]).

- In the 1910s, the **relative links** method, developed by Persons [59], was the one favoured by mathematical economists. Its principle consists in calculating, for each monthly value X_t of a series, the ratio X_t/X_{t-1}; creating a frequency table of the values of these ratios for 12 months; then determining the medians, $\{M_i,\ i = 1, \ldots, 12\}$, of these 12 series. These medians are then chained by multiplication, using a base of 100 for January: $S_1 = 100, S_i = M_i S_{i-1}$. Finally, these seasonal factors are corrected by a factor $(S_{13}/S_1)^{1/12}$ such that $S_1 = S_{13} = 100$.

- The second way to determine the seasonal factors is the **moving average**[5] method, applied since 1922 by the Federal Reserve[6] and

[5] A full discussion of moving averages will come later.

[6] The English physicist Poynting [60] is often credited with being the first to use, in 1884, a moving average to eliminate trend and isolate the fluctuations of the series that could then be processed by harmonic analysis.

later popularized by Macaulay [51]. It is based on the calculation of a centered moving average of order 12 to obtain a trend estimate. The ratio between the original data and this trend estimate provides an initial estimate of the seasonal component. To eliminate the irregular component, the medians (or averages) of this initial seasonal component are then calculated for each month. These new factors are then adjusted so that they sum to 1, and the final seasonal factors are thus obtained.

Although very popular, these methods came under much criticism from the theoretical standpoint. Thus, Slutsky [64] and Yule [69] showed that the use of moving averages could introduce artificial cycles into the data. Fisher [24] deplored the fact that *ad hoc* "empirical" methods were used when there were adequate mathematical tools. The most important event of this period was without a doubt the appearance of autoregressive (Yule [69]) and moving average (Slutsky [64]) models for analysing time series; in other words, means whereby mathematical economists could go beyond the traditional deterministic framework by using the first stochastic processes. But it would be many years before these models would enjoy some success in seasonal adjustment.

In the 1930s, seasonal adjustment methods based on regression techniques had varying degrees of success. These techniques were based generally on an additive decomposition of the initial series or of a simple transformation of it, a modelling of the initial series and of each component by simple parametric functions, and the estimation of parameters by ordinary least squares methods. The difficulty of finding a good specification of the model, and in particular the strong assumptions necessary about the unobservable components, doubtless explain why these methods were set aside for a while (Bell and Hillmer [6]).

The development in computers after the Second World War greatly contributed to the spread and improvement of seasonal adjustment methods. Thus, in 1954, Julius Shiskin developed the **Census Method I** at the US Bureau of the Census. This seasonal adjustment technique would be followed by the **Census Method II** in 1957 and eleven experimental versions (X-1, X-2, etc.) culminating in the **X-11 Variant of the Census Method II Seasonal Adjustment Program** in 1965 (Shiskin, Young and Musgrave [63]). Directly inspired by smoothing using moving averages and the works of Macaulay [51], these various versions constituted the first automatic seasonal adjustment methods and **Census X-11** quickly became a standard used around the world. The new computational possibilities facilitated the use of parametric regression techniques to estimate and correct calendar effects (trading-days, statutory holidays, annual leave, etc.). Furthermore, the automatic processing of these effects, based on the works of Young [67], was integrated into the Census X-11 software.

At the same time, significant advances were made in the parametric modelling of time series and in spectral analysis, owing essentially to the development of the theory of stochastic processes, advances from which seasonal adjustment gradually benefited.

Harmonic analysis, used very early on to solve problems of series decomposition, was set into a firmly deterministic context and was used to reveal exact periodicities, though it was well recognized that the cycles might not be strictly periodic or that seasonalities could evolve. Here again, it was necessary to wait for the computers of the 1960s to benefit from advances in the theory and to use spectral analysis more effectively: better estimates of spectral densities (Bartlett [3], Tukey [66]), study of non-stationary processes (Priestley [61]), fast Fourier transform (Cooley and Tukey [17]) and so on.

With the popularization of ARIMA models by Box and Jenkins in 1970 [9], it became possible to advance seasonal adjustment tools in two directions. Firstly, an important development to Census X-11 evolved into **X-11-ARIMA** in 1975 (Dagum [19, 20]). In this new version, ARIMA models are used to extend the initial series before seasonal adjustment with the X-11 method. This tends to reduce the revisions of estimates near the end of the series. Secondly, ARIMA modelling was also introduced into seasonal adjustment methods based on the theory of signal extraction. There are many examples of works using ARIMA modelling and spectral analysis for seasonal adjustment purposes (see Bell and Hillmer [6]). As a result of all these technical and theoretical advances, model-based decomposition methods are now being developed and popularized.

Today, the two main philosophies of seasonal adjustment, i.e., the **empirical approach** and the **modelling approach**, inspire diverse methods, some blending the two. The main criticisms that can be made about each of these two approaches are difficult to avoid. Thus, empirical methods are criticized for being suboptimal and for not being based on explicit models, making it especially difficult, may be even impossible, to know the statistical properties of the estimators used. In contrast, model-based methods are satisfactory in this regard, but one questions the relevance of modelling, especially when univariate models are used for economic series, which typically depends on many external factors, and the robustness of the estimation methods in the case of very noisy series. The difficulties of modelling *a priori* components about which little is known and the relative weakness of statistical theory for non-stationary series are also pointed out.

Consequently, the improvements made by recent studies do not have to do with the principle of the existing methods per se, but rather are intended to correct some of their flaws. The main concerns are about, firstly, the problems related to the estimation of the components at the ends of the series, and secondly, the elimination of the various disruptive effects that affect the results of seasonal adjustment (extreme observations, changes in regime, calendar effects, etc.).

Schematically, as summarized in Figure 1.1, seasonal adjustment methods may be classified into two main categories: **non-parametric** methods, and **parametric** methods.

Non-parametric, so-called empirical, methods decompose a series into unobservable components using a procedure, often iterative, based on successive smoothings. All the smoothers used can be regrouped under the heading **local regressions**. Local regressions consist in fitting polynomials, generally by least squares, weighted or unweighted, over sliding intervals (shifted a point at a time). In the middle of the interval, the smoothed data point is the value, on that date, of the fitted polynomial (the smoothed data point for the following date is obtained by fitting a polynomial over the next interval). It can be shown that local regressions amount to applying moving averages when observations are regularly spaced. The methods are distinguished essentially by their degree of robustness: One has, in a first group, **STL** (Cleveland et al. [16]), a LOWESS-based method, a robust smoothing technique by local regressions (Cleveland [15]) and **SABL** where robustness is achieved by the use of running medians; and in a second group, the famous **X-11** method of seasonal adjustment (US Bureau of the Census), **X-11-ARIMA** (Statistics Canada, Dagum [21]) and **X-12-ARIMA** (US Bureau of the Census, Findley et al. [23]).

Parametric methods may also be divided into two main groups: regression methods and methods based on stochastic models.

Regression methods, inspired by Buys-Ballot [11], fit for each component, other than the irregular, a deterministic function of time. Among them, we might mention **BV4** (Technische Universität Berlin, Deutsche Institut für Wirtschaftsforschüng) and **DAINTIES**, used in the eighties at the European Economic Commission (Hylleberg [37]).

The methods based on stochastic (non-deterministic) models use primarily ARIMA models to model unobservable components. These are further divided into two subgroups: those that derive the components models from the ARIMA model of the initial series (Burman [10], Hillmer and Tiao [35]), **SEATS** (Gomez and Maravall [27]) being the most recent; and those that model and estimate them directly (Akaike [1], Kitagawa and Gersch [41]), such as the **STAMP** method (Koopman et al. [42]), **BAYSEA** and **DECOMP** (Institute of Statistical Mathematics, Japan).

Table 1.1 provides URLs for the quoted softwares [7].

[7] See also the software section of "The Econometric Journal online", http://www.econometriclinks.com, for updated links.

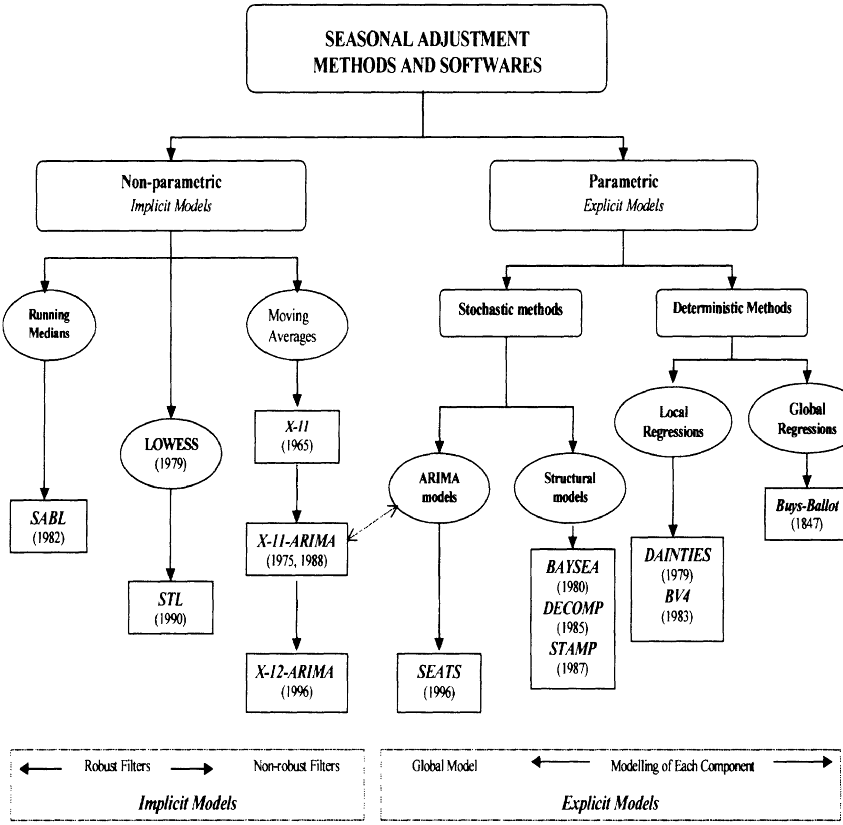


FIGURE 1.1. Seasonal adjustment methods and softwares.

Software	URL
BAYSEA	http://www.ism.ac.jp/software/products-e.html
BV4	http://www.statistik-bund.de/mve/e/bv4.htm
Web DECOMP	http://ssnt.ism.ac.jp/inets2/title.html
STL and SABL in S-PLUS	http://www.splus.mathsoft.com/splus/splsprod/default.htm
STAMP	http://stamp-software.com
TRAMO/SEATS	http://www.bde.es/servicio/software/softwaree.htm
X-11-ARIMA	http://www.statcan.ca/english/IPS/Data/10F0003XDE.htm
X-12-ARIMA	http://ftp.census.gov/pub/ts/x12a/final/pc/

TABLE 1.1. URLs for seasonal adjustment softwares.

2

Outline of the X-11 Method

The X-11 method is based on an iterative principle of estimation of the different components, this estimation being done at each step using appropriate moving averages. The method is designed for the decomposition and seasonal adjustment of monthly and quarterly series.

2.1 Components and Decomposition Models

The components that may appear at one time or another in the decomposition are:

1. The **trend**, representing the long-term evolution of the series;

2. The **cycle**, a smooth movement around the trend, revealing a succession of phases of expansion and recession.

   X-11 does not separate these two components: the series studied are generally too short for both components to be easily estimated. Consequently, hereafter we will only speak of the **trend-cycle** component, written as C_t to follow the usual X-11 notation.

3. The **seasonal** component, written as S_t, representing intra-year fluctuations, monthly or quarterly, that are repeated more or less regularly year after year;

4. A so-called **trading-day** component, written as D_t, that measures the impact on the series of the day-of-the-week composition of the month or quarter;

5. A component measuring the effect of the **Easter holiday**, written as E_t;

6. And finally, the **irregular** component, written as I_t, combining all the other more or less erratic fluctuations not covered by the previous components.

We would point out that these definitions are qualitative and rather imprecise. They are still today the subject of controversy and various interpretations. For example, here are two quotations of eminent statisticians who apparently do have neither the same objective nor the same definitions:

- Sir Kendall [40, page 29]: *"The essential idea of trend is that it shall be smooth."*

- Andrew Harvey [30, page 284]: *"There is no fundamental reason, though, why a trend should be smooth."*

In the X-11 method, the components are in fact defined implicitly by the tools used to estimate them.

The X-11 method proposes two decomposition models[1]:

1. The **additive** model: $X_t = C_t + S_t + D_t + E_t + I_t$;

2. The **multiplicative** model: $X_t = C_t \times S_t \times D_t \times E_t \times I_t$;

In addition, X-11-ARIMA proposes the **log additive** model:

$$\log X_t = \log C_t + \log S_t + \log D_t + \log E_t + \log I_t.$$

And X-12-ARIMA further adds a **pseudo-additive** model:

$$X_t = C_t \times (S_t + D_t + E_t + I_t - 1).$$

2.2 Moving Averages

Moving averages, which are the basic tool of the X-11 seasonal adjustment method, are used to estimate the main components of the series: trend-cycle and seasonality. They are, above all, **smoothing** tools designed to

[1] The Census X-11 software does not allow for the estimation of the Easter holiday effect.

eliminate an undesirable component from the series. Let us take the simple example of a series consisting of a trend and an irregular component: if the trend is smooth, then the values of the series around date t must contain information about the value of this trend at instant t and it must be possible to use an average of these values as an estimate.

A **moving average** of coefficients $\{\theta_i\}$ is thus defined[2] as:

$$M(X_t) = \hat{C}_t = \sum_{i=-p}^{+f} \theta_i X_{t+i}$$

and the whole problem is then to find the "right" set of coefficients $\{\theta_i\}$. The very limited computing facilities at the end of the 19th century prompted statisticians to look for coefficients independent of the values of the series using methods that will be examined in detail in Chapter 3.

2.3 A Simple Seasonal Adjustment Algorithm

Let a monthly unadjusted series be X_t. Assume that X_t can be decomposed into a trend-cycle, a seasonality and an irregular component according to the additive model: $X_t = C_t + S_t + I_t$. A simple four-step seasonal adjustment algorithm can be devised as follows:

1. **Estimation of the trend-cycle by moving average:**

$$C_t^{(1)} = M_0(X_t).$$

   The moving average used here should therefore reproduce, at best, the trend-cycle component while eliminating the seasonal component and minimizing the irregular component.

2. **Estimation of the seasonal-irregular component:**

$$(S_t + I_t)^{(1)} = X_t - C_t^{(1)}.$$

3. **Estimation of the seasonal component by moving average over each month:**

$$S_t^{(1)} = M_1\left[(S_t + I_t)^{(1)}\right]$$

[2]The value at instant t of the unadjusted series is replaced by a weighted average of p "past values" of the series, the current value, and f "future values" of the series. Moving averages are also called linear filters whose purpose is to minimize or suppress waves or oscillations of certain frequencies. More details are given in Chapter 3.

and therefore also

$$I_t^{(1)} = (S_t + I_t)^{(1)} - S_t^{(1)}.$$

Here, it is a question of smoothing out the values of the seasonal-irregular component separately for each month to extract the evolution of the seasonal factor of the month concerned. The moving average used here must reproduce, as best as possible, the seasonal component of each month by minimizing the irregular component.

A normalizing constraint, for example that they sum to zero, may be imposed on the factors.

4. **Estimation of the seasonally adjusted series:**

$$A_t^{(1)} = (C_t + I_t)^{(1)} = X_t - S_t^{(1)}.$$

The whole difficulty lies, then, in the choice of the moving averages used in steps 1 and 3.

2.4 The Basic Algorithm of the X-11 Method

It is this simple algorithm that the X-11 method implements using judiciously chosen moving averages, and gradually refining, by iteration of the algorithm, the estimates of the components.

It is thus possible to define the basic algorithm of the X-11 method. It actually corresponds to using the previous simple algorithm twice, changing the moving averages each time.

1. **Estimation of the trend-cycle by a 2×12 moving average:**

$$C_t^{(1)} = M_{2\times 12}(X_t).$$

The moving average used here is a so-called 2×12 moving average, of coefficients $\frac{1}{24}\{1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1\}$, which preserves linear trends, eliminates order-12 constant seasonalities and minimizes the variance of the irregular component.

2. **Estimation of the seasonal-irregular component:**

$$(S_t + I_t)^{(1)} = X_t - C_t^{(1)}.$$

3. **Estimation of the seasonal component by a 3×3 moving average over each month:**

$$S_t^{(1)} = M_{3\times 3}\left[(S_t + I_t)^{(1)}\right].$$

The moving average used here is a so-called 3×3 moving average over 5 terms, of coefficients $\frac{1}{9}\{1, 2, 3, 2, 1\}$, which preserves linear trends. The seasonal factors are then normalized so that their sum over each 12-month period is approximately zero.

$$\tilde{S}_t^{(1)} = S_t^{(1)} - M_{2\times12}\left(S_t^{(1)}\right).$$

4. **Estimation of the seasonally adjusted series:**

$$A_t^{(1)} = (C_t + I_t)^{(1)} = X_t - \tilde{S}_t^{(1)}.$$

This first estimate of the seasonally adjusted series must, by construction, contain less seasonality. The X-11 method again uses our simple algorithm, changing the moving averages to take this property into account.

5. **Estimation of the trend-cycle by a 13-term Henderson moving average:**

$$C_t^{(2)} = H_{13}\left(A_t^{(1)}\right).$$

Henderson moving averages, while they do not have special properties in terms of eliminating seasonality (limited or none at this stage), are very good smoothers and preserve a locally polynomial trend of degree 2[3].

6. **Estimation of the seasonal-irregular component:**

$$(S_t + I_t)^{(2)} = X_t - C_t^{(2)}.$$

7. **Estimation of the seasonal component by a 3×5 moving average over each month:**

$$S_t^{(2)} = M_{3\times5}\left[(S_t + I_t)^{(2)}\right].$$

The moving average used here is a so-called 3×5 moving average over 7 terms, of coefficients $\frac{1}{15}\{1, 2, 3, 3, 3, 2, 1\}$, which preserves linear trends. The seasonal factors are then normalized so that their sum over each 12-month period is approximately zero.

$$\tilde{S}_t^{(2)} = S_t^{(2)} - M_{2\times12}\left(S_t^{(2)}\right).$$

8. **Estimation of the seasonally adjusted series:**

$$A_t^{(2)} = (C_t + I_t)^{(2)} = X_t - \tilde{S}_t^{(2)}.$$

This X-11 basic algorithm is further summarized in Table 2.1.

---

[3]Because the Henderson moving average is symmetric, it also preserves a locally polynomial trend of degree 3 (see Chapter 3).

Monthly unadjusted series: $X_t = C_t + S_t + I_t$

1. **Estimation of the trend-cycle by a 2×12 moving average:**
   $C_t^{(1)} = M_{2\times12}(X_t)$
2. **Estimation of the seasonal-irregular component:**
   $(S_t + I_t)^{(1)} = X_t - C_t^{(1)}$
3. **Estimation of the seasonal component by 3×3 moving average over each month:**
   $S_t^{(1)} = M_{3\times3}\left[(S_t + I_t)^{(1)}\right]$
   and normalization
   $\tilde{S}_t^{(1)} = S_t^{(1)} - M_{2\times12}\left(S_t^{(1)}\right)$
4. **Estimation of the seasonally adjusted series:**
   $A_t^{(1)} = (C_t + I_t)^{(1)} = X_t - \tilde{S}_t^{(1)}$

5. **Estimation of the trend-cycle by a 13-term Henderson moving average:**
   $C_t^{(2)} = H_{13}\left(A_t^{(1)}\right)$
6. **Estimation of the seasonal-irregular component:**
   $(S_t + I_t)^{(2)} = X_t - C_t^{(2)}$
7. **Estimation of the seasonal component by 3×5 moving average over each month:**
   $S_t^{(2)} = M_{3\times5}\left[(S_t + I_t)^{(2)}\right]$
   and normalization
   $\tilde{S}_t^{(2)} = S_t^{(2)} - M_{2\times12}\left(S_t^{(2)}\right)$
8. **Estimation of the seasonally adjusted series:**
   $A_t^{(2)} = (C_t + I_t)^{(2)} = X_t - \tilde{S}_t^{(2)}$

TABLE 2.1. X-11 basic algorithm.

2.5 Extreme Observations and Calendar Effects

As with any linear operator, moving averages respond dramatically to the presence of extreme observations. The X-11 method therefore incorporates a **tool for the detection and modification of extreme values** used to clean up the series prior to seasonal adjustment.

Also, effects other than seasonality may explain the variations observed in the series; the most common are effects related to the calendar: trading-day effect, Easter effect, and so on. These components are estimated using linear regression models, based on the irregular component[4].

The X-11 basic algorithm, described in Table 2.1, produces 3 different estimates of the irregular component:

- At step 3 by subtracting the estimate of the seasonal component from the estimate of the seasonal-irregular component obtained in step 2:

  $$I_t^{(1)} = (S_t + I_t)^{(1)} - \tilde{S}_t^{(1)}.$$

  X-11 will use this estimate to detect and correct the extreme observations and obtain a better estimate of the seasonal component.

- At step 7, by subtracting the estimate of the seasonal component from the estimate of the seasonal-irregular component obtained in step 6:

  $$I_t^{(2)} = (S_t + I_t)^{(2)} - \tilde{S}_t^{(2)}.$$

  X-11 will use this estimate again to detect and correct the extreme observations and obtain a more reliable estimate of the seasonal component.

- At step 8, by subtracting from the estimate of the seasonally adjusted series the estimate of the trend-cycle component obtained in step 5:

  $$I_t^{(3)} = A_t^{(2)} - C_t^{(2)}.$$

  X-11 will use this estimate to evaluate, by linear regression, the trading-day component and to detect and correct extreme observations[5].

2.6 The Iterative Principle of X-11

To evaluate the different components of the series, while taking into account the possible presence of extreme values, X-11 proceeds iteratively: estima-

---

[4]X-12-ARIMA has a "Reg-ARIMA" module by which it is possible to estimate directly these effects on the unadjusted series before performing the seasonal adjustment. This module will not be studied here.

[5]These various methods will be presented in Chapter 4.

tion of components, search for disruptive effects in the irregular component, estimation of components from a corrected series, search for disruptive effects in the irregular component, and so on.

The Census X-11 program presents 4 processing stages (A, B, C, and D), plus 3 stages, E, F, and G, that present statistics and charts that are not part of the decomposition per se.

2.6.1 Part A: Pre-Adjustments

This part, which is optional, allows the user to correct *a priori* the series by introducing adjustment factors. The user can thus:

- introduce monthly (or quarterly) adjustments factors that will allow him to correct the effect of certain statutory holidays, change the series level (effect of a strike for example), etc.;
- for monthly series only, introduce 7 weights, one of each day of the week, to take into account the variations attributable to the trading-day composition of months.

Based on these data, the program calculates prior adjustment factors that are applied to the raw series. The series thus corrected, Table B1 of the printouts, then proceeds to Part B.

2.6.2 Part B: First Automatic Correction of the Series

This stage comprises a first estimation and downweighting of the extreme observations and, if requested, a first estimation of the trading-day effects. This stage is performed by applying the basic algorithm detailed in Section 2.4.

These operations lead to tables B19, evaluation of trading-day effects, and B20, adjustment values for extreme observations, that are used to correct the prior-adjusted series in Table B1, and result in the series shown in Table C1.

2.6.3 Part C: Second Automatic Correction of the Series

Still applying the basic algorithm, this part leads to a more precise estimation of trading-day effects (Table C19) and replacement values for the extreme observations (Table C20).

The series, finally "cleaned up," is shown in Table D1 of the printouts.

2.6.4 Part D: Seasonal Adjustment

This part, at which the basic algorithm is applied for the last time, is that of the seasonal adjustment per se, as it leads to final estimates:

- of the seasonal component (Table D10),
- of the seasonally adjusted series (Table D11),
- of the trend-cycle component (Table D12),
- of the irregular component (Table D13).

2.6.5 Parts E, F and G: Statistics and Charts

Parts E and F propose statistics for judging the quality of the seasonal adjustment.

Part G proposes graphics in character mode. It can be ignored, as today it can be replaced by the usual office-based graphics software.

A summary of the processing stages of the X-11 method is shown in Table 2.2.

Part A:	**Prior adjustments**
	• for significant known extremes
	• for trading-day
Part B:	**First automatic correction of the series**
	Estimation of the irregular component
	Detection and automatic correction of extreme observations
	Correction of trading-day effects
Part C:	**Second automatic correction of the series**
	Estimation of the irregular component
	Detection and automatic correction of extreme observations
	Correction of trading-day effects
Part D:	**Seasonal adjustment**
1	Calculation of the temporary seasonally adjusted series (Tables D1 to D6)
2	Smoothing out of the seasonally adjusted series by a Henderson moving average and new estimation of seasonal factors (Tables D7 to D10)
3	Calculation of the final seasonally adjusted series (Table D11) and extraction of the trend-cycle (Table D12) and irregular component (Table D13)
Part E:	**Components modified for large extreme values**
Part F:	**Seasonal adjustment quality measures**
Part G:	**Graphics**

TABLE 2.2. Simplified diagram of X-11 operation.

2.7 From Census X-11 to X-11-ARIMA and X-12-ARIMA

The use of moving averages, as we will see in Chapter 3, poses problems at the series' ends, notably with regard to the stability of the estimates. Thus, when you have one or more additional points and the series is again seasonally adjusted using the X-11 procedure, it is not uncommon to note appreciable revisions in the estimates for the most recent dates.

As early as 1975, Estella B. Dagum [19] proposed an improvement using the ARIMA models that became popular some years earlier through the works of Box and Jenkins [9]. She thus showed that revisions were reduced appreciably by fitting an ARIMA model to the series, predicting the future values of the series using this model, and applying the X-11 seasonal adjustment procedure to the series thus extended. This is the idea that underlies the X-11-ARIMA software (Dagum [21])[6].

Unfortunately, the estimation of ARIMA models is made tricky by the presence of extreme observations, level changes, calendar effects, and so on. X-11-ARIMA is therefore based on the following scheme:

1. **First seasonal adjustment by the X-11 method:**

   At this stage, it is possible to estimate the extreme values, the trading-day effects, as we have seen, and also the Easter effects using the estimate of the irregular component from Table D13.

2. **ARIMA modelling and forecasting of the series corrected for all these effects.**

3. **Second seasonal adjustment by the X-11 method, applied to the extended series.**

X-12-ARIMA is based on the same principle but proposes, in addition, a complete module, called Reg-ARIMA, that allows for the initial series to be corrected for all sorts of undesirable effects. These effects are estimated using regression models with ARIMA errors (Findley et al. [23]).

[6]This idea was already implicitly expressed by Frederick Macaulay [51] in 1931:

> *However, graduation of the ends of almost any series is necessarily extremely hypothetical unless facts outside the range covered by the graduation are used in obtaining the graduation .... Though mathematically inelegant, the most desirable procedure in a majority of the cases of graduation is to graduate not only the actual data, but extrapolated data which sometimes may be extremely crude estimates.*

We pay tribute to Estella B. Dagum for having successfully applied it.

3

Moving Averages

The X-11 method of seasonal adjustment uses moving averages to estimate the main components of the series: trend-cycle and seasonality. These filters, which do not involve *a priori* the use of sophisticated concepts or model, are very simple in principle and especially flexible in their application: it is possible to construct a moving average that has good properties in terms of trend preservation, elimination of seasonality, noise reduction, and so on.

In this chapter we will study their properties and the principles that guided the construction of the moving averages used in X-11.

3.1 Some Definitions and a Little Theory

A time series may be considered from two standpoints: time, and frequency.

- In the *time domain*, the series $\{X_t\}$ is regarded as a succession of T observed values at instants t, t varying from 1 to T. This is how a time series is generally approached, and it is easy to show graphically, as in Figure 3.1, its evolution over time. Note that this series is characterized by strong seasonality expressing the drop in industrial activity in the month of August.

  The modelling of the series or its components, relating the value at instant t to those of past instant, is especially easy to formalize. This is the case, for example, for the modelling of the series using a seasonal ARIMA model, or using the expression of a linear, exponential or

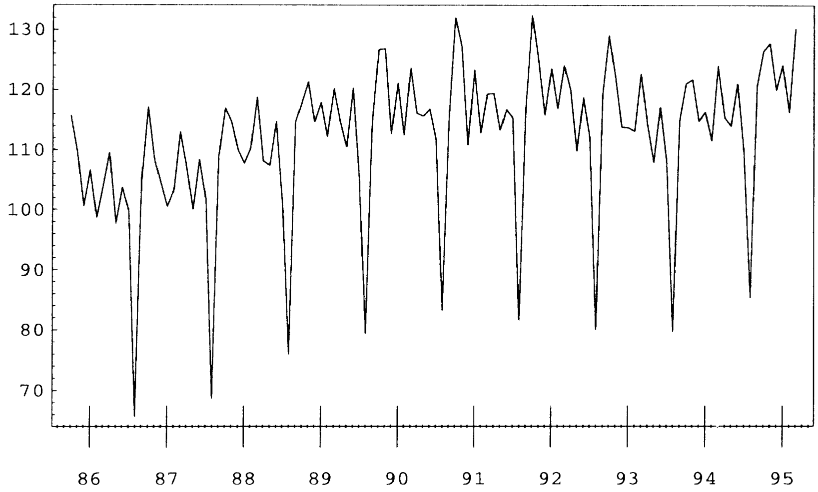


FIGURE 3.1. French Industrial Production Index, monthly, October 1985 to March 1995.

even locally polynomial trend, and the modelling of the irregular component by a white noise.

- In contrast, in the *frequency domain*, one begins with the series $\{X_t\}$ expressed as the sum of trigonometric functions[1]. The importance of each frequency in the composition of the series is measured: the graph that associates with each frequency its importance in the series is called the spectrum of the series. Thus, Figure 3.2 shows the spectrum of the French Industrial Production Index.

  As can be seen, this spectrum reveals a strong contribution (called a spectral peak) at frequency $\pi/6 = 30°$, and its multiples $2\pi/6 = 60°$, $3\pi/6 = 90°$, ..., $6\pi/6 = 180°$. The period associated with this frequency is $\omega = 2\pi/f = 2\pi/(\pi/6) = 12$ and we find the monthly seasonality observed in the previous chart.

  The low frequencies correspond naturally to slowly changing components, for example trend and cycle, and the high frequencies to more quickly changing components, such as the irregular component.

These two approaches often prove complementary and we will subsequently use one or the other to show the qualities and defects of moving average filters.

---

[1] In his *Théorie Analytique de la Chaleur* (first reading on 21 December 1807, first publication in 1822) Jean-Baptiste Fourier [25] established that any mathematical function could be decomposed into a sum of sine and cosine functions. This theorem gave rise initially to harmonic analysis, and then, once generalized, to spectral analysis.

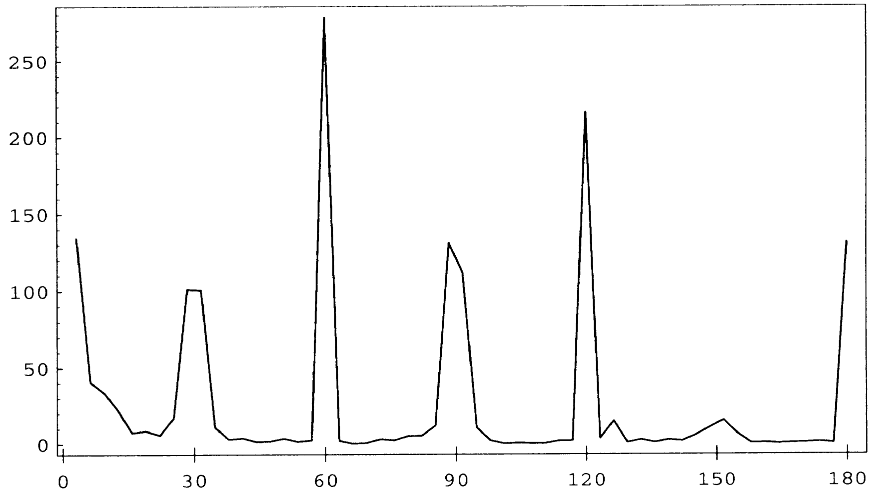


FIGURE 3.2. Spectrum of the French Industrial Production Index.

3.1.1 Definitions and Example

We call the **moving average of coefficients** $\{\theta_k\}$ the operator written as $M\{\theta_k\}$, or simply M, defined by:

$$M(X_t) = \sum_{k=-p}^{+f} \theta_k X_{t+k}$$

The value at instant t of the unadjusted series is therefore replaced by a weighted average of p "past" values of the series, the current value, and f "future" values of the series.

- The quantity $p + f + 1$ is called the **moving average order**.
- When p is equal to f, that is, when the number of points in the past is the same as the number of points in the future, the moving average is said to be **centered**.
- If, in addition, $\theta_{-k} = \theta_k$ for any k, the moving average M is said to be **symmetric**. In this case, when listing the coefficients of the moving average, it will suffice to specify the order of the moving average and the $k + 1$ first coefficients (Kendall [40]). For example:

  $$\frac{1}{24}\{1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1\}$$

  is written simply as

  $$[13]; \frac{1}{24}\{1, 2, 2, 2, 2, 2, 2\}.$$

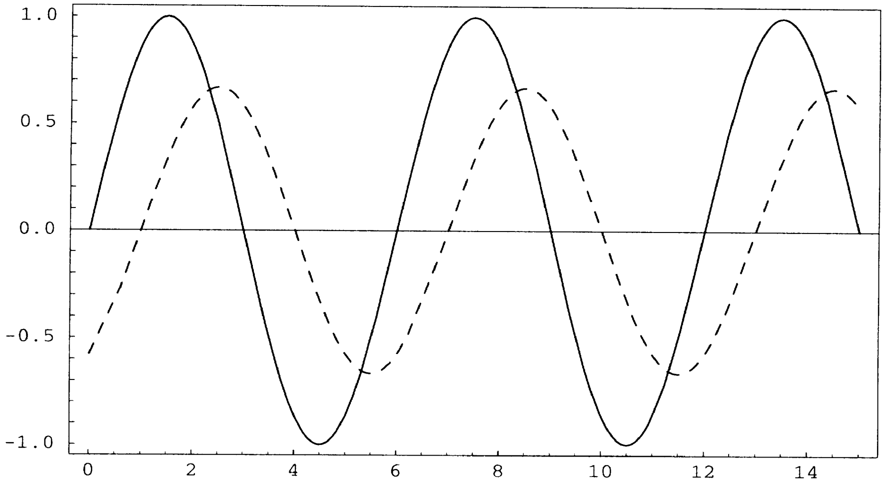


FIGURE 3.3. Smoothing of the series $X_t = \sin(\pi t/3)$ by the moving average $(X_{t-2} + X_{t-1} + X_t)/3$. The original series is graphed with a solid line, and the smoothed series is graphed with a broken line.

Generally, with a moving average of order $p+f+1$ calculated for instant t with p points in the past and f points in the future, it will be impossible to smooth out the first p values and the last f values of the series.

In the X-11 method, symmetric moving averages play an important role; to avoid losing information at the series ends, they are supplemented by ad hoc asymmetric moving averages.

3.1.2 Gain and Phase Shift Functions

Let us consider the series $X_t = \sin(\pi t/3)$ and transform it using the asymmetric moving average defined by $M(X_t) = (X_{t-2} + X_{t-1} + X_t)/3$ which replaces the value at instant t with the simple average of the values at the present instant and the two previous instants.

Figure 3.3 shows the result of the smoothing and reveals two phenomena:

- First of all, a reduction in the amplitude of the series, which effectively meets our objective of smoothing;
- But also a time lag, called a **phase shift**: the two series do not show turning points on the same dates.

This phase shift phenomenon is disagreeable insofar as it transforms the changes in the series themselves. It can be shown that symmetric moving averages do not result in a phase shift (see, for example, Koopmans [43]).

More generally, let $X_t = R\sin(\omega t + \phi)$ be a series of frequency ω (or period $2\pi/\omega$), of amplitude R and of phase ϕ. The transform of $\{X_t\}$ by any moving average will still be a sine curve, but with a modified amplitude and a phase shift in relation to the original series:

$$M(X_t) = M[R\sin(\omega t + \phi)] = G(\omega) R \sin[\omega t + \phi + \Gamma(\omega)].$$

- The function $|G(\omega)|$ is called the **gain function** of the moving average.
- The function $\Gamma(\omega)$ is called the **phase shift function** of the moving average. It is sometimes represented as $\Gamma(\omega)/\omega$ which allows the phase shift to be measured in number of periods.

In the case of the asymmetric 3-term moving average considered before, we have:

$$\begin{aligned} M(X_t) &= \frac{1}{3}(X_{t-2} + X_{t-1} + X_t) \\ &= \frac{1}{3}R\{\sin[\omega(t-2)+\phi] + \sin[\omega(t-1)+\phi] + \sin(\omega t + \phi)\} \\ &= \frac{1}{3}R(1 + 2\cos\omega)\sin(\omega t + \phi - \omega) \end{aligned}$$

and so:

$$\begin{aligned} G(\omega) &= \frac{1 + 2\cos\omega}{3}, \\ \Gamma(\omega) &= -\omega \quad \text{so that} \quad \Gamma(\omega)/\omega = -1. \end{aligned}$$

The gain function, represented in Figure 3.4, shows that the moving average cancels out the frequencies $120° = 2\pi/3 = 4 \times 2\pi/12$. It would be well suited for surveys done every 4 months (therefore, period 3) as it would thus eliminate seasonality while retaining the long term variations corresponding to low frequencies. In contrast, this average introduces a systematic phase shift of one period that could result in noticing a trend turning point one period too late.

The gain function therefore basically shows the frequencies eliminated or preserved by the moving average.

The phase shift function shows the lags introduced by the use of asymmetric moving averages, as it is zero for symmetric moving averages. Since the X-11 method emphasizes symmetric moving averages, we will not discuss this function any further.

For smoothing, the "ideal" filter would be one that would leave low frequencies unchanged, such as, for example, the periodic functions of period greater than a year (trend and cycle), but would eliminate all high frequencies corresponding to periodicities less than or equal to the year (seasonality and irregular). The gain function of this ideal filter, known as the **low-pass** filter, would therefore be as follows:

$$G(\omega) = \begin{cases} 1 & \text{for} \quad \omega \le \omega_0, \\ 0 & \text{for} \quad \omega > \omega_0. \end{cases}$$

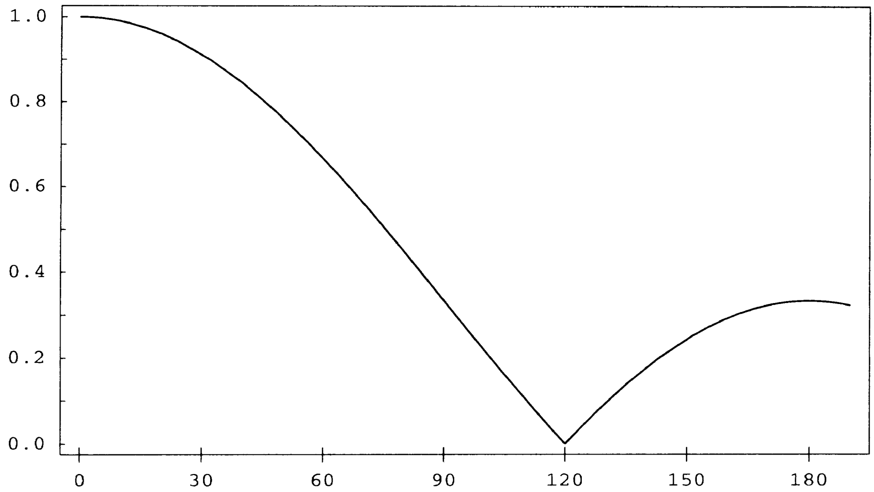


FIGURE 3.4. Gain function of the moving average $(X_{t-2} + X_{t-1} + X_t)/3$.

3.1.3 *Trend Preservation*

The phase shift introduced by the moving average $(X_{t-2} + X_{t-1} + X_t)/3$ can also be seen by applying this asymmetric average to a simple straight line $X_t = at + b$. We have:

$$\begin{aligned} M(X_t) &= \frac{1}{3}(X_{t-2} + X_{t-1} + X_t) \\ &= \frac{1}{3}[a(t-2) + b + a(t-1) + b + at + b] \\ &= a(t-1) + b \\ &= X_{t-1}. \end{aligned}$$

It would, however, be desirable for a moving average to preserve simple trends, particularly polynomials.

- Now, for any moving average to preserve constant series $X_t = a$, it is necessary that:

$$M(X_t) = \sum_{k=-p}^{+f} \theta_k X_{t+k} = \sum_{k=-p}^{+f} \theta_k a = a \sum_{k=-p}^{+f} \theta_k = a,$$

  and therefore that the sum of the coefficients of the moving average $\sum_{k=-p}^{+f} \theta_k$ be equal to 1.

- For any moving average to preserve straight lines, it is necessary, for any t, that:

$$M(X_t) = \sum_{k=-p}^{+f} \theta_k X_{t+k} = \sum_{k=-p}^{+f} \theta_k [a(t+k) + b]$$

$$= \quad at \sum_{k=-p}^{+f} \theta_k + a \sum_{k=-p}^{+f} k\theta_k + b \sum_{k=-p}^{+f} \theta_k = at + b,$$

which results in $\sum_{k=-p}^{+f} \theta_k = 1$ and $\sum_{k=-p}^{+f} k\theta_k = 0$.

- Generally, it can be shown that for a moving average to preserve a polynomial of degree d, it is necessary and sufficient that its coefficients satisfy:

$$\sum_{k=-p}^{+f} \theta_k = 1$$

and

$$\sum_{k=-p}^{+f} k^j \theta_k = 0, \ j = 1, \ldots, d.$$

Thus, for the asymmetric 3-term moving average defined above, we have:

$$\sum_{k=-2}^{0} \theta_k = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$$

and

$$\sum_{k=-2}^{0} k\theta_k = -2 \times \frac{1}{3} - 1 \times \frac{1}{3} + 0 \times \frac{1}{3} = -1$$

so that this average, while it preserves the constants, does not retain the straight lines.

It is easy, however, to establish that the following symmetric moving averages preserve straight lines:

$$\begin{aligned} M_1(X_t) &= \frac{1}{3}(X_{t-1} + X_t + X_{t+1}) \\ M_2(X_t) &= \frac{1}{8}(X_{t-2} + 2X_{t-1} + 2X_t + 2X_{t+1} + X_{t+2}). \end{aligned}$$

3.1.4 Elimination of Seasonality

As we saw when defining the gain function in Section 3.1.2, moving averages can eliminate certain frequencies and therefore certain seasonal components. It is the gain function, moreover, that most readily identifies the frequencies eliminated by a moving average.

Generally, a simple moving average of order k (whose coefficients are all equal to $1/k$) cancels out the fixed seasonalities of period k and its gain function therefore is 0 at $2\pi/k$.

Moreover, it is possible to treat the case of seasonalities that vary linearly, or even polynomially, with time (Grun-Rehomme and Ladiray [29]). To eliminate these seasonalities, specific linear constraints are required on the coefficients.

3.1.5 Reduction of the Irregular Component

After the trend and the seasonality, it remains to see the effect of a moving average on the irregular component. The residual, in the decomposition of the unadjusted series, is often modelled in the form of a **white noise**, a sequence of random variables, ϵ_t, of zero expectation, non-correlated, and of constant variance σ^2. This white noise is transformed by the moving average into a sequence of random variables, ϵ_t^*, of constant variance $\sigma^2 \sum_{k=-p}^{+f} \theta_k^2$. Reducing the irregular component, and therefore its variance, amounts to reducing the quantity $\sum_{k=-p}^{+f} \theta_k^2$.

3.1.6 An Example of Construction of a Moving Average

Let us look, for example, for a centered 3-term moving average, of coefficients $\{\theta_{-1}, \theta_0, \theta_1\}$, that minimizes the irregular component and preserves straight lines. According to the foregoing, this amounts to solving the problem: Minimize $\sum_{k=-1}^{+1} \theta_k^2$ under the constraints $\sum_{k=-1}^{+1} \theta_k = 1$ and $\sum_{k=-1}^{+1} k\theta_k = 0$.

The last constraint implies that $\theta_{-1} = \theta_1$. Substituting this into the first constraint gives $\theta_0 = 1 - 2\theta_1$, and the minimization problem becomes $\min_{\theta_1} \left[2\theta_1^2 + (1 - 2\theta_1)^2\right]$. The derivative, with respect to θ_1, of the function to be minimized is $12\theta_1 - 4$ which therefore cancels out when $\theta_1 = 1/3$. Hence, we get the simple 3-term moving average of coefficients all equal to $1/3$, an average which, as we have already seen, eliminates order-3 seasonalities.

3.2 The Symmetric Moving Averages Used in X-11

3.2.1 Composite Simple Moving Average

A so-called $\mathbf{P \times Q}$ **moving average** is obtained by composing a simple moving average of order P, whose coefficients are all equal to $1/P$, and a simple moving average of order Q, whose coefficients are all equal to $1/Q$. In concrete terms, this amounts to applying both simple moving averages in succession.

Thus, the 3×3 moving average that results from the double application of the simple arithmetic 3-term moving average is a moving average of coefficients $\{1, 2, 3, 2, 1\}/9$. Generally, a $P \times Q$ moving average is a symmetric moving average of order $P + Q - 1$[2].

[2]Simple composite moving averages are weighted moving averages. In pre-computer days, they were easier to use than weighted moving averages, while providing similar benefits.

When the order Q is even, for example equal to $2q$, there is a slight ambiguity in the definition, in that one can choose either q points in the past and $q-1$ points in the future, or $q-1$ points in the past and q points in the future. The problem is usually resolved by using a symmetric $2 \times Q$ composite average that corresponds to the average of the two possible moving averages.

Estimation of the Trend-Cycle: 2×4 *and* 2×12 *Averages*

When X-11 performs an initial estimation of trend-cycle (Tables B2, C2 and D2), it uses a 2×4 moving average in the quarterly case, and a 2×12 moving average in the monthly case. At that point, the series to be smoothed is composed of the trend-cycle, seasonal and irregular components. In the case of an additive decomposition model, it may be written: $X_t = C_t + S_t + I_t$.

The 2×4 *Average*

This is a moving average of order 5, with coefficients $\{1, 2, 2, 2, 1\}/8$. The coefficients curve and the gain function, shown in Figure 3.5, reveal the properties of this moving average:

- It eliminates frequency $90° = 2\pi/4 = \pi/2$ corresponding to period 4 and is therefore well suited to quarterly series having a constant seasonality.
- The sum of its coefficients is equal to 1 and it is symmetric: it therefore preserves linear trends.
- The sum of the squares of its coefficients is equal to 0.250 and it therefore reduces the variance of white noise by 75%.

Using this moving average will work optimally when the trend-cycle of our series is linear, or locally linear, that the seasonal factors are constant or vary little over time, and that the irregular component has no structure and is of limited amplitude. In this case, we will have:

$$
\begin{aligned}
M_{2\times 4}(X_t) &= M_{2\times 4}(C_t + S_t + I_t) \\
&= M_{2\times 4}(C_t) + M_{2\times 4}(S_t) + M_{2\times 4}(I_t) \\
&\approx C_t + 0 + \epsilon_t \\
&\approx C_t
\end{aligned}
$$

This moving average, however, restores rather poorly the low frequencies associated with periods greater than a year. Thus, 3-year periodic functions, which correspond, in the quarterly case, to frequencies of $30° = 2\pi/12 = \pi/6$, are only about 80% restored.

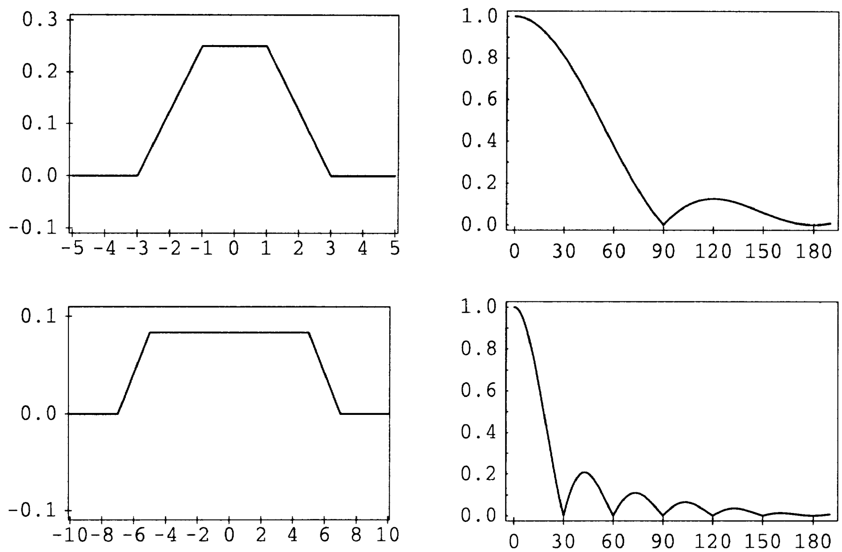


FIGURE 3.5. Coefficient curves (on the left) and gain functions (on the right) of the composite moving averages used in X-11 trend-cycle estimation. The 2×4 is displayed in the upper panel, and the 2×12 in the lower panel.

The 2×12 *Average*

This average is based on the same notions as those explained for the 2×4 average and is used in the monthly case. Its coefficients are:

$$\{1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1\}/24.$$

This average is also known as a centered 12-term moving average. It therefore preserves the straight lines and, as its gain function shows, eliminates annual seasonalities (which correspond to a frequency of $30° = 2\pi/12 = \pi/6$). Moreover, the sum of the squares of its coefficients being equal to $23/288$, it reduces the variance of white noise by more than 90%.

But here again, not all periodic series of periods less than a year are very well restored. Thus, a 3-year periodic function, which corresponds here to a frequency of $10° = 2\pi/36 = \pi/18$, will be only 80% restored.

Comments

- The 2×4 and 2×12 moving averages are also used in the X-11 method to normalize the seasonal factors.

- X-11-ARIMA and X-12-ARIMA also offer as an option centered 8-term and 24-term moving averages due to Pierre Cholette [13]. These averages have been constructed based on criteria that are fairly different from, and more complex than, those presented here. They will not be studied in this book.

Estimation of Seasonality: 3×3, 3×5 *and* 3×9 *Averages*

These averages[3] are used by X-11 to extract the seasonal component based on an estimation of the seasonal-irregular component. They are therefore used in the construction of Tables B4, B5, B9, B10, C5, C10, D5 and D10.

At that point, the series to be smoothed is composed of the seasonal and irregular components. In the case of an additive decomposition model, it may be written: $SI_t = S_t + I_t$ and, unlike previously, the problem is strictly one of smoothing.

The coefficients of the $3 \times 3, 3 \times 5$ and 3×9 averages are as follows:

$$
\begin{aligned}
M_{3\times 3} &: \{1, 2, 3, 2, 1\}/9 \\
M_{3\times 5} &: \{1, 2, 3, 3, 3, 2, 1\}/15 \\
M_{3\times 9} &: \{1, 2, 3, 3, 3, 3, 3, 3, 3, 2, 1\}/27.
\end{aligned}
$$

As can easily be established, each of these symmetric moving averages preserves straight lines. But the gain functions, shown in Figure 3.6, are fairly different from the ideal form associated with a low-pass filter. Here, it is not a hindrance insofar as the seasonal-irregular component is not supposed to present a cyclical component whose periodicity is of the order of 3 to 6 years.

Applying, for example, the 3×3 filter separately to each month of an estimate of the seasonal-irregular component amounts to applying to this selfsame component a 49-term moving average of coefficients:

$$
\begin{aligned}
\{&1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \\
&2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \\
&3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \\
&2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \\
&1\}/9.
\end{aligned}
$$

The same is true with respect to the 3×5 and 3×9 filters with which are associated 73-term and 121-term moving averages respectively! The gain functions of these new moving averages are shown in Figure 3.7. As can be seen, these moving averages preserve the annual seasonality since they exactly restore the multiple frequencies of $30° = 2\pi/12 = \pi/6$.

[3] X-11-ARIMA also allows for the use of a simple moving average of order 3 (a 3×1). X-12-ARIMA further adds a 3×15 moving average.

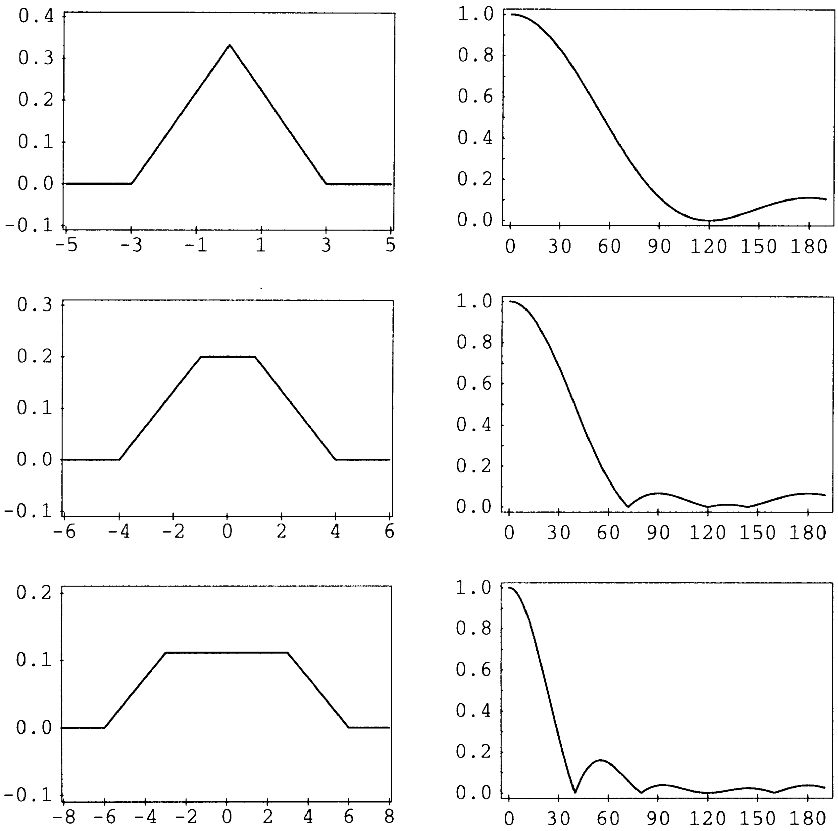


FIGURE 3.6. Coefficient curves (on the left) and gain functions (on the right) of the composite moving averages used in X-11 seasonal factor estimation. The 3×3 is displayed in the upper panel, the 3×5 in the middle panel and the 3×9 in the lower panel.

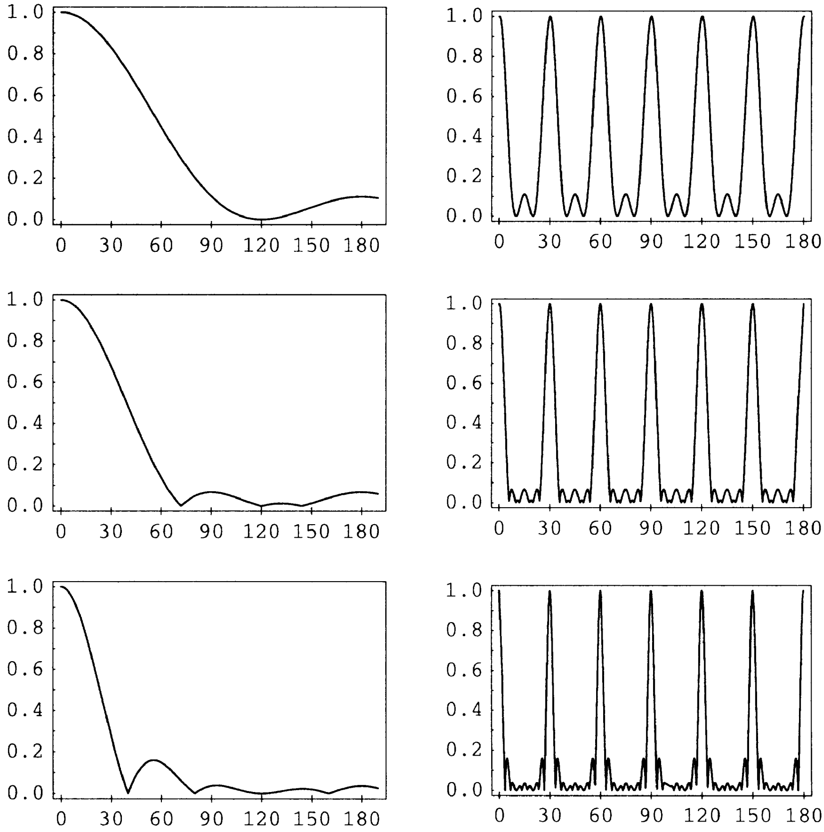


FIGURE 3.7. Gain functions of the $3 \times 3, 3 \times 5$ and 3×9 averages. The 3×3 is displayed in the upper panel, the 3×5 in the middle panel and the 3×9 in the lower panel. The gain functions over each month are displayed on the left, and over the whole series on the right.

i	2 × 4	2 × 12	3 × 3	3 × 5	3 × 9
-6		1/24			
-5		1/12			1/27
-4		1/12			2/27
-3		1/12		1/15	3/27
-2	1/8	1/12	1/9	2/15	3/27
-1	1/4	1/12	2/9	3/15	3/27
0	1/4	1/12	3/9	3/15	3/27
1	1/4	1/12	2/9	3/15	3/27
2	1/8	1/12	1/9	2/15	3/27
3		1/12		1/15	3/27
4		1/12			2/27
5		1/12			1/27
6		1/24			
$\sum \theta_i^2$	0.2188	0.0799	0.2346	0.1644	0.1001
$\sum (\nabla^3\theta_i)^2$	0.1250	0.0139	0.1481	0.0356	0.0110

TABLE 3.1. Coefficients of the composite moving averages used in X-11, their variance reducing power, $\sum \theta_i^2$, and Henderson criterion, $\sum (\nabla^3\theta_i)^2$.

3.2.2 Henderson Moving Averages

Henderson moving averages are used in X-11 to extract the trend-cycle from an estimate of the seasonally adjusted series (Tables B7, C7, D7, D12). In the additive case, we therefore have a model of the type: $A_t = C_t + I_t$. What criterion can be used to ensure a smooth estimation of the trend-cycle? Let us consider the series

$$X_t = \begin{cases} 1 & \text{if} \quad t = 0 \\ 0 & \text{if} \quad t \neq 0. \end{cases}$$

Its transform by a centered moving average M of order $2p+1$ with coefficients $\{\theta_i\}$, is given by:

$$M(X_t) = \begin{cases} 0 & \text{if} \quad t < -p. \\ \theta_i & \text{if} \quad -p \leq t \leq p \\ 0 & \text{if} \quad t > p. \end{cases}$$

This transform will therefore be smooth if the coefficient curve of the moving average is smooth. As can be seen in Figures 3.5 and 3.6, the coefficient curves of the composite moving averages are not smooth. Henderson [31, 32] proposed using the quantity $H = \sum (\nabla^3\theta_i)^2$, where ∇ represents the first difference operator[4], to measure the "flexibility" of the coefficient curve. This quantity vanishes when the coefficients $\{\theta_i\}$ are located along a parabola and, in the general case, it measures the difference between a parabola and the form of the coefficient curve. Henderson then looked for centered averages of order $2p+1$ that preserve quadratic polynomials and minimize quantity H.

[4] $\nabla X_t = X_t - X_{t-1}$

With the notation of Section 3.1.6, the order $2p+1$ Henderson moving average will be the solution of the following minimization program: Minimize, with respect to the θ_i's, $\sum_{i=-p}^{p}\left(\nabla^3\theta_i\right)^2$ with the constraints

$$\sum_{i=-p}^{p}\theta_i=1,\ \sum_{i=-p}^{p}i\theta_i=0,\ \sum_{i=-p}^{p}i^2\theta_i=0.$$

The coefficients of these moving averages may also be calculated explicitly and, for an order $2p+1$ average, by letting $n=p+2$, we have:

$$\theta_i=\frac{315\left[(n-1)^2-i^2\right]\left[n^2-i^2\right]\left[(n+1)^2-i^2\right]\left[3n^2-16-11i^2\right]}{8n(n^2-1)(4n^2-1)(4n^2-9)(4n^2-25)}.$$

Using this formula, it is therefore possible to calculate, in rational form, the coefficients of the Henderson moving averages used in X-11. We thus have, showing only the necessary coefficients because of symmetry:

5-term: $[5];\ \frac{1}{286}\{-21,84,160\}$,

7-term: $[7];\ \frac{1}{715}\{-42,42,210,295\}$,

9-term: $[9];\ \frac{1}{2431}\{-99,-24,288,648,805\}$,

13-term: $[13];\ \frac{1}{16796}\{-325,-468,0,1100,2475,3600,4032\}$,

23-term: $[23];\ \frac{1}{4032015}\{-17250,-44022,-63250,-58575,-19950,$
$54150,156978,275400,392700,491700,557700,580853\}$.

Table 3.2 provides, in decimal form, the coefficients of the Henderson moving averages used in X-11[5] as well as their variance reduction power and Henderson criterion. The coefficient curves, shown in Figure 3.8, are smooth and the gain functions of these averages are closer to the ideal low-pass form than those of the composite moving averages seen earlier.

3.3 Musgrave Asymmetric Moving Averages

When applying a centered moving average of order $2p+1$, it is not possible to have, by construction, estimates of the smoothed series for the first and last p observations of the series, which is, at the very least, bothersome as usually the most important point in a series is the last point. One is

[5]With X-12-ARIMA, it is possible to use any odd Henderson moving average of order less than 101.

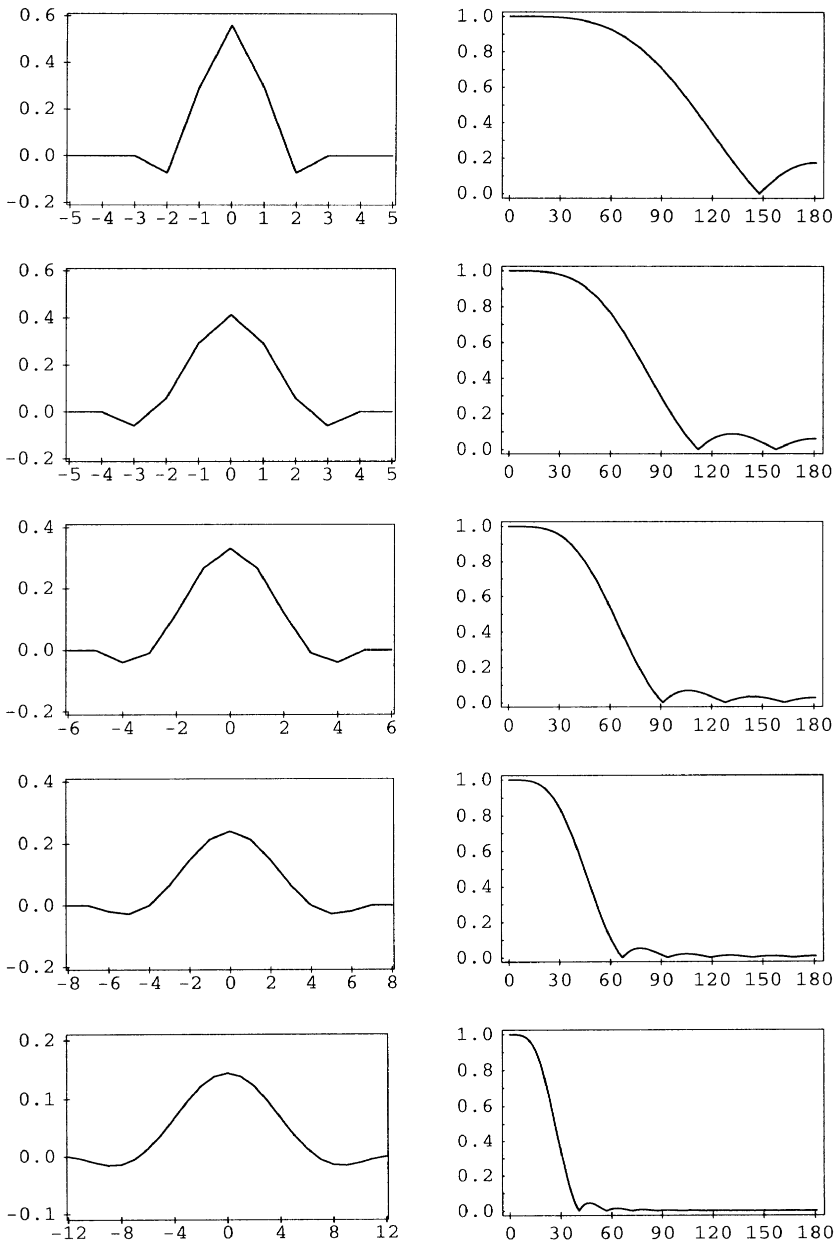


FIGURE 3.8. Coefficient curves and gain functions, from top to bottom, of the centered 5, 7, 9, 13 and 23-term Henderson moving averages used in X-11. Coefficient curves are on the left, gain functions on the right.

i	5-term	7-term	9-term	13-term	23-term
-11	.	.	.	.	-0.00428
-10	.	.	.	.	-0.01092
-9	.	.	.	.	-0.01569
-8	.	.	.	.	-0.01453
-7	.	.	.	.	-0.00495
-6	.	.	.	-0.01935	0.01343
-5	.	.	.	-0.02786	0.03893
-4	.	.	-0.04072	0.00000	0.06830
-3	.	-0.05874	-0.00987	0.06549	0.09740
-2	-0.07343	0.05874	0.11847	0.14736	0.12195
-1	0.29371	0.29371	0.26656	0.21434	0.13832
0	0.55944	0.41259	0.33114	0.24006	0.14406
1	0.29371	0.29371	0.26656	0.21434	0.13832
2	-0.07343	0.05874	0.11847	0.14736	0.12195
3	.	-0.05874	-0.00987	0.06549	0.09740
4	.	.	-0.04072	0.00000	0.06830
5	.	.	.	-0.02786	0.03893
6	.	.	.	-0.01935	0.01343
7	.	.	.	.	-0.00495
8	.	.	.	.	-0.01453
9	.	.	.	.	-0.01569
10	.	.	.	.	-0.01092
11	.	.	.	.	-0.00428
$\sum \theta_i^2$	0.4963	0.3566	0.2833	0.2038	0.1217
$\sum \left(\nabla^3 \theta_i\right)^2$	1.4965	0.2629	0.0675	0.0083	0.0003

TABLE 3.2. Coefficients of the Henderson moving averages used in X-11, their variance reducing power, $\sum \theta_i^2$, and Henderson criterion, $\sum \left(\nabla^3 \theta_i\right)^2$.

therefore prompted in practice to use non-centered moving averages to perform these estimations.

Musgrave [55, 56] studied this problem in the context of the X-11 method and proposed a set of asymmetric averages that supplement the Henderson moving averages.

Let us suppose that the series to be smoothed ends in July 1999. If we use a 13-term symmetric moving average, the last smoothed data point will be that of January 1999. It will therefore be necessary to smooth the last six months using asymmetric moving averages. Six months later, in January 2000, it will be possible to calculate the smoothed value for July 1999 using the 13-term symmetric moving average. We will then have two different estimates of this value, and it would be desirable that they not be too different.

Musgrave's idea is precisely to construct asymmetric moving averages that minimize the revisions to the estimates. To this end, he formulates the problem as follows:

- The series to be smoothed can be modelled linearly in the form: $X_t = a + bt + \epsilon_t$ where a and b are constants, and ϵ_t are non-correlated random variables, of average 0 and variance σ^2.
- We are given a set of weights $\{w_1, \ldots, w_N\}$ whose sum is equal to 1 (for example, a centered Henderson moving average) and are looking

for another set of weights $\{v_1, \ldots, v_M\}$, with $M < N$, whose sum is also equal to 1.

- This new moving average must also minimize the revisions to the estimates; that is, it must minimize the criterion:

$$\mathrm{E}\left(\sum_{i=1}^{M} v_i X_i - \sum_{i=1}^{N} w_i X_i\right)^2 .$$

Under these assumptions, it may be shown (Doherty [22], Findley et al. [23]) that the weights may be calculated explicitly as a function of the ratio $D = b^2/\sigma^2$:

$$v_j = w_j + \frac{1}{M} \sum_{i=M+1}^{N} w_i + \frac{\left(j - \frac{M+1}{2}\right) D}{1 + \frac{M(M-1)(M+1)}{12} D} \sum_{i=M+1}^{N} \left(i - \frac{M+1}{2}\right) w_i. \tag{3.1}$$

3.3.1 Musgrave Asymmetric Moving Averages Associated with Henderson Symmetric Moving Averages

The value of D, required in equation (3.1), is unknown, but Musgrave points out that the order of the Henderson moving averages in X-11 is chosen based on the value of the ratio $R = \bar{I}/\bar{C}$ where $\bar{I}$ designates the average of the absolute monthly variations in the irregular portion of the series and $\bar{C}$ designates the average of the absolute monthly variations in the trend-cycle of the series[6]. Assuming the normality of ϵ_t, it can be shown that $D = 4/\left(\pi R^2\right)$, which makes it possible to calculate the asymmetric moving averages numerically.

Tables 3.5 to 3.10 show the coefficients of these moving averages calculated based on the values of the ratio $R = \bar{I}/\bar{C}$ provided in Table 3.3. It is these asymmetric moving averages that are used in X-12-ARIMA. X-11-ARIMA uses the same asymmetric moving averages except for the 5-term Henderson average. These are given in Table 3.4. For deriving these weights, missing observations are forecasted by the average of the last two observations, and the symmetric 5-term Henderson is applied to the extended series.

3.3.2 Comment About Musgrave Moving Averages

While the series to be smoothed is assumed to follow a linear model at its end, Musgrave moving averages do not preserve straight lines, only constant

[6] Refer to the explanation of Table B7, Section 4.1.7, for more details.

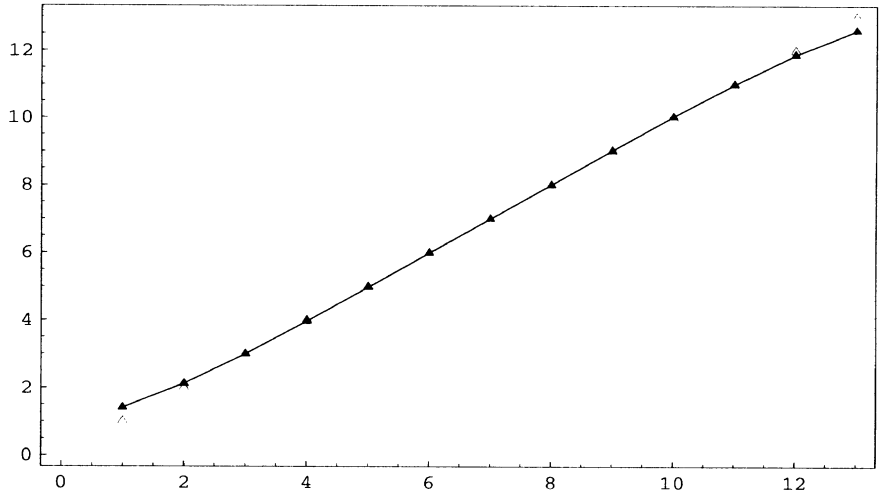


FIGURE 3.9. Smoothing of a straight line by a 13-term Henderson moving average supplemented by Musgrave asymmetric moving averages. The straight line is displayed with empty triangles. The smoothed line is displayed with solid triangles joined by a solid line.

series. In order for it do so, it would also have been necessary to impose on the coefficients the additional constraint $\sum_{i=-p}^{+f} iv_i = 0$.

As Figure 3.9 shows, Musgrave asymmetric averages provide "cautious" estimates of the smoothed series by dampening the growth observed in the last points of the series.

3.3.3 Asymmetric Moving Averages Associated With Composite Moving Averages

Curiously, while Musgrave's first work concerned the generation of asymmetric filters associated with the composite moving averages used to estimate the seasonal factors, his recommendations have not been applied in the X-11 method.

The asymmetric filters associated with the $3 \times 3, 3 \times 5$ and 3×9 averages are shown in Tables 3.11 to 3.13. We do not know the rationale behind these asymmetric filters, and we are not aware of any publication that explains the choice of the coefficients. The 2×4 and 2×12 filters are not supplemented by asymmetric averages.

3.4 The X-11 Moving Average Filter

If the procedure for the detection and correction of extreme values and the procedure for estimating calendar effects are not taken into account,

5-term Henderson	$R = .001$
7-term Henderson	$R = 4.5$
9-term Henderson	$R = 1$
13-term Henderson	$R = 3.5$
23-term Henderson	$R = 4.5$

TABLE 3.3. X-12-ARIMA default values of the ratio $R = \bar{I}/\bar{C}$ used in the computation of Musgrave asymmetric moving averages.

i	$H2\_2$	$H2\_1$	$H2\_0$
-2	-0.07343	-0.073	-0.073
-1	0.29371	0.294	0.403
0	0.55944	0.522	0.670
1	0.29371	0.257	0
2	-0.07343	0	0

TABLE 3.4. Coefficients of the asymmetric moving averages associated with the 5-term Henderson average (X-11-ARIMA). The notation Hp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$H2\_2$	$H2\_1$	$H2\_0$
-2	-0.07343	-0.03671	-0.18357
-1	0.29371	0.29371	0.36713
0	0.55944	0.52273	0.81643
1	0.29371	0.22028	0
2	-0.07343	0	0

TABLE 3.5. Coefficients of Musgrave asymmetric moving averages associated with the 5-term Henderson average (X-12-ARIMA), $R = 0.001$. The notation Hp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$H3\_3$	$H3\_2$	$H3\_1$	$H3\_0$
-3	-0.05874	-0.05314	-0.05421	-0.03379
-2	0.05874	0.05818	0.06101	0.11601
-1	0.29371	0.28699	0.29371	0.38329
0	0.41259	0.39972	0.41032	0.53449
1	0.29371	0.27468	0.28917	0
2	0.05874	0.03356	0	0
3	-0.05874	0	0	0

TABLE 3.6. Coefficients of Musgrave asymmetric moving averages associated with the 7-term Henderson average (X-12-ARIMA), $R = 4.5$. The notation Hp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$H4\_4$	$H4\_3$	$H4\_2$	$H4\_1$	$H4\_0$
-4	-0.04072	-0.03082	-0.02262	-0.04941	-0.15554
-3	-0.00987	-0.00426	-0.00021	-0.01056	-0.03384
-2	0.11847	0.11980	0.11969	0.12578	0.18536
-1	0.26656	0.26361	0.25933	0.28187	0.42429
0	0.33114	0.32391	0.31547	0.35445	0.57972
1	0.26656	0.25504	0.24244	0.29786	0
2	0.11847	0.10267	0.08590	0	0
3	-0.00987	-0.02995	0	0	0
4	-0.04072	0	0	0	0

TABLE 3.7. Coefficients of Musgrave asymmetric moving averages associated with the 9-term Henderson average (X-12-ARIMA), $R = 1$. The notation Hp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$H6\_6$	$H6\_5$	$H6\_4$	$H6\_3$	$H6\_2$	$H6\_1$	$H6\_0$
-6	-0.01935	-0.01643	-0.01099	-0.00813	-0.01603	-0.04271	-0.09186
-5	-0.02786	-0.02577	-0.02204	-0.02019	-0.02487	-0.03863	-0.05811
-4	0	0.00127	0.00330	0.00413	0.00267	0.00182	0.01202
-3	0.06549	0.06594	0.06626	0.06608	0.06784	0.07990	0.11977
-2	0.14736	0.14698	0.14559	0.14441	0.14939	0.17436	0.24390
-1	0.21434	0.21314	0.21004	0.20784	0.21605	0.25392	0.35315
0	0.24006	0.23803	0.23324	0.23002	0.24144	0.29223	0.42113
1	0.21434	0.21149	0.20498	0.20076	0.21540	0.27910	0
2	0.14736	0.14368	0.13547	0.13024	0.14810	0	0
3	0.06549	0.06099	0.05108	0.04483	0	0	0
4	0	-0.00532	-0.01694	0	0	0	0
5	-0.02786	-0.03401	0	0	0	0	0
6	-0.01935	0	0	0	0	0	0

TABLE 3.8. Coefficients of Musgrave asymmetric moving averages associated with the 13-term Henderson average (X-12-ARIMA), $R = 3.5$. The notation Hp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$H11\_11$	$H11\_10$	$H11\_9$	$H11\_8$	$H11\_7$	$H11\_6$
-11	-0.00428	-0.00390	-0.00282	-0.00103	0.00108	0.00268
-10	-0.01092	-0.01059	-0.00968	-0.00817	-0.00642	-0.00511
-9	-0.01569	-0.01542	-0.01467	-0.01344	-0.01205	-0.01103
-8	-0.01453	-0.01431	-0.01372	-0.01279	-0.01175	-0.01101
-7	-0.00495	-0.00479	-0.00436	-0.00372	-0.00303	-0.00258
-6	0.01343	0.01354	0.01380	0.01416	0.01448	0.01465
-5	0.03893	0.03898	0.03908	0.03916	0.03913	0.03900
-4	0.06830	0.06830	0.06823	0.06802	0.06764	0.06723
-3	0.09740	0.09734	0.09711	0.09661	0.09587	0.09517
-2	0.12195	0.12184	0.12144	0.12066	0.11956	0.11858
-1	0.13832	0.13815	0.13759	0.13652	0.13507	0.13380
0	0.14406	0.14384	0.14312	0.14176	0.13995	0.13839
1	0.13832	0.13804	0.13716	0.13551	0.13334	0.13150
2	0.12195	0.12162	0.12057	0.11864	0.11611	0.11399
3	0.09740	0.09701	0.09580	0.09358	0.09070	0.08829
4	0.06830	0.06786	0.06649	0.06398	0.06075	0.05805
5	0.03893	0.03844	0.03690	0.03411	0.03052	0.02753
6	0.01343	0.01288	0.01118	0.00810	0.00415	0.00088
7	-0.00495	-0.00555	-0.00742	-0.01078	-0.01509	0
8	-0.01453	-0.01519	-0.01721	-0.02087	0	0
9	-0.01569	-0.01640	-0.01859	0	0	0
10	-0.01092	-0.01169	0	0	0	0
11	-0.00428	0	0	0	0	0

TABLE 3.9. Coefficients of Musgrave asymmetric moving averages associated with the 23-term Henderson average (X-12-ARIMA), $R = 4.5$, Part A. The notation Hp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$H11\_5$	$H11\_4$	$H11\_3$	$H11\_2$	$H11\_1$	$H11\_0$
-11	0.00258	-0.00065	-0.00861	-0.02293	-0.04520	-0.07689
-10	-0.00519	-0.00776	-0.01396	-0.02486	-0.04130	-0.06385
-9	-0.01109	-0.01300	-0.01744	-0.02491	-0.03554	-0.04893
-8	-0.01106	-0.01230	-0.01500	-0.01904	-0.02385	-0.02808
-7	-0.00261	-0.00319	-0.00413	-0.00475	-0.00373	0.00119
-6	0.01464	0.01472	0.01554	0.01834	0.02518	0.03925
-5	0.03902	0.03976	0.04233	0.04856	0.06121	0.08444
-4	0.06726	0.06866	0.07299	0.08264	0.10112	0.13350
-3	0.09522	0.09729	0.10337	0.11644	0.14074	0.18228
-2	0.11865	0.12137	0.12921	0.14571	0.17583	0.22652
-1	0.13389	0.13728	0.14687	0.16679	0.20273	0.26258
0	0.13850	0.14255	0.15390	0.17724	0.21901	0.28801
1	0.13163	0.13634	0.14945	0.17622	0.22380	0
2	0.11413	0.11951	0.13437	0.16456	0	0
3	0.08845	0.09449	0.11111	0	0	0
4	0.05823	0.06493	0	0	0	0
5	0.02773	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0

TABLE 3.10. Coefficients of Musgrave asymmetric moving averages associated with the 23-term Henderson average (X-12-ARIMA), $R = 4.5$, Part B. The notation Hp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$S2\_2$	$S2\_1$	$S2\_0$
-2	1/9	3/27	5/27
-1	2/9	7/27	11/27
0	3/9	10/27	11/27
1	2/9	7/27	0
2	1/9	0	0

TABLE 3.11. Asymmetric moving averages associated with the 3×3 symmetric moving average. The notation Sp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$S3\_3$	$S3\_2$	$S3\_1$	$S3\_0$
-3	1/15	4/60	4/60	9/60
-2	2/15	8/60	11/60	17/60
-1	3/15	13/60	15/60	17/60
0	3/15	13/60	15/60	17/60
1	3/15	13/60	15/60	0
2	2/15	9/60	0	0
3	1/15	0	0	0

TABLE 3.12. Asymmetric moving averages associated with the 3×5 symmetric moving average. The notation Sp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future.

i	$S5\_5$	$S5\_4$	$S5\_3$	$S5\_2$	$S5\_1$	$S5\_0$
-5	1/27	35/1026	35/1026	33/1026	29/1026	52/1026
-4	2/27	75/1026	77/1026	81/1026	94/1026	115/1026
-3	3/27	114/1026	116/1026	136/1026	148/1026	177/1026
-2	3/27	116/1026	120/1026	136/1026	164/1026	202/1026
-1	3/27	117/1026	126/1026	147/1026	181/1026	227/1026
0	3/27	119/1026	131/1026	158/1026	197/1026	252/1026
1	3/27	120/1026	135/1026	167/1026	213/1026	0
2	3/27	121/1026	141/1026	177/1026	0	0
3	3/27	123/1026	145/1026	0	0	0
4	2/27	86/1026	0	0	0	0
5	1/27	0	0	0	0	0

TABLE 3.13. Asymmetric moving averages associated with the 3×9 symmetric moving average. The notation Sp_f means that the moving average has order $p + f + 1$ with p points in the past and f points in the future. In X-11-ARIMA and X-12-ARIMA, the asymmetric moving averages are coded in decimal form with 3 figures only after the decimal point. The fractional form given here is the one closest to this decimal expression. A good approximation can also be obtained using Musgrave's formula with $D = 9.8$.

the X-11 method may be seen as the successive application of several moving averages. The operator that transforms the unadjusted series into the seasonally adjusted series is therefore itself a moving average.

Thus, in the case of a monthly series, the basic algorithm described in Table 2.1 amounts to the application of a single moving average which can be calculated (Gouriéroux and Monfort [28]).

1. **Estimation of the trend-cycle by a 2×12 moving average:**

$$C^{(1)} = M_{2\times 12}X \text{ with } M_{2\times 12} \ : [13]; \frac{1}{24}\{1, 2, 2, 2, 2, 2, 2\}.$$

2. **Estimation of the seasonal-irregular component:**

$$(S+I)^{(1)} = X - C^{(1)} = [I_d - M_{2\times 12}]\, X$$

where I_d represents the identity operator that transforms the series into itself. Here, I_d would be the moving average $[13]; \{0, 0, 0, 0, 0, 0, 1\}$.

3. **Estimation of the seasonal component by a 3×3 moving average over each month:**

Applying the 3×3 moving average to the values of each month separately amounts to applying, to the series, the average M_3 over 49 months, defined by:

$$M_3 \ : [49]; \frac{1}{9}\{1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3\}.$$

Consequently,

$$\begin{aligned} S^{(1)} &= M_{3\times 3}\,(S+I)^{(1)} \\ &= M_3\,[I_d - M_{2\times 12}]\, X. \end{aligned}$$

The factors are then normalized such that their sum over each consecutive 12-month period is approximately zero.

$$\begin{aligned} \tilde{S}^{(1)} &= S^{(1)} - M_{2\times 12}S^{(1)} \\ &= [I_d - M_{2\times 12}]\, M_3\,[I_d - M_{2\times 12}]\, X \\ &= M_3\,[I_d - M_{2\times 12}]^2\, X. \end{aligned}$$

4. **Estimation of the seasonally adjusted series:**

$$\begin{aligned} A^{(1)} &= X - \tilde{S}^{(1)} \\ &= X - M_3\,[I_d - M_{2\times 12}]^2\, X \\ &= \left(I_d - M_3\,[I_d - M_{2\times 12}]^2\right) X. \end{aligned}$$

5. **Estimation of the trend-cycle by a 13-term Henderson moving average:**

$$\begin{aligned} C^{(2)} &= H_{13}A^{(1)} \\ &= H_{13}\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)X. \end{aligned}$$

6. **Estimation of the seasonal-irregular component:**

$$\begin{aligned} (S+I)^{(2)} &= X - C^{(2)} \\ &= \left[I_d - H_{13}\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)\right]X. \end{aligned}$$

7. **Estimation of the seasonal component by a 3×5 moving average over each month:**

   Applying the 3×5 moving average to the values of each month separately amounts to applying, to the series, the average M_5 over 73 months, defined by:

$$\begin{aligned} M_5 \quad : \quad & [73]; \frac{1}{27}\{1,0,0,0,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0, \\ & 3,0,0,0,0,0,0,0,0,0,0,0,3\}. \end{aligned}$$

   Consequently,

$$\begin{aligned} S^{(2)} &= M_5\,(S+I)^{(2)} \\ &= M_5\left[I_d - H_{13}\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)\right]X. \end{aligned}$$

   The factors are then normalized such that their sum over each consecutive 12-month period is approximately zero.

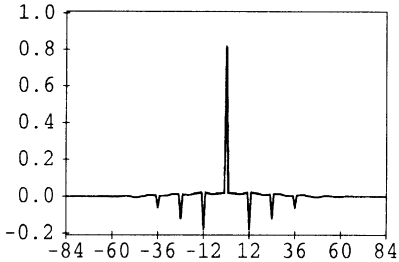
$$\begin{aligned} \tilde{S}^{(2)} &= S^{(2)} - M_{2\times 12}S^{(2)} \\ &= (I_d - M_{2\times 12})\,M_5\left[I_d - H_{13}\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)\right]X. \end{aligned}$$

8. **Estimation of the seasonally adjusted series:**

$$\begin{aligned} A &= X - \tilde{S}^{(2)} \\ &= \{I_d - (I_d - M_{2\times 12})\,M_5\,[I_d- \\ &\quad H_{13}\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)]\}\,X. \end{aligned}$$

The order of this moving average can be calculated step by step:

- Order $[I_d - M_{2\times 12}]^2 = 2\times$ Order $[I_d - M_{2\times 12}] - 1 = 2\times 13 - 1 = 25$,



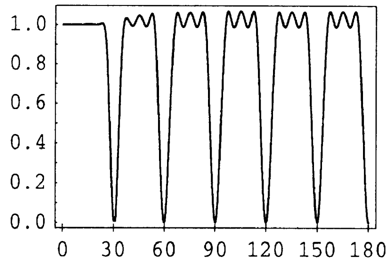


FIGURE 3.10. Coefficient curve and gain function of the symmetric monthly moving average of X-11. The coefficient curve is on the left, the gain function on the right.

- Order $\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)$ = Order $M_3 + 25 - 1 = 49 + 25 - 1 = 73$,
- Order $H_{13}\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)$ = Order $H_{13} + 73 - 1 = 13 + 73 - 1 = 85$,
- Order $M_5\left[I_d - H_{13}\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)\right]$ = Order $M_5 + 85 - 1 = 73 + 85 - 1 = 157$,
- Order $\left\{I_d - (I_d - M_{2\times 12})\, M_5\left[I_d - H_{13}\left(I_d - M_3\left[I_d - M_{2\times 12}\right]^2\right)\right]\right\}$ = Order $(I_d - M_{2\times 12}) + 157 - 1 = 169$.

It is therefore a moving average of order 169 whose coefficient curve and gain function are shown in Figure 3.10. Strictly speaking, one needs to have 84 observations, or 7 years, on either side of a point to be able to use this filter. It must therefore be supplemented with 84 asymmetric moving averages.

Figure 3.11 shows the coefficient curve and the gain function of the central X-11 filter used in the quarterly case. This is a moving average of order 57 and also requires 7 years on either side of a point in order to be used.

Comment

The central symmetric composite X-11 moving averages can be obtained by adjusting a dummy variable with the X-11-ARIMA or X-12-ARIMA softwares. For example, in the monthly case, where the global filter is a moving average of order 169 (when the 13-term Henderson, the 3×3 and the 3×5 moving averages are used), the moving average coefficients are obtained by seasonally adjusting a series of $193 (= 169 + 24)$ observations that are all equal to zero except the 97th observation that is set equal to one. The seasonal adjustment options must be set to an additive decomposition model, specifying both the trend (for instance, Henderson 13) and

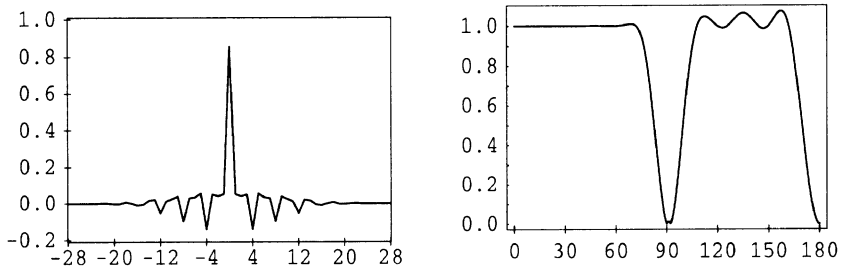


FIGURE 3.11. Coefficient curve and gain function of the symmetric quarterly moving average of X-11. The coefficient curve is on the left, the gain function on the right.

the seasonal moving averages (for instance, the 3×3 and the 3×5) under study, without the automatic correction for extreme observations (for instance, the values for the two sigma limits can be set equal to 9.9). The resulting seasonally adjusted series in Table D11 is the series of coefficients corresponding to the central symmetric moving average. The extra 12 zeroes at the beginning and the end are necessary because of the way X-11 applies the 2×12 moving average in various tables.

Another algorithm needs to be used to obtain the full matrix of asymmetric and symmetric moving averages corresponding to the X-11 linear approximation applied to a given series of length T. In the example above, it first consists in seasonally adjusting each column of the identity matrix of order 169 with the appropriate options. Next, each column of the identity matrix is replaced by its corresponding seasonally adjusted series. The rows of the resulting matrix provide the 169 moving averages with the symmetric one being in row 85. This algorithm works for the following reasons.

Denote by $X = (x_1, \ldots, x_T)$ the series to be seasonally adjusted, by $A = (a_1, \ldots, a_T)$ its seasonally adjusted version, and by W the weight matrix such that $A = WX$. The goal is to find W.

The seasonally adjusted series at time t is $a_t = \sum_{j=1}^{T} w_{t,j}x_j$. Hence, if $X = (1, 0, \ldots, 0)$ is used, then $a_t = w_{t,1}$, and $(w_{1,1}, w_{2,1}, \ldots, w_{T,1})$ represents the first column of W. Similarly, if $X = (0, 1, 0, \ldots, 0)$ is used, then $a_t = w_{t,2}$, and $(w_{1,2}, w_{2,2}, \ldots, w_{T,2})$ represents the second column of W. This is continued until $X = (0, 0, \ldots, 0, 1)$ is used, to obtain $a_t = w_{t,T}$, and $(w_{1,T}, w_{2,T}, \ldots, w_{T,T})$ as the last column of W.

4

The Various Tables

This chapter presents a complete and detailed example of seasonal adjustment with the X-11 method. The series that is used in this example is a monthly series; it is in such cases that the softwares' options are most numerous and complex. The series studied X_t is the monthly index of industrial production in France between October 1985 and March 1995[1]. The series is represented in the top panel of Figure 4.1, which gives a decomposition plot. The seasonal factors S_t (Table D10) are graphed in the third panel, and, in the case at hand, the trading-day factors D_t (Table C18) are provided in the fourth panel. These two sets of factors are used to compute the seasonally adjusted series A_t (Table D11), which is shown with the original series in the top panel, and with the trend-cycle C_t (Table D12) in the second panel. Finally, the irregular component I_t (Table D13) is graphed in the bottom panel. It is obtained by removing the trend-cycle from the seasonally adjusted series. Because, in this example, the decomposition is multiplicative, we can write:

$$
\begin{aligned}
X_t &= C_t \times S_t \times D_t \times I_t, \\
A_t &= C_t \times I_t.
\end{aligned}
$$

[1] The data are given in Table B1 (see Table 4.1).

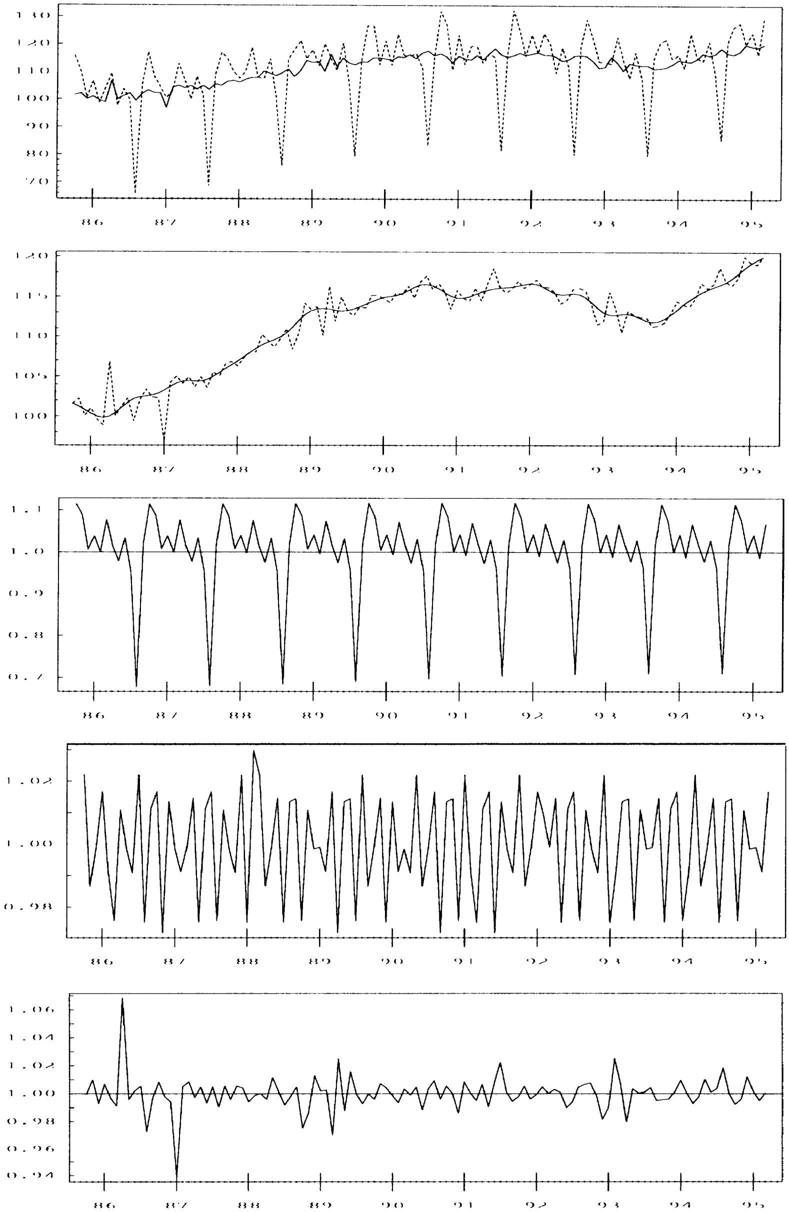


FIGURE 4.1. Decomposition plot for the French Industrial Production Index. The top panel shows both the **original** (dashed line) and **seasonally adjusted** series (solid line). The second panel shows the **seasonally adjusted** series (dashed line) and its **trend-cycle** (solid line). The third panel shows the **seasonal factors**. The fourth panel shows the **trading-day factors**. Finally, the fifth panel displays the **irregulars**. Note the change in the values of the y-axis for the various graphs.

Foreword

1. Emphasis will be placed here on the X-11 portion of the current software, i.e. without reference to *a priori* ARIMA modelling of the series to be seasonally adjusted[2].

2. In the following, we will be referring almost exclusively to the multiplicative decomposition model to describe the tables' contents. However, in formulating the explanations mathematically, we will be using a symbolic language capable of handling both models. We will thus use the following notation:

Symbol	Add. Model	Mult. Model	Meaning
op	$-$	$/$	These first two lines represent the basic operations for each of the two models.
invop	$+$	$\times$	
xbar	0	1	Some estimators, e.g. the seasonal factors, are assumed to have an average of zero in the additive case and an average equal to 1 in the multiplicative case. These average values are used in several places in the computational algorithm of the X-11 method.
mult		$100\times$	In certain tables, and only for the multiplicative case, the estimates are multiplied by 100 and interpreted in terms of percentages.

3. The series X is supposed to have n observations. An observation could be referred to as:

   - X_t where t varies from 1 to n;
   - X_{ij} where $i = 1, \dots, n_j$, $j = 1, \dots, k$, and $k = 4$ or 12. Subscript i refers to the year. It varies from 1 to n_j, the number of observations in month j (or quarter j). Subscript j refers to the period (quarter or month). It varies from 1 to k. The number of period is either $k = 4$ in the quarterly case or $k = 12$ in the monthly case. The total number of observations is $n = \sum_{j=1}^{k} n_j$. When we have only complete years of data, N refers to this number of years. In this case, N is also the common number of observations in a month (or quarter), and $n_j = N, j = 1, \dots, k$.

[2]When an ARIMA model for the series is used, the softwares run on a slightly different principle (see Section 2.7) and most of the calculations we are going to detail are done on a forecasted, and possibly backcasted, series.

4. In its present version, this work describes Parts B, C, D, E and F. Part G (charts) can be replaced by existing graphic software packages[3]. Part A (prior adjustments of the series) differs markedly between X-11-ARIMA and X-12-ARIMA.

5. A number of relatively complex processes are repeated in the X-11 method. Details are provided in the first table where they are implemented. We have in particular:

Process	Tables
Identification and replacement of extreme values	B4, B9, B17, C17
Regression to adjust for trading-day effects	B15, C15
Trend-cycle extraction	B7, C7, D7, D12

6. The output explained in this chapter can be obtained by submitting the following instructions to the X-11-ARIMA and X-12-ARIMA softwares[4]:

X-12-ARIMA:

```
series{data=(115.7 109.8 ... 130.2)
       start= 1985.10
       period= 12
       print=none
       decimals=3}
X11{mode=mult
    print=(all)}
X11regression{variables=td
              print=(all)}
```

X-11-ARIMA:

```
DATA ipi    12 85 10 ;
   (the data)
;
TITLE ipi;
RANGE 12 85  10  95  3  ;
SA (ipi, 0 ,1) TDR 2 00 00 CHART 1 PRTDEC 3 PRINT 5;
END;
```

However, these two programs do not provide exactly the same output; the differences are explained at each step. The data are those of Table B1 (see Table 4.1).

[3]A program named X-12-Graph that uses *SAS/GRAPH*© (SAS Institute [62]) to produce graphics from X-12-ARIMA is distributed with X-12-ARIMA (see Hood [36]).

[4]X-11-ARIMA version 2000 and X-12-ARIMA version 0.2.7.

4.1 PART B: Preliminary Estimation of Extreme Values and Calendar Effects

4.1.1 Table B1: Raw Series or Raw Series Adjusted a priori

Description and method of calculation

This table shows the raw series or the series adjusted *a priori* for elements of Part A. The adjustment factors are the permanent or temporary prior adjustment factors[5] and the trading-day effect calculated from user supplied daily weights.

Comments

- The *a priori* adjustment factors of Part A appear in different tables depending on the software used.

Adjustment factors	X-11-ARIMA	X-12-ARIMA
Permanent monthly adjustments	A2	A2p
Temporary monthly adjustments	A4	A2t
Trading-day effects computed from user-supplied prior daily weights	A6	A4

- Following Table B1, X-11-ARIMA and X-12-ARIMA provide a statistical test for the presence of seasonality. This test is computed from the first estimate of the seasonal-irregular component in Table B3 (see Section 4.1.3).

Example

The original series is given in Table B1 (Table 4.1), and the test for the presence of stable seasonality in Table 4.2.

4.1.2 Table B2: Preliminary Estimation of the Trend-Cycle

Description and method of calculation

The first estimate of the trend-cycle component is obtained by applying a centered moving average of order 12 to the data in Table B1. This moving

[5]With a temporary adjustment, the adjusted series becomes the input on which all transformations will be made. The temporary modifications will be re-introduced in the final seasonally adjusted, trading-day adjusted and (or) Easter adjusted data.

With a permanent adjustment, the new series is thus permanently modified and becomes the input on which all other transformations are made.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	115.700	109.800	100.600
1986	106.600	98.700	103.900	109.500	97.700	103.700	99.700	65.700	105.200	117.100	108.300	104.400
1987	100.500	103.200	112.900	107.100	100.000	108.300	101.800	68.700	108.700	116.900	114.700	110.000
1988	107.700	110.200	118.700	108.100	107.400	114.700	101.200	76.000	114.600	117.900	121.300	114.700
1989	117.900	112.200	120.200	114.700	110.500	120.300	105.600	79.400	114.200	126.700	126.800	112.700
1990	121.100	112.500	123.600	116.100	115.600	116.800	111.800	83.300	114.600	132.000	127.100	110.800
1991	123.300	112.800	119.300	119.400	113.300	116.700	115.300	81.600	116.400	132.400	124.800	115.800
1992	123.500	116.900	124.000	120.000	109.800	118.700	112.100	80.000	119.300	129.000	122.100	113.800
1993	113.700	113.100	122.700	114.200	107.900	117.100	108.100	79.700	114.800	121.000	121.700	114.800
1994	116.300	111.500	124.000	115.400	114.000	121.000	109.500	85.400	120.600	126.400	127.700	120.000
1995	124.100	116.300	130.200	.	.	.	.	.	.	.	.	.

TABLE 4.1. B1: Original series.

	SS	DF	RMSE	F	PROB>F
Inter-month	10897.091	11	990.645	183.698	0.000
Residual	485.351	90	5.393		
Total	11382.442	101			

TABLE 4.2. Test for the presence of stable seasonality.

average, which by design eliminates constant monthly seasonality in an additive composition model, is a composite of a simple 12-term moving average (i.e. of coefficients 1/12) and of a simple 2-term moving average used to "re-centre" the result. The moving average used is therefore a 13-term 2×12 moving average of coefficients $\{1,2,2,2,2,2,2,2,2,2,2,2,1\}/24$.

Comments

- X-11-ARIMA and X-12-ARIMA also include a centered 24-term moving average due to Cholette [13].
- The first six and last six points in the series, for which no trend-cycle estimate can be obtained because of the symmetry of the moving average, are not imputed at this stage of the calculations.

Example

The trend-cycle value for April 1986, the first that can be calculated, is obtained from the values in Table B1 between October 1985 and October

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**101.458**	101.454	101.550	101.454	101.388	101.950	102.225	102.221	102.508
1987	102.788	103.000	103.271	103.408	103.667	104.167	104.700	105.292	105.825	106.108	106.458	107.033
1988	107.275	107.554	108.104	108.392	108.708	109.179	109.800	110.308	110.454	110.792	111.196	111.558
1989	111.975	112.300	112.425	112.775	113.371	113.517	113.567	113.713	113.867	114.067	114.338	114.404
1990	114.517	114.938	115.117	115.354	115.588	115.521	115.533	115.638	115.471	115.429	115.471	115.371
1991	115.513	115.588	115.592	115.683	115.604	115.717	115.933	116.113	116.479	116.700	116.579	116.517
1992	116.467	116.267	116.321	116.300	116.046	115.850	115.358	114.792	114.579	114.283	113.963	113.817
1993	113.583	113.404	113.204	112.683	112.333	112.358	112.508	112.550	112.538	112.642	112.946	113.363
1994	113.583	113.879	114.358	114.825	115.300	115.767	116.308	116.833	117.292	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.3. B2: Trend-cycle, centered 12-term moving average.

1986 (six months before and six months after):

$$\begin{aligned} APR86 &= \frac{115.7}{24} + \frac{109.8 + 100.6 + 106.6 + 98.7 + 103.9 + 109.5}{12} + \\ & \frac{97.7 + 103.7 + 99.7 + 65.7 + 105.2}{12} + \frac{117.1}{24} \\ &= 101.458. \end{aligned}$$

4.1.3 Table B3: Preliminary Estimation of the Unmodified Seasonal-Irregular Component

Description and method of calculation

The trend-cycle component is removed from the analysed series by subtraction or division depending on the decomposition model used, and the result is a first estimate of the seasonal-irregular (SI) component. We thus have: $B3 = B1 \ op\ B2$.

This table forms the basis for the *stable seasonality test* appearing after Table B1. This is a one-way analysis of variance test; we have k samples (in this case the seasonal-irregular estimates for each of our $k = 12$ months or $k = 4$ quarters) of size $n_1, n_2, \ldots, n_k$ respectively. Each of these samples corresponds to a different level of factor A[6], in this case the *seasonality*. This factor is assumed to affect only the means of the distribution and not their variance. This then is a test for the equality of k averages $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$. Assuming that each sample is derived from a random variable X_j following a normal distribution with mean m_j and standard deviation σ, the problem is to test:

$$\begin{aligned} H_0 &: m_1 = m_2 = \ldots = m_k \\ H_1 &: m_p \neq m_q \text{ for at least one pair } (p, q). \end{aligned}$$

The so-called *analysis of variance* equation is written as follows:

$$\sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 = \sum_{j=1}^{k} n_j (\bar{x}_{.j} - \bar{x}_{..})^2 + \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_{.j})^2$$

or

$$S^2 = S_A^2 + S_R^2.$$

The total variance is therefore broken down into a variance of the averages, due to the seasonality, and a residual variance. If hypothesis H_0 holds true, it can be shown that the quantity

$$F_S = \frac{S_A^2/(k-1)}{S_R^2/(n-k)}$$

[6] Letters "X" and "A" were already used to denote the original series and the seasonally adjusted series. Do not confuse.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**107.926**	96.300	102.117	98.271	64.801	103.188	114.551	105.947	101.845
1987	97.775	100.194	109.324	103.570	96.463	103.968	97.230	65.247	102.717	110.170	107.742	102.772
1988	100.396	102.460	109.802	99.731	98.796	105.057	92.168	68.898	103.753	106.416	109.087	102.816
1989	105.291	99.911	106.916	101.707	97.468	105.976	92.985	69.825	100.293	111.075	110.900	98.510
1990	105.749	97.879	107.369	100.647	100.011	101.107	96.769	72.035	99.246	114.356	110.071	96.038
1991	106.742	97.588	103.208	103.213	98.007	100.850	99.454	70.277	99.932	113.453	107.052	99.385
1992	106.039	100.545	106.602	103.181	94.618	102.460	97.175	69.691	104.120	112.877	107.141	99.985
1993	100.103	99.732	108.388	101.346	96.053	104.220	96.082	70.813	102.010	107.420	107.751	101.268
1994	102.392	97.911	108.431	100.501	98.873	104.521	94.146	73.096	102.821	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.4. B3: Unmodified seasonal-irregular (SI) ratios.

follows an F distribution $F(k-1; n-k)$ with $k-1$ and $n-k$ degrees of freedom. Hence the following test: if the statistic F_S calculated on data in Table B3 is greater than the critical value of an F distribution, we conclude there is significant seasonality (i.e. the months' or quarters' averages are not all equal).

Comments

- There is no estimate for the first six and last six values of the series.
- Some assumptions of the analysis of variance model used to test for the presence of seasonality are undoubtedly violated. For instance the irregular component may be autocorrelated at this stage of the analysis. It is therefore customary to choose high critical values and to reject the null hypothesis only if the probability associated with the value of the statistic F_S is less than $1/1000 = 0.001$.
- Rejection of the null hypothesis suggests the existence of seasonality. Acceptance of the null hypothesis does not mean that there is no seasonality; it may be that the seasonality is so distorted over time that we cannot statistically distinguish between the means.
- This test is not part of the default output of X-12-ARIMA.

Example

The value for April 1986, in Table B3, is obtained simply as:

$$APR86 = 100 \times 109.500/101.458 = 107.926.$$

Since the value of the statistic F_S, from Table 4.2, is very high, the hypothesis of equality of the seasonal means is rejected. As Figure 4.2 indicates, this is due to the fact that the mean for the month of August is much smaller than the others.

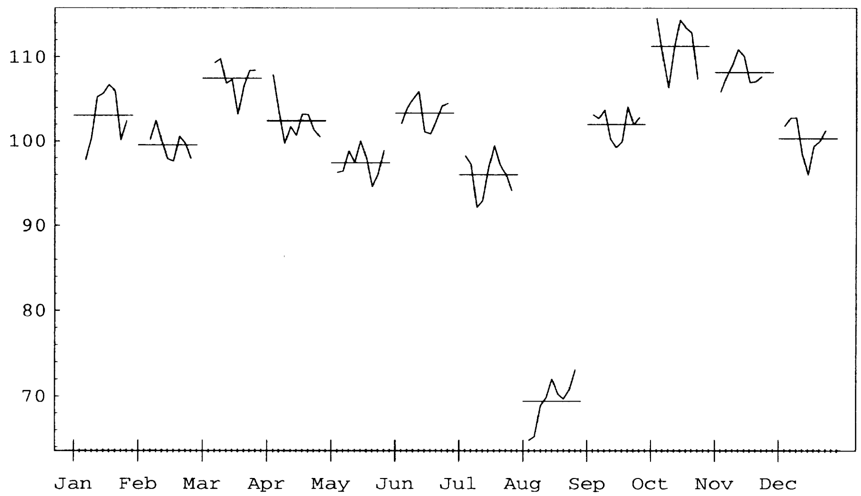


FIGURE 4.2. Monthly subplots of the seasonal-irregular (SI) component. This graph shows the 12 columns of Table B3: the first subplot shows the SI component for January between 1987 and 1994, while the straight line represents the average for these values; the other curves represent the other months.

4.1.4 Table B4: Replacement Values for Extreme Values of the Seasonal-Irregular Component

Description and method of calculation

This table shows the results of the automatic identification and replacement procedure for "extreme" points of the seasonal-irregular component present in the X-11 method. It includes the proposed replacement values for the detected extreme points and, in its margin, the moving standard deviations used to identify them.

This is certainly the most difficult table to reconstruct given that it is the result of a fairly complex algorithm comprising six steps.

Step 1: Estimating the seasonal component.

This seasonal component is estimated by smoothing the seasonal-irregular component one month at a time; first we smooth the values corresponding to the month of January, then the values corresponding to February, and so on, using a 3×3 moving average with coefficients $\{1, 2, 3, 2, 1\}/9$. The symmetric moving average has five terms and therefore does not allow for the estimations of the seasonal factors of the first and last two years. These are then calculated by means of ad hoc asymmetric moving averages shown in Table 3.11. The result then is a series of preliminary seasonal factors denoted *fspro*.

Step 2: Normalizing the seasonal factors.

The preliminary seasonal factors are then normalized in such a way that, for one year of observations, their average is roughly equal to zero (for an additive model) or to unity (for a multiplicative model). To this end, a centered 12-term moving average, i.e. 2×12, of the series *fspro*, written $M_{2\times12}(fspro)$ is calculated; the six missing values at the beginning (end) of the series are considered equal to the first (last) value calculated with this moving average. The *fsnorm* series of normalized factors is then defined as: $fsnorm = fspro\ op\ M_{2\times12}(fspro)$.

Step 3: Estimating the irregular component.
The initial normalized seasonal factors are removed from the seasonal-irregular component to provide an estimate of the irregular component: $Irreg = SI\ op\ fsnorm$.

Step 4: Calculating a moving standard deviation.
A moving standard deviation of the irregular component is calculated at five-year intervals. Each standard deviation is associated with the central year used to calculate it. The values in this central year which, in absolute value, deviate from the average *xbar* by more than 2.5 standard deviations are considered "extremes" and assigned zero weight. The moving standard deviation is then recalculated excluding these values, resulting in a more robust estimate.

For the first two years, the comparisons are based on the standard deviation associated with the third year. For the last two years, the standard deviation of the year two years away from the end is used.

Step 5: Detecting extreme values and weighting the irregular.
Each value of the irregular component is assigned a weight, function of the standard deviation associated with that value, calculated as follows (see Figure 4.3):

- Values which are more than 2.5σ away in absolute value from the average *xbar* are assigned zero weight.
- Values which are less than 1.5σ away in absolute value from the average *xbar* are assigned a weight equal to 1.
- Values which lie between 1.5σ and 2.5σ in absolute value from the average *xbar* are assigned a weight that varies linearly between 0 and 1, depending on their position.

Step 6: Adjusting extreme values of the seasonal-irregular component.
A value of the seasonal-irregular component whose irregular does not receive full weight and which is therefore considered extreme is adjusted and replaced by a weighted average of five values:

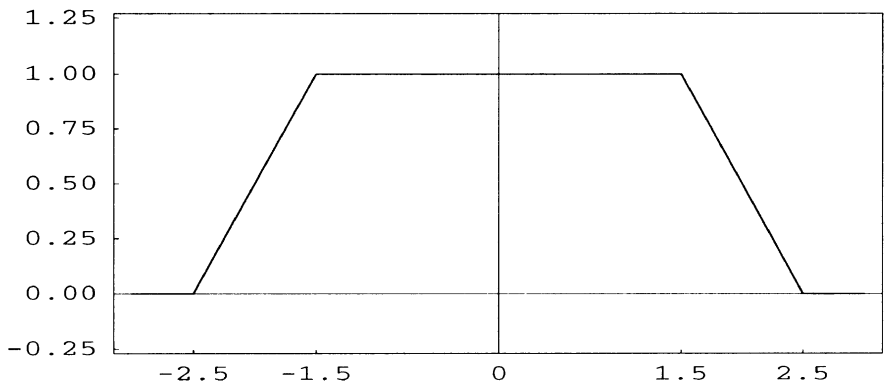


FIGURE 4.3. X-11 Weight function for replacement of extreme values.

- the value itself with its assigned weight,
- the two preceding values, for the same month, having full weight,
- and the next two values, for the same month, having full weight.

For the first two years and the last two years, the replacement values are calculated as the weighted average of the value in question and the four nearest values, for the same month, assigned a weight of 1[7].

Comments

About moving averages:

- We can always select the moving average used. In such a case, X-11-ARIMA provides the choice between a simple 3-term, a 3×3, a 3×5, a 3×9 and constant seasonality (a simple average). X-12-ARIMA also includes a 3×15.

- The asymmetric moving averages used to complete the chosen moving average differ slightly between the various softwares because of rounding problems. The coefficients of these averages are given in Chapter 3.

About calculating weights and adjusting extreme values:

- The extreme nature of a value I_t of the irregular is determined by comparing the value of $|I_t - xbar|$ to bounds $\lambda_L \times \sigma_t$ and $\lambda^U \times \sigma_t$

[7] At the beginning and end of the series, i.e. in the first two years and the last two years, Census X-11 uses the three nearest values of full weight to correct an extreme point (and not the four nearest ones as in X-11-ARIMA or X-12-ARIMA). Moreover, there are some computational errors here in the Census X-11 software.

where σ_t is the standard deviation for the year of the observation I_t and λ_L and λ^U are parameters that can be set by the user (by default 1.5 and 2.5).

- In the absence of four points in the same month with full weight, the extreme value is replaced by the average of the values for the month.

About calculating moving standard deviations:

- The standard deviation is calculated[8] by assuming that the average of the irregular $xbar$ is known, i.e. 0 for an additive model and 1 for a multiplicative model. In this case, an unbiased estimator of the variance is given as: $\sigma^2 = \sum_{t=1}^{n} (I_t - xbar)^2 / n$, where n is the number of observations used (see below).

  It should be recalled that we are trying to detect extreme values. If an estimate of the average was used, it might be strongly influenced by these same extreme values.

- In calculating the second estimate of the standard deviation corresponding to a year, we exclude values the irregulars meeting the condition $|I_t - xbar| > \lambda^U \times \sigma_t$ where σ_t is the first estimate of the standard deviation for the year of the observations I_t.

- In principle, the standard deviation is calculated over five complete years of observations. At the beginning and end of the series, there is a slight exception, mostly because of missing values generated by the use of symmetric moving averages.

  Thus, if the raw series begins in January 1970, the first estimate of the seasonal-irregular component in Table B3 will start in July 1970. For X-11-ARIMA and X-12-ARIMA, the standard deviation for 1972 will be calculated on the basis of 1970 observations and of the first five complete years (from 1971 to 1975), i.e. on the basis of 66 observations. This is the standard deviation that will then be assigned to the years 1970 and 1971. The standard deviation for 1973 will be calculated using fewer observations, i.e. the 60 observations corresponding to the years 1971 to 1975[9].

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	StD
1985	.	.	.	.	.	.	.	.	.	.	.	.	1.427
1986	.	.	.	102.584	.	.	.	.	.	112.451	.	.	1.427
1987	**103.375**	.	.	.	.	.	.	.	.	.	.	.	1.427
1988	.	.	.	101.798	.	.	95.684	.	.	112.038	.	.	1.427
1989	.	.	.	.	.	**103.387**	.	.	.	.	.	.	1.371
1990	.	.	.	.	.	.	.	70.119	.	.	.	99.580	1.396
1991	.	.	106.783	.	.	.	96.339	.	.	.	.	.	1.294
1992	.	.	.	.	97.354	.	.	.	101.594	.	.	.	1.285
1993	104.841	.	.	.	.	.	.	.	.	112.788	.	.	1.285
1994	.	.	.	.	98.075	.	.	70.649	.	.	.	.	1.285
1995	.	.	.	.	.	.	.	.	.	.	.	.	1.285

TABLE 4.5. B4: Replacement values for extreme seasonal-irregular ratios.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**104.634**	96.829	103.416	96.717	65.741	103.101	111.260	107.260	102.403
1987	100.235	101.065	109.073	**103.497**	97.137	103.993	95.716	66.587	102.838	110.433	107.976	102.070
1988	101.580	100.703	108.659	**102.035**	97.840	104.254	94.664	68.186	102.190	110.072	108.951	100.988
1989	103.631	99.799	107.513	101.407	98.276	103.897	94.835	69.652	101.058	110.933	109.534	99.488
1990	105.305	99.071	106.529	101.632	98.266	102.721	96.058	70.544	100.673	112.380	109.149	98.523
1991	105.466	98.806	105.874	102.261	97.422	102.209	97.257	70.547	100.981	112.591	108.248	98.887
1992	104.438	99.118	106.533	102.201	96.762	102.572	97.057	70.709	102.035	111.776	107.601	99.724
1993	102.973	99.232	107.361	**101.810**	96.629	103.467	96.238	71.054	102.537	110.761	107.373	100.397
1994	102.135	99.140	108.075	**101.342**	96.936	104.017	95.496	71.535	102.731	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.6. B4a: Preliminary seasonal factors (3x3 ma).

Example

The X-11 output table is shown in Table B4. To better understand it, we will describe in detail the steps used to calculate it on the basis of tables that are unfortunately not printed in the current versions of the X-11 software family (Tables numbered here B4a to B4f).

Step 1: Estimating the seasonal component.
The data in Table B3 are smoothed column by column (month by month), using a 3×3 moving average of coefficients $\{1,2,3,2,1\}/9$ to produce Table B4a. Thus, the available values of the seasonal-irregular component for the months of April 1986 to 1994 are as follows (from Table B3):

$$107.926, 103.570, 99.731, 101.707, 100.647, 103.213, 103.181, 101.346, 100.501.$$

The seasonal factor for the month of April 1988 is therefore estimated as follows:

$$APR88 = \frac{107.926 + 2\times 103.570 + 3\times 99.731 + 2\times 101.707 + 100.647}{9}$$

[8] The standard deviation is calculated from the irregulars including the extremes. In the old Census X-11, the moving standard deviation was calculated from the irregulars with the extremes in previous year replaced.

[9] Note that, for the Census X-11 software, the first five-year standard deviation that can be calculated is the one for 1973, and its value will likewise be assigned to the years 1970, 1971 and 1972. The first six observations will therefore not be included in the calculation.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.097	100.097	100.097	100.097	100.097	100.097	**100.097**	100.062	100.099
1987	100.082	100.075	100.099	100.054	100.050	100.066	100.108	100.149	100.116	100.038	100.007	100.047
1988	100.014	100.037	100.076	100.034	100.060	100.055	100.096	100.144	100.058	99.984	99.976	99.979
1989	99.972	100.040	100.054	100.043	100.103	100.064	100.072	100.111	100.040	100.008	100.017	99.968
1990	99.970	100.058	100.079	100.123	100.168	100.111	100.078	100.073	100.035	100.034	100.025	99.969
1991	99.997	100.047	100.060	100.082	100.053	100.031	100.003	99.973	100.014	100.039	100.009	99.996
1992	100.003	100.001	100.052	100.062	100.001	100.009	99.983	99.927	99.966	99.984	99.962	99.994
1993	99.997	99.977	100.013	99.991	99.939	99.958	99.951	99.912	99.938	99.948	99.942	99.977
1994	99.969	99.958	99.987	99.987	99.987	99.987	99.987	99.987	99.987	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.7. B4b: Centered 12-month moving average.

$$= \quad 102.035.$$

This symmetric moving average can be used to estimate the seasonal factors for the years 1988 to 1992. For the beginning of the series (years 1986 and 1987) and the end of the series (years 1993 and 1994), predefined asymmetric averages are used (see Table 3.11):

$$APR86 = \frac{107.926 \times 11 + 103.570 \times 11 + 99.731 \times 5}{11} = 104.634$$

(the current point and two points in the future)

$$APR87 = \frac{107.926 \times 7 + 103.570 \times 10 + 99.731 \times 7 + 101.707 \times 3}{27} = 103.497$$

(one point in the past, the current point and two points in the future)

$$APR93 = \frac{100.501 \times 7 + 101.346 \times 10 + 103.181 \times 7 + 103.213 \times 3}{27} = 101.810$$

(one point in the future, the current point and two points in the past)

$$APR94 = \frac{100.501 \times 11 + 101.346 \times 11 + 103.181 \times 5}{27} = 101.342$$

(the current point and two points in the past).

Step 2: Normalizing the seasonal factors.

From Table B4a, a centered 12-month moving average is used to produce Table B4b. The first computable term is therefore that for October 1986, and the last that for March 1994. Thus:

$$\begin{aligned} OCT86 &= \frac{104.634}{24} + \frac{96.829 + 103.416 + 96.717 + 65.741 + 103.101}{12} \\ &\quad + \frac{111.260 + 107.260 + 102.403 + 100.235 + 101.065 + 109.073}{12} \\ &\quad + \frac{103.497}{24} \\ &= 100.097. \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**104.532**	96.735	103.315	96.623	65.678	103.001	111.152	107.193	102.301
1987	100.153	100.989	108.964	103.441	97.089	103.925	95.613	66.488	102.719	110.391	107.969	102.022
1988	101.566	100.666	108.577	102.001	97.782	104.196	94.573	68.088	102.131	110.089	108.977	101.009
1989	103.660	99.759	107.456	101.363	98.176	103.830	94.767	69.574	101.018	110.924	109.515	99.521
1990	105.337	99.014	106.445	101.507	98.102	102.607	95.984	70.493	100.638	112.342	109.122	98.554
1991	105.469	98.760	105.811	102.177	97.370	102.178	97.254	70.566	100.967	112.548	108.238	98.890
1992	104.435	99.117	106.478	102.138	96.761	102.563	97.074	70.761	102.070	111.794	107.642	99.730
1993	102.976	99.255	107.347	101.819	96.688	103.511	96.285	71.117	102.600	110.818	107.435	100.419
1994	102.166	99.182	108.089	101.355	96.949	104.031	95.509	71.545	102.745	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.8. B4c: Normalized seasonal factors.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**103.246**	99.550	98.840	101.705	98.665	100.182	103.058	98.838	99.555
1987	97.625	99.213	100.330	100.125	99.355	100.041	101.692	98.134	99.998	99.800	99.790	100.735
1988	98.848	101.782	101.128	97.775	101.038	100.826	97.456	101.189	101.589	96.663	100.101	101.789
1989	101.573	100.152	99.498	100.339	99.279	102.066	98.120	100.360	99.282	100.137	101.264	98.985
1990	100.391	98.854	100.869	99.152	101.946	98.539	100.818	102.189	98.617	101.793	100.870	97.447
1991	101.207	98.814	97.540	101.014	100.654	98.701	102.261	99.590	98.974	100.804	98.904	100.500
1992	101.535	101.441	100.116	101.022	97.785	99.900	100.105	98.488	102.009	100.969	99.534	100.256
1993	97.210	100.481	100.970	99.535	99.344	100.685	99.789	99.573	99.425	96.934	100.294	100.845
1994	100.221	98.719	100.316	99.157	101.984	100.471	98.574	102.168	100.074	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.9. B4d: Preliminary irregular component.

The first six values, from April to September 1986, which cannot be computed using this symmetric moving average, are considered equal to the first computable value, that for October 1986. The same procedure is used for the end of the series; the value calculated for March 1994 (99.987) is repeated for the next six months. Normalized seasonal factors are then obtained by dividing Table B4a by Table B4b, to produce Table B4c. For example:

$$APR86 = 100 \times 104.634/100.097 = 104.532.$$

Step 3: Estimating the irregular component

We simply divide the seasonal-irregular component of Table B3 by the normalized seasonal factors of Table B4c to obtain Table B4d. For example:

$$APR86 = 100 \times 107.926/104.532 = 103.246.$$

Step 4: Calculating a moving standard deviation.

The standard deviation corresponding to the year 1989 is calculated on the basis of data for the years 1987 to 1991 (two years before, two years

Year	Standard Deviation 1	Standard Deviation 2
1985		
1986	1.4265	1.4265
1987	1.4265	1.4265
1988	1.4265	1.4265
1989	1.3705	1.3705
1990	1.3958	1.3958
1991	1.2941	1.2941
1992	1.2847	1.2847
1993	1.2847	1.2847
1994	1.2847	1.2847
1995		

TABLE 4.10. B4c: Five-year moving standard deviations.

after) using the formula[10]:

$$\sigma_{89} = \left[\frac{1}{60} \sum_{t=Jan87}^{Dec91} (I_t - 100)^2 \right]^{1/2} = 1.3705.$$

Those for the years 1990 and 1991 are calculated using the same principle.

For X-11-ARIMA and X-12-ARIMA, the standard deviation for 1988 is calculated on the basis of all available observations from 1986 to 1991, i.e. 69 observations. This is the standard deviation that will be associated with the years 1986 and 1987. Likewise, the standard deviation for 1992 uses all the data from 1989 to 1994, and will be associated with the years 1993 and 1994. These first estimates of the standard deviations are shown in Table B4e, in the Standard Deviation 1 column.

They are used to detect possible extreme values. For a given year, a value is considered extreme if, in absolute value, it deviates by more than 2.5 times the standard deviation corresponding to that year from its theoretical average (here 1). Figure 4.4 shows the deviation of the irregular from its theoretical mean and the two detection thresholds. As can be seen, no value is considered very extreme. Otherwise, the values would be removed, and the standard deviation would be recalculated (for an example, see Section 4.1.17). This new calculation results in the Standard Deviation 2 column in Table B4e, identical to the first column.

Step 5: Detecting extreme values and weighting the irregular.

Figure 4.4 shows where the values for the irregular are in relation to the upper and lower detection thresholds calculated on the basis of the previously estimated standard deviations. No value is considered very extreme. All values between the lower and upper detection thresholds are considered extreme, and are therefore adjusted; the weights (multiplied by 100) associated with each of these values are shown in Table B4f.

[10] Here the theoretical mean is assumed to be 100 as the irregular values have been multiplied by 100.

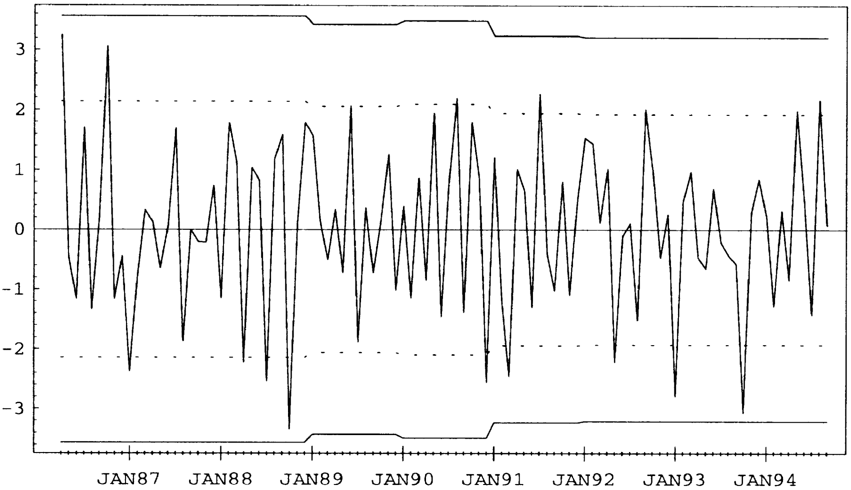


FIGURE 4.4. B4d: Deviation of the irregular component from its theoretical average with the detection thresholds corresponding to $\pm 1.5\sigma$ and $\pm 2.5\sigma$.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	22.419	100.000	100.000	100.000	100.000	100.000	35.633	100.000	100.000
1987	**83.535**	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	94.011	100.000	100.000	71.692	100.000	100.000	16.096	100.000	100.000
1989	100.000	100.000	100.000	100.000	100.000	**99.217**	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	93.192	100.000	100.000	100.000	67.113
1991	100.000	100.000	59.928	100.000	100.000	100.000	75.258	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	77.601	100.000	100.000	100.000	93.624	100.000	100.000	100.000
1993	32.820	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	11.350	100.000	100.000
1994	100.000	100.000	100.000	100.000	95.570	100.000	100.000	81.282	100.000	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.11. B4f: Weights associated with values of the irregular.

We have, e.g. for January 1987: $|Jan87 - 100| = |97.625 - 100| = 2.375$, and $1.5 \times \sigma_{87} = 1.5 \times 1.4265 = 2.13975 < 2.375 < 2.5 \times \sigma_{87} = 2.5 \times 1.4265 = 3.5663$.

This value, considered moderately extreme, is assigned a weight proportional to the deviation from its average, which is:

$$weight(Jan87) = \frac{3.5663 - 2.375}{3.5663 - 2.1398} = 0.835.$$

Likewise, for the June 1989 value, slightly above the lower confidence limit, we have: $|Jun89 - 100| = |102.066 - 100| = 2.066$, and $1.5 \times \sigma_{89} = 1.5 \times 1.3705 = 2.056 < 2.066 < 2.5 \times \sigma_{87} = 2.5 \times 1.3705 = 3.426$. Hence,

$$weight(Jun89) = \frac{3.426 - 2.066}{3.426 - 2.056} = 0.992.$$

Step 6: Adjusting extreme values of the seasonal-irregular component.

Finally, the seasonal-irregular component (Table B3) is adjusted on the basis of these weights. Thus, the value for June 1989 is replaced by the average of this value with its assigned weight and the two preceding and subsequent values for the same month receiving full weight, i.e. not considered extreme. As can be seen from Table B4f, these are the values for the months of June 1987, 1988, 1990 and 1991. This leads to:

$$\begin{aligned} SI(Jun89) &= \frac{103.968 + 105.057 + 105.976 \times 0.992 + 101.107 + 100.850}{4 + 0.992} \\ &= 103.387. \end{aligned}$$

The value for January 1987 is also considered extreme. However, as it appears at the beginning of the series, it is adjusted differently: it is replaced by the average of this value with its assigned weight and the four nearest values for the same month receiving full weight. In this case then, according to Table B4f. These are the values for January 1988, 1989, 1990 and 1991. And so we have:

$$\begin{aligned} SI(Jan87) &= \frac{97.775 \times 0.835 + 100.396 + 105.291 + 105.749 + 106.742}{4 + 0.835} \\ &= 103.375. \end{aligned}$$

These replacement values are printed in Table B4 presented at the beginning of the example.

4.1.5 Table B5: Estimation of the Seasonal Component

Description and method of calculation

The estimation is done on the basis of the values for the seasonal-irregular component in Table B3, corrected with the values in Table B4. We follow three steps; the first two steps are identical to those in Table B4:

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	102.584	96.300	102.117	98.271	64.801	103.188	112.451	105.947	101.845
1987	**103.375**	100.194	109.324	103.570	96.463	103.968	97.230	65.247	102.717	110.170	107.742	102.772
1988	100.396	102.460	109.802	101.798	98.796	105.057	95.684	68.898	103.753	112.038	109.087	102.816
1989	105.291	99.911	106.916	101.707	97.468	**103.387**	92.985	69.825	100.293	111.075	110.900	98.510
1990	105.749	97.879	107.369	100.647	100.011	101.107	96.769	70.119	99.246	114.356	110.071	99.580
1991	106.742	97.588	106.783	103.213	98.007	100.850	96.339	70.277	99.932	113.453	107.052	99.385
1992	106.039	100.545	106.602	103.181	97.354	102.460	97.175	69.691	101.594	112.877	107.141	99.985
1993	104.841	99.732	108.388	101.346	96.053	104.220	96.082	70.813	102.010	112.788	107.751	101.268
1994	102.392	97.911	108.431	100.501	98.075	104.521	94.146	70.649	102.821	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.12. B4g: Corrected seasonal-irregular component.

- Step 1: Estimating the seasonal component using a 3x3 moving average.
- Step 2: Normalizing the seasonal factors using a centered 12-term moving average.

The third step is:

- **Step 3: Imputing the missing seasonal factors.**

  Seasonal factors for the six missing values at either end of the series, due to the use of the centered 12-term moving average in Table B2, are estimated by the nearest available factor for that particular month.

Comment

We can specify the moving average to be used. In such a case, X-11-ARIMA gives the choice between a simple 3-term, a 3×3, a 3×5, a 3×9 and constant seasonality (a simple average). X-12-ARIMA also includes a 3×15.

Example

The estimation is based on the corrected seasonal-irregular component given in Table B4g.

Step 1: Estimating the seasonal component.

The data in Table B4g are smoothed column by column (month by month), using a 3×3 moving average of coefficients $\{1,2,3,2,1\}/9$, resulting in Table B5a. The seasonal factor for the month of April 1988 is therefore estimated as follows:

$$\begin{aligned} APR88 &= \frac{102.584 + 2\times 103.570 + 3\times 101.798 + 2\times 101.707 + 100.647}{9} \\ &= 102.131. \end{aligned}$$

This symmetric moving average can be used to estimate the seasonal factors for the years 1988 to 1992. For the beginning of the series (years 1986

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**102.840**	96.829	103.416	97.368	65.741	103.101	111.445	107.260	102.403
1987	102.516	101.065	109.073	**102.648**	97.137	103.706	96.627	66.587	102.838	111.346	107.976	102.070
1988	103.032	100.703	108.659	**102.131**	97.840	103.678	95.836	67.973	102.190	111.712	108.951	101.382
1989	104.253	99.799	107.911	101.866	98.276	103.034	95.270	69.226	101.058	112.182	109.534	100.276
1990	105.305	99.071	107.323	101.862	98.570	102.146	95.757	69.906	100.393	113.004	109.149	99.704
1991	105.993	98.806	107.066	102.261	98.030	101.921	96.219	70.121	100.420	113.188	108.248	99.674
1992	105.492	99.118	107.328	102.201	97.585	102.572	96.365	70.225	101.192	113.168	107.601	100.117
1993	104.728	99.232	107.758	**101.810**	97.132	103.467	95.892	70.420	101.881	112.948	107.373	100.397
1994	104.065	99.140	108.075	**101.342**	97.118	104.017	95.496	70.538	102.263	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.13. B5a: Preliminary seasonal factors (3×3 ma).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.247	100.247	100.247	100.247	100.247	100.247	**100.247**	100.251	100.276
1987	100.258	100.262	100.286	100.271	100.297	100.313	100.321	100.327	100.295	100.256	100.264	100.292
1988	100.258	100.282	100.313	100.301	100.357	100.369	100.392	100.405	100.336	100.294	100.301	100.292
1989	100.242	100.270	100.275	100.248	100.292	100.270	100.268	100.281	100.226	100.202	100.214	100.189
1990	100.172	100.221	100.221	100.228	100.246	100.206	100.211	100.229	100.207	100.213	100.207	100.175
1991	100.185	100.213	100.223	100.232	100.202	100.163	100.141	100.133	100.157	100.166	100.145	100.153
1992	100.186	100.197	100.233	100.265	100.237	100.229	100.215	100.188	100.211	100.212	100.177	100.196
1993	100.213	100.202	100.239	100.258	100.239	100.242	100.226	100.194	100.203	100.197	100.177	100.199
1994	100.206	100.194	100.215	100.215	100.215	100.215	100.215	100.215	100.215	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.14. B5b: Centered 12-month moving average.

and 1987) and the end of the series (years 1993 and 1994), predefined asymmetric averages are used (see Table 3.11):

$$APR86 = \frac{102.584 \times 11 + 103.570 \times 11 + 101.798 \times 5}{27} = 102.840$$

(the current point and two points in the future)

$$APR87 = \frac{102.584 \times 7 + 103.570 \times 10 + 101.798 \times 7 + 101.707 \times 3}{27} = 102.648$$

(one point in the past, the current point and two points in the future)

$$APR93 = \frac{100.501 \times 7 + 101.346 \times 10 + 103.181 \times 7 + 103.213 \times 3}{27} = 101.810$$

(one point in the future, the current point and two points in the past)

$$APR94 = \frac{100.501 \times 11 + 101.346 \times 11 + 103.181 \times 5}{27} = 101.342$$

(the current point and two points in the past).

Step 2: Normalizing the seasonal factors.

On Table B5a, a centered 12-month moving average is used to produce Table B5b. The first computable term is therefore that for October 1986, and the last that for March 1994. Thus:

$$OCT86 = \frac{102.840}{24}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.171	106.991	102.120
1986	102.253	100.801	108.761	**102.587**	96.590	103.161	97.128	65.580	102.847	111.171	106.991	102.120
1987	102.253	100.801	108.761	102.370	96.850	103.382	96.319	66.370	102.536	111.062	107.692	101.773
1988	102.767	100.419	108.320	101.824	97.492	103.297	95.462	67.699	101.848	111.385	108.624	101.087
1989	104.002	99.530	107.614	101.614	97.991	102.757	95.016	69.032	100.830	111.956	109.301	100.086
1990	105.124	98.853	107.086	101.630	98.328	101.936	95.555	69.746	100.185	112.764	108.924	99.529
1991	105.797	98.596	106.828	102.024	97.832	101.755	96.083	70.028	100.262	113.000	108.091	99.521
1992	105.295	98.924	107.078	101.931	97.355	102.338	96.158	70.093	100.980	112.928	107.411	99.922
1993	104.505	99.033	107.501	101.548	96.900	103.218	95.676	70.284	101.675	112.725	107.183	100.197
1994	103.852	98.948	107.843	101.124	96.910	103.793	95.291	70.387	102.044	112.725	107.183	100.197
1995	103.852	98.948	107.843	.	.	.	.	.	.	.	.	.

TABLE 4.15. B5: Seasonal factors.

$$
\begin{aligned}
& + \frac{96.829 + 103.416 + 97.368 + 65.741 + 103.101}{12} \\
& + \frac{111.445 + 107.260 + 102.403 + 102.516 + 101.065 + 109.073}{12} \\
& + \frac{102.648}{24} \\
= \; & 100.247.
\end{aligned}
$$

The first six values, from April to September 1986, which cannot be computed using this symmetric moving average, are considered equal to the first computable value, that for October 1986. The same procedure is used for the end of the series: the value calculated for March 1994 (100.215) is repeated for the next six months. The normalized seasonal factors are then obtained by dividing Table B5a by Table B5b. Thus

$$APR86 = 100 \times 102.840/100.247 = 102.587.$$

Step 3: Imputing the missing seasonal factors.

Finally, to obtain the seasonal factors for the six missing values at either end of the series (October 1985 to March 1986 and October 1994 to March 1995) due to the use of the centered 12-term moving average in Table B2, we repeat the nearest available factor for that particular month, resulting in Table B5.

4.1.6 Table B6: Estimation of the Seasonally Adjusted Series

Description and method of calculation

The estimation is done simply by removing from the starting series, in Table B1, the estimate of the seasonal component in Table B5: $B6 = B1 \; op \; B5$.

Example

For example, $APR86 = 100 \times 109.500/102.587 = 106.739$.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	104.074	102.626	98.511
1986	104.251	97.916	95.530	**106.739**	101.149	100.522	102.648	100.183	102.288	105.333	101.224	102.232
1987	98.286	102.380	103.805	104.620	103.253	104.757	105.691	103.511	106.011	105.257	106.507	108.084
1988	104.800	109.740	109.583	106.164	110.163	111.039	106.011	112.261	112.520	105.849	111.669	113.467
1989	113.363	112.730	111.695	112.878	112.766	117.072	111.140	115.019	113.260	113.169	116.010	112.603
1990	115.197	113.805	115.421	114.238	117.566	114.582	117.000	119.433	114.388	117.058	116.687	111.324
1991	116.544	114.406	111.675	117.031	115.810	114.687	120.000	116.525	116.096	117.168	115.458	116.357
1992	117.289	118.172	115.804	117.727	112.784	115.988	116.579	114.134	118.143	114.232	113.676	113.889
1993	108.799	114.205	114.138	112.459	111.352	113.449	112.985	113.398	112.909	107.340	113.544	114.574
1994	111.987	112.685	114.982	114.117	117.635	116.578	114.911	121.329	118.184	112.131	119.142	119.764
1995	119.498	117.536	120.731	.	.	.	.	.	.	.	.	.

TABLE 4.16. B6: Seasonally adjusted series.

4.1.7 Table B7: Estimation of the Trend-Cycle Component

Description and method of calculation

This table shows an estimate of the trend-cycle component obtained from the seasonally adjusted series in Table B6. This is a problem of smoothing, and to solve it, the program uses a Henderson moving average.

Step 1: Selecting the moving average, calculating the $\bar{I}/\bar{C}$ ratio.
For a monthly series, X-11 uses at this stage a Henderson moving average (9 or 13 terms). Unless specified by the user, the choice of the order of the moving average is automatic and based on the value of a statistic called $\bar{I}/\bar{C}$ **ratio**, which measures the size of the irregular component in the series; the greater it is, the higher the order of the moving average selected.

In order to calculate this ratio, we compute a first decomposition of the seasonally adjusted series (SA) using a 13-term Henderson moving average. At this stage of the calculation, the six "lost" points at the beginning and end of the series are ignored.

Thus we have: trend-cycle $= C = H_{13}(SA)$ and irregular $= I = SA \; op \; C$. We then calculate, for both the C and I series, the average absolute monthly growth rate (multiplicative model) or the absolute monthly change (additive model), written $\bar{C}$ and $\bar{I}$. Thus we have:

$$
\begin{aligned}
\bar{C} &= \frac{1}{n-1}\sum_{t=2}^{n} |C_t \; op \; C_{t-1} - xbar| \\
\bar{I} &= \frac{1}{n-1}\sum_{t=2}^{n} |I_t \; op \; I_{t-1} - xbar| .
\end{aligned}
$$

We then calculate the $\bar{I}/\bar{C}$ ratio and:

- If the ratio is smaller than 1, a 9-term Henderson moving average is selected.

- Otherwise, a 13-term Henderson moving average is selected.

Step 2: Smoothing the SA series using a Henderson moving average.

The SA series in Table B6 is then smoothed using the selected Henderson moving average. At this stage, the points that cannot be computed using the symmetric average at the beginning and end of the series (4 or 6 as the case may be) are estimated by means of ad hoc asymmetric moving averages.

Comments

- Note that the $\bar{I}/\bar{C}$ ratio is calculated independently of the first six and last six months for which the 13-term Henderson moving average does not provide an estimate of the trend-cycle.

- At this stage, the program chooses only between a 9-term average and a 13-term average.

- It is possible to specify the length of the Henderson moving average to be used. X-11-ARIMA provides a choice between a 9-term, a 13-term or a 23-term moving average. X-12-ARIMA allows any odd-numbered average less than 101.

- The coefficients of the moving averages used (symmetric or not) are, to the nearest rounded digit, identical in X-11-ARIMA and X-12-ARIMA. The coefficients of the symmetric moving averages are calculated using the exact Henderson formula. As for the asymmetric filters, X-11-ARIMA uses 7-decimal place coefficients derived from a formula developed by Laniel [46], and X-12-ARIMA makes direct use of a formula attributed to Doherty [22] and based on the work of Musgrave [55, 56]. These various filters are presented in sections 3.2.2 and 3.3.1[11].

- The estimate of the trend-cycle is calculated without adjusting the seasonally adjusted series for extreme values. X-11-ARIMA proposes an optional adjustment of the trend-cycle for strikes. The effects of extreme values are reduced with an algorithm similar to the one used in Table B4 : remove the estimate of the trend-cycle from the seasonally adjusted series to obtain an estimate of the irregular, compute a five-year standard deviation of the irregular etc. X-12-ARIMA does not propose this option.

---

[11] In Census X-11, the coefficients of the symmetric and asymmetric moving averages are rounded to the third decimal place. The method of calculating the asymmetric filters is not known.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**100.809**	101.258	101.649	102.031	102.287	102.241	102.092	101.939	101.700
1987	101.671	102.029	102.691	103.528	104.218	104.567	104.799	104.992	105.302	105.774	106.319	106.848
1988	107.460	107.972	108.320	108.737	109.126	109.403	109.568	109.760	110.159	110.671	111.282	111.855
1989	112.343	112.722	113.066	113.268	113.389	113.645	113.835	113.913	113.901	113.920	113.989	114.155
1990	114.366	114.521	114.877	115.348	115.889	116.472	116.816	116.862	116.614	116.065	115.337	114.704
1991	114.232	114.160	114.487	115.084	115.871	116.538	116.921	117.008	116.831	116.619	116.632	116.735
1992	116.829	116.824	116.503	116.091	115.767	115.602	115.688	115.709	115.381	114.694	113.877	113.108
1993	112.640	112.448	112.498	112.798	112.965	112.853	112.539	112.212	112.024	111.941	111.996	112.314
1994	112.954	113.648	114.346	115.193	116.069	116.819	117.188	117.307	117.362	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.17. B7a: Trend-cycle (13-term Henderson ma).

Example

Step 1: Selecting the moving average, calculating the $\bar{I}/\bar{C}$ ratio.

First, Table B6 is smoothed using a 13-term Henderson moving average whose coefficients are shown in Table 3.8 (column $H6\_6$).

The first computable term with the symmetric Henderson filter is that for April 1986, and we have:

$$\begin{aligned} APR86 &= 104.074 \times (-0.0193) + 102.626 \times (-0.0279) + \\ &\quad 98.511 \times (0.0000) + 104.251 \times (0.0655) + \\ &\quad 97.916 \times (0.1474) + 95.530 \times (0.2143) + \\ &\quad 106.739 \times (0.2401) + 101.149 \times (0.2143) + \\ &\quad 100.522 \times (0.1474) + 102.648 \times (0.0655) + \\ &\quad 100.183 \times (0.0000) + 102.288 \times (-0.0279) + \\ &\quad 105.333 \times (-0.0193) \\ &= 100.809. \end{aligned}$$

At this step in the calculation, there is no attempt to estimate the six points that cannot be calculated with the symmetric filter at the beginning and end of the series. An estimate is derived for the trend-cycle (Table B7a), and one for the irregular component (Table B7b) by division using Table B6. Thus the irregular value associated with April 86 is:

$$APR86 = 100 \times 106.739/100.809 = 105.882.$$

Since we use a multiplicative model, we calculate:

$$\begin{aligned} \bar{C} &= \frac{1}{n-1} \sum_{t=2}^{n} |C_t/C_{t-1} - 1| \\ \bar{I} &= \frac{1}{n-1} \sum_{t=2}^{n} |I_t/I_{t-1} - 1|. \end{aligned}$$

Using the row totals from Tables B7c and B7d, we have:

$$\bar{C} = \frac{2.033 + 5.008 + 4.589 + 2.060 + 4.207}{101} +$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**105.882**	99.893	98.891	100.605	97.944	100.046	103.174	99.298	100.524
1987	96.670	100.344	101.085	101.055	99.074	100.182	100.851	98.590	100.674	99.510	100.178	101.157
1988	97.525	101.638	101.166	97.634	100.951	101.495	96.753	102.279	102.143	95.643	100.348	101.441
1989	100.908	100.007	98.787	99.656	99.451	103.016	97.632	100.971	99.437	99.341	101.773	98.640
1990	100.727	99.375	100.473	99.037	101.447	98.378	100.158	102.200	98.091	100.856	101.171	97.053
1991	102.024	100.216	97.544	101.692	99.947	98.412	102.634	99.587	99.371	100.471	98.994	99.676
1992	100.394	101.154	99.400	101.409	97.423	100.334	100.770	98.639	102.393	99.598	99.823	100.691
1993	96.590	101.563	101.458	99.700	98.573	100.528	100.396	101.057	100.790	95.890	101.382	102.012
1994	99.144	99.152	100.556	99.066	101.350	99.793	98.057	103.428	100.701	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.18. B7b: Irregular component.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1985	.	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	0.445	0.387	0.375	0.251	0.044	0.146	0.150	0.235	2.033
1987	0.028	0.352	0.649	0.814	0.667	0.335	0.222	0.184	0.295	0.449	0.514	0.498	5.008
1988	0.573	0.477	0.322	0.385	0.358	0.255	0.151	0.175	0.364	0.465	0.552	0.514	4.589
1989	0.437	0.337	0.305	0.178	0.106	0.226	0.168	0.069	0.011	0.017	0.061	0.146	2.060
1990	0.184	0.136	0.311	0.410	0.469	0.503	0.295	0.040	0.212	0.471	0.627	0.549	4.207
1991	0.412	0.063	0.286	0.522	0.684	0.575	0.329	0.075	0.151	0.182	0.011	0.089	3.379
1992	0.081	0.005	0.274	0.354	0.279	0.143	0.075	0.018	0.284	0.596	0.712	0.675	3.496
1993	0.414	0.170	0.045	0.266	0.148	0.099	0.278	0.291	0.167	0.074	0.049	0.284	2.285
1994	0.570	0.615	0.614	0.741	0.760	0.647	0.316	0.102	0.047	.	.	.	4.410
1995	.	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.19. B7c: Absolute growth rates (in %) of the trend-cycle.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1985	.	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	5.657	1.002	1.733	2.645	2.146	3.128	3.757	1.234	21.302
1987	3.833	3.801	0.738	0.029	1.961	1.118	0.668	2.242	2.114	1.156	0.671	0.978	19.307
1988	3.591	4.217	0.464	3.492	3.397	0.539	4.672	5.711	0.133	6.364	4.920	1.090	38.591
1989	0.526	0.893	1.219	0.879	0.205	3.585	5.226	3.420	1.519	0.097	2.448	3.079	23.096
1990	2.116	1.342	1.105	1.429	2.433	3.025	1.810	2.039	4.020	2.818	0.312	4.070	26.519
1991	5.122	1.772	2.666	4.252	1.715	1.536	4.290	2.968	0.218	1.107	1.470	0.689	27.805
1992	0.720	0.758	1.735	2.022	3.931	2.988	0.434	2.115	3.806	2.730	0.227	0.869	22.334
1993	4.073	5.148	0.103	1.733	1.130	1.984	0.131	0.658	0.264	4.862	5.728	0.622	26.435
1994	2.812	0.009	1.416	1.482	2.305	1.536	1.740	5.478	2.637	.	.	.	19.415
1995	.	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.20. B7d: Absolute growth rates (in %) of the irregular.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	**102.405**	101.784	101.095
1986	100.543	100.309	100.463	100.809	101.258	101.649	102.031	102.287	102.241	102.092	101.939	101.700
1987	101.671	102.029	102.691	103.528	104.218	104.567	104.799	104.992	105.302	105.774	106.319	106.848
1988	107.460	107.972	108.320	108.737	109.126	109.403	109.568	109.760	110.159	110.671	111.282	111.855
1989	112.343	112.722	113.066	113.268	113.389	113.645	113.835	113.913	113.901	113.920	113.989	114.155
1990	114.366	114.521	114.877	115.348	115.889	116.472	116.816	116.862	116.614	116.065	115.337	114.704
1991	114.232	114.160	114.487	115.084	115.871	116.538	116.921	117.008	116.831	116.619	116.632	116.735
1992	116.829	116.824	116.503	116.091	115.767	115.602	115.688	115.709	115.381	114.694	113.877	113.108
1993	112.640	112.448	112.498	112.798	112.965	112.853	112.539	112.212	112.024	111.941	111.996	112.314
1994	112.954	113.648	114.346	115.193	116.069	116.819	117.188	117.307	117.362	117.495	117.801	118.258
1995	118.787	119.246	119.901	.	.	.	.	.	.	.	.	.

TABLE 4.21. B7: Trend-cycle, I/C ratio is 7.14, a 13-term Henderson moving average has been selected.

$$
\begin{aligned}
& \frac{3.379 + 3.496 + 2.285 + 4.410}{101} \\
= \; & 0.312,
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{I} &= \frac{21.302 + 19.307 + 35.591 + 23.096 + 26.519}{101} + \\
& \quad \frac{27.805 + 22.334 + 26.435 + 19.415}{101} \\
&= 2.226.
\end{aligned}
$$

Thus $\bar{I}/\bar{C} = 2.226/0.312 = 7.14$.

Step 2: Smoothing the SA series using a Henderson moving average.

Since the ratio is greater than 1, we select a 13-term Henderson moving average whose coefficients and those of the associated asymmetric moving averages are shown in Table 3.8. The trend-cycle for October 1985 is estimated, from the SA series in Table B6, using the current point and six points in the future to which are assigned the coefficients of the $H6\_0$ moving average in Table 3.8.

$$
\begin{aligned}
OCT85 &= 104.074 \times (0.42113) + 102.626 \times (0.35315) + \\
& \quad 98.511 \times (0.24390) + 104.251 \times (0.11977) + \\
& \quad 97.916 \times (0.01202) + 95.530 \times (-0.05811) + \\
& \quad 106.739 \times (-0.09186) \\
&= 102.405.
\end{aligned}
$$

This leads to Table B7.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	112.983	107.876	99.511
1986	106.024	98.396	103.421	**108.622**	96.487	102.017	97.716	64.231	102.894	114.700	106.240	102.655
1987	98.848	101.148	109.941	103.451	95.953	103.570	97.138	65.434	103.227	110.518	107.883	102.950
1988	100.224	102.064	109.583	99.414	98.419	104.841	92.362	69.242	104.031	106.532	109.002	102.544
1989	104.946	99.537	106.309	101.264	97.453	105.856	92.766	69.702	100.263	111.218	111.239	98.725
1990	105.888	98.235	107.593	100.652	99.751	100.282	95.706	71.281	98.273	113.729	110.199	96.596
1991	107.938	98.809	104.204	103.750	97.781	100.139	98.614	69.739	99.631	113.533	107.003	99.199
1992	105.710	100.065	106.435	103.367	94.846	102.680	96.898	69.139	103.396	112.474	107.221	100.612
1993	100.941	100.580	109.069	101.243	95.517	103.763	96.055	71.026	102.478	108.092	108.664	102.213
1994	102.962	98.110	108.443	100.180	98.218	103.579	93.439	72.800	102.759	107.579	108.403	101.473
1995	104.473	97.530	108.590	.	.	.	.	.	.	.	.	.

TABLE 4.22. B8: Unmodified seasonal-irregular ratios.

4.1.8 Table B8: Estimation of the Unmodified Seasonal-Irregular Component

Description and method of calculation

This table is similar to Table B3: the trend-cycle component is removed from the analysed series, by subtracting or dividing depending on the decomposition model that is selected, so as to provide an estimate of the seasonal-irregular component. We thus have: $B8 = B1 \; op \; B7$.

Comment

Unlike Table B3, the seasonal-irregular component is available for all the data points. This is because the trend-cycle component at the beginning and end of the series has been estimated using asymmetric moving averages.

Example

The value for April 1986 is therefore obtained simply:

$$APR86 = 100 \times 109.500/100.809 = 108.622.$$

4.1.9 Table B9: Replacement Values for Extreme Values of the Seasonal-Irregular Component

Description and method of calculation

For the second time in Part B, extreme values in the seasonal-irregular component will be detected and adjusted automatically by the program. The strategy used is similar to that followed for Table B4 (see Section 4.1.4). Detection is based on the data in Table B8, and unlike what occurs for Table B4, a 3×5 moving average of coefficients $\{1, 2, 3, 3, 3, 2, 1\}/15$ is used for the first estimation of the seasonal component.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	StD
1985	.	.	.	.	.	.	.	.	.	.	.	.	2.077
1986	**104.457**	.	107.611	101.329	.	.	.	68.245	.	.	.	.	2.077
1987	103.337	.	.	.	.	.	.	.	.	.	.	.	2.077
1988	.	.	.	.	.	.	.	.	.	**111.877**	.	.	2.104
1989	.	.	.	.	.	.	.	.	.	.	.	.	1.885
1990	.	.	.	.	.	.	.	.	101.123	.	.	99.679	1.808
1991	105.353	.	106.753	.	.	.	95.836	.	.	.	.	.	1.609
1992	.	.	.	.	.	.	.	.	.	.	.	.	1.625
1993	104.314	.	.	.	.	.	.	.	.	.	.	.	1.603
1994	.	.	.	.	.	.	95.015	70.697	.	.	.	.	1.603
1995	.	.	.	.	.	.	.	.	.	.	.	.	1.603

TABLE 4.23. B9: Replacement values for extreme seasonal-irregular ratios.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.803	107.583	101.831
1986	102.186	100.386	107.447	103.444	97.028	103.833	95.293	66.812	102.916	111.496	107.899	101.660
1987	102.573	100.199	107.399	**103.147**	97.231	103.700	95.192	67.258	102.471	111.238	108.370	101.295
1988	103.089	100.040	107.403	102.525	97.520	103.441	95.150	68.070	101.809	111.212	108.809	100.658
1989	103.899	99.859	107.240	**102.025**	97.711	103.004	95.241	68.959	101.314	111.314	108.970	100.077
1990	104.532	99.715	107.024	101.817	97.530	102.747	95.532	69.692	101.004	111.471	108.955	99.669
1991	104.905	99.449	106.899	101.861	97.314	102.464	95.807	70 265	101.078	111.463	108.699	99.674
1992	104.726	99.208	107.073	102.010	96.961	102.391	96.134	70.550	101.394	111.103	108.329	100.121
1993	104.205	99.083	107.544	101.929	96.722	102.550	96.058	70.779	101.975	110.432	108.036	100.701
1994	103.670	98.987	107.985	101.920	96.432	102.861	95.937	70.801	102.391	110.004	107.932	101.098
1995	103.230	98.939	108.361	.	.	.	.	.	.	.	.	.

TABLE 4.24. B9a: Preliminary seasonal factors (3×5 ma).

Comments

- The 7-term 3×5 symmetric moving average cannot be used to estimate the seasonal factors for the first three and last three years; ad hoc asymmetric moving averages are used (see Table 3.12).

- The use of these asymmetric moving averages could be problematic in the absence of a sufficient number of observations. Assuming that, for a given month, we had only five years of observations, the central point could not be estimated since we would have neither three points in the future, nor three points in the past, as required for the use of an asymmetric moving average. In such a case, it would be estimated using a simple average of the five available observations.

- Other comments at this stage are the same as those made for Table B4 (see Section 4.1.4).

Example

The X-11 output table is shown in Table B9. To better understand it, we will describe in detail the steps used to calculate it on the basis of tables that are unfortunately not printed in the current versions of the X-11 software family (Tables numbered here B9a to B9f).

Step 1: Estimating the seasonal component.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.034	100.034	100.034
1986	100.034	100.034	100.034	**100.034**	100.034	100.041	100.050	100.058	100.048	100.034	100.030	100.033
1987	100.023	100.037	100.037	100.008	100.017	100.021	100.028	100.042	100.036	100.010	99.996	99.998
1988	99.985	100.017	100.023	99.995	100.012	100.004	100.011	100.037	100.023	99.995	99.982	99.972
1989	99.958	99.998	100.015	99.999	100.010	99.992	99.994	100.015	100.000	99.982	99.966	99.948
1990	99.949	99.992	100.009	100.003	100.009	99.991	99.990	99.994	99.978	99.974	99.967	99.946
1991	99.946	99.981	100.008	100.011	100.000	99.990	99.982	99.965	99.962	99.975	99.967	99.949
1992	99.960	99.985	100.010	100.009	99.978	99.981	99.978	99.951	99.966	99.982	99.969	99.965
1993	99.969	99.975	100.009	100.005	99.965	99.977	99.979	99.953	99.967	99.985	99.973	99.973
1994	99.981	99.977	99.995	99.995	99.973	99.985	99.983	99.963	99.976	99.976	99.976	99.976
1995	99.976	99.976	99.976	.	.	.	.	.	.	.	.	.

TABLE 4.25. B9b: Centered 12-month moving average.

The data in Table B8 are smoothed column by column (month by month), using a 3×5 moving average whose coefficients, and those of the associated asymmetric averages, are shown in Table 3.12. Thus, the values of the seasonal-irregular component for the months of April 1986 to 1994 are the following:

108.622, 103.451, 99.414, 101.264, 100.652, 103.750, 103.367, 101.243, 100.180.

The seasonal factor for the month of April 1989 is therefore estimated as follows:

$$
\begin{aligned}
APR89 &= \frac{108.622 + 103.451 \times 2 + 99.414 \times 3 + 101.264 \times 3}{15} + \\
&\quad \frac{100.652 \times 3 + 103.750 \times 2 + 103.367}{15} \\
&= 102.025.
\end{aligned}
$$

This symmetric moving average can be used to estimate the seasonal factors for the years 1989 to 1991. For the beginning of the series (years 1986 to 1988) and the end of the series (years 1992 to 1994), predefined asymmetric averages are used, e.g.:

$$
\begin{aligned}
APR87 &= \frac{108.622 \times 15 + 103.451 \times 15 + 99.414 \times 15}{60} + \\
&\quad \frac{101.264 \times 11 + 100.652 \times 4}{60} \\
&= 103.147
\end{aligned}
$$

(one point in the past, the current point and three points in the future).

Step 2: Normalizing the seasonal factors.

From Table B9a, a centered 12-month moving average is used to produce Table B9b. The first computable term is therefore that for April 1986, and the last that for September 1994. Thus:

$$
APR86 = \frac{111.803}{24} +
$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.765	107.547	101.796
1986	102.151	100.352	107.411	**103.409**	96.994	103.791	95.246	66.774	102.866	111.458	107.867	101.627
1987	102.550	100.162	107.359	103.139	97.215	103.678	95.165	67.229	102.434	111.226	108.374	101.298
1988	103.104	100.023	107.378	102.530	97.509	103.437	95.140	68.045	101.786	111.217	108.828	100.686
1989	103.943	99.860	107.224	102.027	97.702	103.012	95.247	68.949	101.314	111.334	109.008	100.130
1990	104.585	99.724	107.014	101.814	97.522	102.756	95.542	69.697	101.027	111.499	108.990	99.723
1991	104.961	99.468	106.890	101.850	97.314	102.475	95.824	70.290	101.117	111.490	108.735	99.725
1992	104.768	99.223	107.062	102.001	96.982	102.410	96.155	70.584	101.429	111.123	108.363	100.155
1993	104.237	99.108	107.534	101.923	96.756	102.574	96.078	70.812	102.009	110.449	108.066	100.728
1994	103.689	99.010	107.990	101.925	96.458	102.876	95.953	70.827	102.415	110.030	107.958	101.122
1995	103.254	98.962	108.386	.	.	.	.	.	.	.	.	.

TABLE 4.26. B9c: Normalized seasonal factors.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	101.090	100.306	97.755
1986	103.792	98.051	96.285	**105.041**	99.477	98.291	102.593	96.192	100.027	102.909	98.491	101.011
1987	96.390	100.984	102.405	100.303	98.702	99.896	102.073	97.330	100.774	99.363	99.547	101.631
1988	97.206	102.040	102.054	96.961	100.933	101.358	97.081	101.760	102.206	95.787	100.160	101.845
1989	100.965	99.676	99.147	99.252	99.745	102.761	97.395	101.092	98.962	99.896	102.047	98.597
1990	101.246	98.507	100.541	98.859	102.286	97.592	100.173	102.273	97.275	102.000	101.109	96.865
1991	102.836	99.338	97.488	101.865	100.480	97.721	102.912	99.216	98.531	101.832	98.408	99.472
1992	100.899	100.849	99.414	101.340	97.798	100.264	100.773	97.952	101.940	101.215	98.946	100.456
1993	96.838	101.485	101.427	99.333	98.720	101.160	99.976	100.302	100.459	97.866	100.554	101.475
1994	99.299	99.091	100.419	98.288	101.824	100.683	97.380	102.785	100.336	97.772	100.413	100.347
1995	101.180	98.553	100.188	.	.	.	.	.	.	.	.	.

TABLE 4.27. B9d: Preliminary irregular component.

$$
\begin{aligned}
& \frac{107.583 + 101.831 + 102.186 + 100.386 + 107.447}{12} + \\
& \frac{103.444 + 97.028 + 103.833 + 95.293 + 66.812 + 102.916}{12} + \\
& \frac{111.496}{24} \\
= \; & 100.034.
\end{aligned}
$$

The first six values, from October 1985 to March 1986, which cannot be computed using this symmetric moving average, are considered equal to the first computable value, that for April 1986. The same procedure is used for the end of the series; the value calculated for September 1994 (99.976) is repeated for the next six months.

Normalized seasonal factors are then obtained by dividing Table B9a by Table B9b to obtain Table B9c. Thus,

$$APR86 = 100 \times 103.444/100.034 = 103.409.$$

Step 3: Estimating the irregular component.

We simply divide the seasonal-irregular component of Table B8 by the normalized seasonal factors of Table B9c to obtain Table B9d. Thus,

$$APR86 = 100 \times 108.622/103.409 = 105.041.$$

Step 4: Calculating a moving standard deviation.

Year	Standard Deviation 1	Standard Deviation 2
1985	2.0774	2.0774
1986	2.0774	2.0774
1987	2.0774	2.0774
1988	2.1038	2.1038
1989	1.8846	1.8846
1990	1.8082	1.8082
1991	1.6093	1.6093
1992	1.6246	1.6246
1993	1.6030	1.6030
1994	1.6030	1.6030
1995	1.6030	1.6030

TABLE 4.28. B9e: Five-year moving standard deviations.

The standard deviation corresponding to the year 1989 is calculated on the basis of data for the years 1987 to 1991 (two years before, two years after) using the formula[12]:

$$\sigma_{89} = \left[\frac{1}{60} \sum_{t=Jan87}^{Dec91} (I_t - 100)^2 \right]^{1/2} = 1.8846.$$

Those for the years 1988, 1990, 1991 and 1992 are calculated using the same principle.

For X-11-ARIMA and X-12-ARIMA, the standard deviation for 1987 is calculated on the basis of all available observations from 1985 to 1990, i.e. 63 observations. These first estimates of the standard deviations are shown in Table B9e, in the Standard Deviation 1 column.

They are used to detect the extreme values. For a given year, a value is considered very extreme if, in absolute value, it deviates by more than 2.5 times the standard deviation corresponding to that year from its theoretical average . In our example, according to Figure 4.5, no value is considered very extreme. The recalculation of the standard deviations therefore leads to the same results (the Standard Deviation 2 column in Table B9e).

Step 5: Detecting extreme values and weighting the irregular.

Values of the irregular are then classified in relation to the upper and lower detection thresholds calculated on the basis of the previously estimated standard deviations. All values beyond the lower detection thresholds are considered extreme, and are therefore adjusted; the weights (multiplied by 100) associated with each of these values are shown in Table B9f. Thus, for October 1988, we have: $|OCT88 - 100| = |95.787 - 100| = 4.213$, and $1.5 \times \sigma_{88} = 1.5 \times 2.1038 = 3.1557 < 4.213 < 2.5 \times \sigma_{88} = 2.5 \times 2.1038 = 5.2595$. This value, considered moderately extreme, is assigned a weight

[12]Here the theoretical mean is assumed to be 100 as the irregular values have been multiplied by 100.

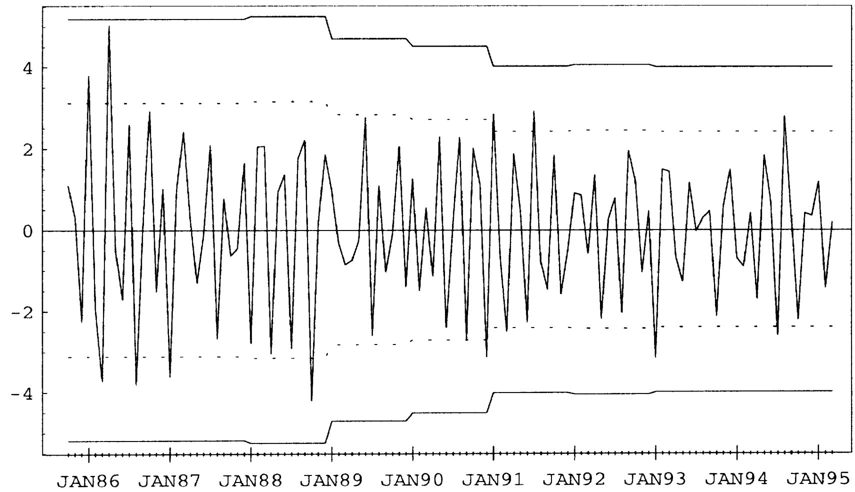


FIGURE 4.5. B9d: Deviation of the irregular component from its theoretical average with the detection thresholds corresponding to $\pm 1.5\sigma$ and $\pm 2.5\sigma$.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	67.475	100.000	71.178	7.340	100.000	100.000	100.000	66.711	100.000	100.000	100.000	100.000
1987	76.235	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	**49.731**	100.000	100.000
1989	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	99.273	100.000	100.000	76.630
1991	73.758	100.000	93.885	100.000	100.000	100.000	69.057	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1993	52.737	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	100.000	100.000	100.000	100.000	86.573	76.236	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

TABLE 4.29. B9f: Weights associated with values of the irregular.

proportional to the deviation from its average, which is:

$$weight(OCT88) = \frac{5.2595 - 4.213}{5.2595 - 3.1557} = 0.497.$$

Step 6: Adjusting extreme values of the seasonal-irregular component.

Finally, the seasonal-irregular component (Table B8) is adjusted on the basis of these weights. Thus, the value for October 1988 is replaced by the average of this value with its assigned weight and the two preceding and subsequent values for the same month receiving full weight, i.e. not considered extreme. As can be seen from Table B9f, these are the values for the months of October 1986, 1987, 1990 and 1991. This leads to:

$$\begin{aligned} SI(OCT88) &= \frac{114.700 + 110.518 + 106.532 \times 0.497 + 111.218 + 113.729}{4 + 0.497} \\ &= 111.877. \end{aligned}$$

The value for January 1986 is also considered extreme. However, as it appears at the beginning of the series, it is adjusted differently; it is replaced by the average of this value with its assigned weight and the four nearest values for the same month receiving full weight. In this case then, according to Table B9f, these are the values for January 1988, 1989, 1990 and 1992. And so we have:

$$\begin{aligned} SI(JAN86) &= \frac{106.024 \times 0.675 + 100.224 + 104.946 + 105.888 + 105.710}{4 + 0.675} \\ &= 104.457. \end{aligned}$$

4.1.10 Table B10: Estimation of the Seasonal Component

Description and method of calculation

The estimation is done, using a procedure similar to that which led to Table B5 (see Section 4.1.5), on the basis of values for the seasonal-irregular component in Table B8, corrected with the values in Table B9. The procedure follows two steps: estimating the seasonal component, then normalizing the seasonal factors.

As in Table B9, and unlike in Table B5, a 3×5 moving average is used to estimate the seasonal component.

Comment

We can specify the moving average to be used. In such a case, X-11-ARIMA gives the choice between a simple 3-term, a 3×3, a 3×5, a 3×9 and constant seasonality (a simple average). X-12-ARIMA also includes a 3×15.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	112.983	107.876	99.511
1986	**104.457**	98.396	107.611	101.329	96.487	102.017	97.716	68.245	102.894	114.700	106.240	102.655
1987	103.337	101.148	109.941	103.451	95.953	103.570	97.138	65.434	103.227	110.518	107.883	102.950
1988	100.224	102.064	109.583	99.414	98.419	104.841	92.362	69.242	104.031	**111.877**	109.002	102.544
1989	104.946	99.537	106.309	101.264	97.453	105.856	92.766	69.702	100.263	111.218	111.239	98.725
1990	105.888	98.235	107.593	100.652	99.751	100.282	95.706	71.281	101.123	113.729	110.199	99.679
1991	105.353	98.809	106.753	103.750	97.781	100.139	95.836	69.739	99.631	113.533	107.003	99.199
1992	105.710	100.065	106.435	103.367	94.846	102.680	96.898	69.139	103.396	112.474	107.221	100.612
1993	104.314	100.580	109.069	101.243	95.517	103.763	96.055	71.026	102.478	108.092	108.664	102.213
1994	102.962	98.110	108.443	100.180	98.218	103.579	95.015	70.697	102.759	107.579	108.403	101.473
1995	104.473	97.530	108.590	.	.	.	.	.	.	.	.	.

TABLE 4.30. B9g: Corrected seasonal-irregular component.

Example

The estimation is based on the corrected seasonal-irregular component from Table B9g.

Step 1: Estimating the seasonal component.
The data in Table B9g are smoothed column by column (month by month), using a 3×5 moving average (see Table 3.12), leading to Table B10a. The seasonal factor for the month of April 1989 is therefore estimated as follows:

$$\begin{aligned} APR89 &= \frac{101.329 + 103.451 \times 2 + 99.414 \times 3 + 101.264 \times 3}{15} + \\ &\quad \frac{100.652 \times 3 + 103.750 \times 2 + 103.367}{15} \\ &= 101.539. \end{aligned}$$

For the beginning of the series (years 1986 to 1988) and the end of the series (years 1992 to 1994), predefined asymmetric averages are used, e.g.:

$$\begin{aligned} APR87 &= \frac{101.329 \times 15 + 103.451 \times 15 + 99.414 \times 15}{60} + \\ &= \frac{101.264 \times 11 + 100.652 \times 4}{60} \\ &= 101.324 \end{aligned}$$

(one point in the past, the current point and three points in the future).

Step 2: Normalizing the seasonal factors.
On Table B10a, a centered 12-month moving average is used to produce Table B10b. The first computable term is therefore that for April 1986, and the last that for September 1994. Thus:

$$\begin{aligned} APR86 &= \frac{112.605}{24} + \\ &\quad \frac{107.583 + 101.831 + 103.013 + 100.386 + 108.635}{12} + \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	112.605	107.583	101.831
1986	103.013	100.386	108.635	101.378	97.028	103.833	95.293	67.950	102.916	112.476	107.899	101.660
1987	103.304	100.199	108.447	**101.324**	97.231	103.700	95.192	68.261	102.661	112.396	108.370	101.501
1988	103.654	100.040	108.201	101.431	97.520	103.441	94.965	68.672	102.189	112.281	108.809	101.069
1989	104.048	99.859	107.859	**101.539**	97.711	103.004	94.871	69.227	101.884	112.383	108.970	100.694
1990	104.539	99.715	107.534	101.817	97.530	102.747	94.976	69.692	101.574	112.184	108.955	100.286
1991	104.837	99.449	107.408	101.861	97.314	102.464	95.356	70.125	101.648	111.819	108.699	100.291
1992	104.883	99.208	107.583	102.010	96.961	102.391	95.769	70.235	101.774	111.103	108.329	100.532
1993	104.591	99.083	107.883	101.929	96.722	102.550	95.943	70.253	102.165	110.432	108.036	100.906
1994	104.341	98.987	108.155	101.920	96.432	102.861	95.967	70.205	102.391	110.004	107.932	101.098
1995	104.185	98.939	108.361	.	.	.	.	.	.	.	.	.

TABLE 4.31. B10a: Preliminary seasonal factors (3 × 5 ma).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.199	100.199	100.199
1986	100.199	100.199	100.199	**100.199**	100.207	100.213	100.218	100.222	100.206	100.196	100.202	100.205
1987	100.196	100.204	100.207	100.193	100.209	100.222	100.230	100.238	100.221	100.215	100.232	100.233
1988	100.213	100.221	100.218	100.194	100.207	100.207	100.206	100.215	100.193	100.183	100.196	100.185
1989	100.163	100.182	100.193	100.184	100.195	100.186	100.191	100.206	100.186	100.184	100.188	100.170
1990	100.164	100.187	100.194	100.173	100.164	100.146	100.141	100.143	100.126	100.123	100.116	100.095
1991	100.099	100.133	100.154	100.142	100.116	100.106	100.108	100.100	100.097	100.110	100.102	100.084
1992	100.098	100.120	100.130	100.105	100.060	100.055	100.053	100.035	100.042	100.052	100.038	100.035
1993	100.049	100.057	100.074	100.062	100.022	100.026	100.031	100.016	100.024	100.035	100.022	100.023
1994	100.037	100.036	100.043	100.035	100.013	100.016	100.018	100.009	100.016	100.016	100.016	100.016
1995	100.016	100.016	100.016	.	.	.	.	.	.	.	.	.

TABLE 4.32. B10b: Centered 12-term moving average.

$$
\begin{aligned}
& \frac{101.378 + 97.028 + 103.833 + 95.293 + 67.950 + 102.916}{12} + \\
& \frac{112.476}{24} \\
= \; & 100.199.
\end{aligned}
$$

The first six values, from October 1985 to March 1986, which cannot be computed using this symmetric moving average, are taken to be equal to the first computable value, that for April 1986. The same procedure is used for the end of the series: the value calculated for September 1994 (100.016) is repeated for the next six months. The normalized seasonal factors, in Table B10, are then obtained by dividing Table B10a by Table B10b. For example,

$$APR86 = 100 \times 101.378/100.199 = 101.177.$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	112.382	107.370	101.629
1986	102.809	100.187	108.419	**101.177**	96.828	103.613	95.086	67.799	102.704	112.255	107.681	101.452
1987	103.102	99.995	108.223	101.129	97.028	103.470	94.973	68.099	102.435	112.154	108.120	101.265
1988	103.433	99.820	107.966	101.235	97.319	103.227	94.770	68.525	101.992	112.076	108.596	100.882
1989	103.879	99.677	107.652	101.352	97.521	102.812	94.690	69.085	101.694	112.177	108.766	100.523
1990	104.368	99.529	107.326	101.641	97.371	102.597	94.842	69.593	101.446	112.046	108.828	100.190
1991	104.734	99.317	107.243	101.717	97.201	102.356	95.253	70.055	101.550	111.696	108.588	100.207
1992	104.780	99.089	107.443	101.902	96.902	102.335	95.718	70.210	101.731	111.046	108.287	100.497
1993	104.540	99.027	107.804	101.865	96.700	102.524	95.913	70.242	102.141	110.394	108.012	100.883
1994	104.302	98.952	108.108	101.884	96.419	102.844	95.950	70.199	102.374	109.987	107.915	101.082
1995	104.169	98.923	108.343	.	.	.	.	.	.	.	.	.

TABLE 4.33. B10: Seasonal factors.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	102.953	102.263	98.988
1986	103.687	98.516	95.832	**108.226**	100.901	100.084	104.853	96.904	102.430	104.316	100.575	102.906
1987	97.476	103.205	104.322	105.905	103.063	104.668	107.188	100.882	106.117	104.231	106.086	108.626
1988	104.125	110.399	109.942	106.781	110.359	111.115	106.785	110.908	112.362	105.197	111.698	113.697
1989	113.498	112.563	111.656	113.169	113.309	117.010	111.522	114.931	112.297	112.947	116.581	112.114
1990	116.032	113.033	115.163	114.225	118.722	113.843	117.881	119.696	112.967	117.809	116.789	110.589
1991	117.727	113.576	111.243	117.385	116.563	114.014	121.046	116.480	114.624	118.536	114.930	115.561
1992	117.866	117.975	115.410	117.760	113.310	115.991	117.114	113.944	117.270	116.168	112.755	113.238
1993	108.762	114.211	113.818	112.109	111.582	114.217	112.706	113.466	112.393	109.607	112.672	113.795
1994	111.503	112.681	114.700	113.266	118.234	117.654	114.123	121.655	117.803	114.923	118.334	118.716
1995	119.134	117.566	120.173	.	.	.	.	.	.	.	.	.

TABLE 4.34. B11: Seasonally adjusted series.

4.1.11 Table B11: Estimation of the Seasonally Adjusted Series

Description and method of calculation

The estimation is done simply by removing from the starting series, in Table B1, the estimate of the seasonal component in Table B10: $B11 = B1 \; op \; B10$.

Example

The value for April 86 is:

$$APR86 = 100 \times 109.500/101.177 = 108.226.$$

4.1.12 Table B13: Estimation of the Irregular Component

Description and method of calculation

The estimation is done simply by removing from the seasonally adjusted series, in Table B11, the estimate of the trend-cycle component in Table B7: $B13 = B11 \; op \; B7$.

Example

The value for April 86 is:

$$APR86 = 100 \times 108.226/100.809 = 107.358.$$

4.1.13 Estimating the Effect of the Daily Composition of the Month (Trading-Day)

Some economic series, e.g. the monthly sales of a retail business, are strongly affected by the daily composition of the month: one Saturday more or less

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.535	100.471	97.916
1986	103.127	98.213	95.390	**107.358**	99.648	98.460	102.766	94.737	100.185	102.178	98.661	101.186
1987	95.874	101.153	101.588	102.296	98.892	100.097	102.280	96.086	100.774	98.541	99.781	101.664
1988	96.897	102.248	101.498	98.202	101.130	101.564	97.460	101.047	101.999	95.053	100.374	101.647
1989	101.028	99.859	98.753	99.913	99.930	102.961	97.968	100.893	98.592	99.146	102.274	98.212
1990	101.457	98.700	100.249	99.026	102.444	97.743	100.912	102.425	96.873	101.503	101.259	96.413
1991	103.060	99.489	97.167	101.999	100.597	97.834	103.528	99.549	98.111	101.645	98.541	98.994
1992	100.887	100.985	99.061	101.438	97.878	100.337	101.233	98.474	101.637	101.286	99.015	100.115
1993	96.558	101.568	101.173	99.389	98.776	101.209	100.148	101.117	100.329	97.915	100.604	101.319
1994	98.715	99.149	100.309	98.327	101.865	100.715	97.384	103.706	100.376	97.811	100.452	100.387
1995	100.292	98.592	100.227	.	.	.	.	.	.	.	.	.

TABLE 4.35. B13: Irregular series.

in a month can change the monthly sales significantly. These **trading-day effects**, like seasonality, can complicate comparisons of values in a series from one month to the next, or those for a given month from one year to the next, and that is why, when these effects are deemed statistically significant, they are generally removed from the series during the seasonal adjustment process.

The irregular component, e.g. in Table B13, does not by design contain any trend or seasonality; however, if there are any trading-day effects, this is where they can be found[13]. It is therefore quite natural to use the estimate of the irregular component to identify any trading-day effects, using for instance a linear regression model.

At the user's request, these effects can be estimated and corrected automatically. That is the purpose of Tables B14, B15, B16, B18 and B19 in Part B and of their counterparts in Part C.

A few properties of our calendar

Our calendar is based on the solar cycle. The earth takes roughly 365 days and six hours to complete one revolution around the Sun. To account for these six extra hours, the calendar uses 365 days three years out of four, and the fourth year (called **leap year**) has one extra day, February 29. The rule assumes that leap years correspond to any year which is numerically divisible by four. However, this correction is too strong. That is why centennial years are not leap years, unless they are divisible by 400 (1600, 2000, 2400, etc.). There remains, nevertheless, a slight error considered to be one day every 4000 years. In terms of months, we therefore have a periodical calendar of period 4, at least between 1901 and 2099.

For a given date, the corresponding day is shifted: if the first day of January in a non-leap year is a Saturday, it will be a Sunday the following year, or a Monday if the reference year is a leap year. We must therefore

[13]The reason is simply that trading-day effects have a characteristic spectral signature which is filtered out by Henderson's moving averages (hence they cannot be in the trend-cycle estimate), and by the moving averages applied month-by-month for extracting the seasonal factors (hence they cannot be in the seasonal component).

wait 28 years (4×7) to find the same annual structure in terms of dates and days.

Trading-day effects

It will be assumed below, following the notation of Findley et al. [23], that the j-th day of the week has an effect α_j where, for example, $j = 1$ refers to Monday, $j = 2$ refers to Tuesday, ..., and $j = 7$ refers to Sunday. Each α_j represents for example the average sales for one day j. If D_{jt} represents the number of days j in the month t, the length of the month will be $N_t = \sum_{j=1}^{7} D_{jt}$ and the cumulative effect for that month, the total sales of the month, will be $\sum_{j=1}^{7} \alpha_j D_{jt}$. We also have $\bar{\alpha} = \sum_{j=1}^{7} \alpha_j/7$, the mean daily effect, the average sales for one day. Since by design we have $\sum_{j=1}^{7} (\alpha_j - \bar{\alpha}) = 0$, we may write:

$$\begin{aligned} \sum_{j=1}^{7} \alpha_j D_{jt} &= \bar{\alpha} N_t + \sum_{j=1}^{7} (\alpha_j - \bar{\alpha}) D_{jt} \\ &= \bar{\alpha} N_t + \sum_{j=1}^{6} (\alpha_j - \bar{\alpha}) (D_{jt} - D_{7t}) . \end{aligned} \tag{4.1}$$

Thus, the cumulative monthly effect is decomposed into an effect directly linked to the length of the month and a net effect for each day of the week.

Note that the sum $\sum_{j=1}^{7} (\alpha_j - \bar{\alpha}) D_{jt}$ involves only the days of the week occurring five times in a month; every month contains four complete weeks, for which by definition the effect linked to the days is cancelled out, plus 0, 1, 2 or 3 days which contribute to the trading-day effect for the month.

Regression model

If equation (4.1) is to be consistent with the irregular component of Table B13 to be analyzed, which contains neither seasonality nor trend, it must be adjusted for these effects.

- Potentially, part $\bar{\alpha} N_t$ of the equation contains such components because the months vary in length and because, as we have seen, variable N_t is periodic (period of 48 months or 4 years). These effects can be summarized by the quantity $\bar{\alpha} N_t^*$ where N_t^* represents the average, over four years, of the length of the month t. In other words, N_t^* is equal to 30 or 31 if the month in question is not a February, and is equal to 28.25 otherwise. Thus, we have: $\bar{\alpha} N_t = \bar{\alpha} N_t^* + \bar{\alpha} (N_t - N_t^*)$, an equation whose second part is zero except for Februarys.

- The second part of the equation includes D_{jt}, the number of times that day j is present in month t. These variables are periodic (period of 336 months or 28 years) with equal means for a given month[14]. In the second part of the equation, the difference $D_{jt} - D_{7t}$ is used, and since these variables show the same behaviour, the difference involves no seasonality and no trend.

The procedure used to adjust equation (4.1) for these effects depends on the decomposition model used.

- For a **multiplicative decomposition**, the seasonal and trend effects are removed by dividing equation (4.1) by $\bar{\alpha}N_t^*$. By writing $\beta_j = \alpha_j/\bar{\alpha} - 1$, we obtain:

$$\frac{1}{N_t^*}\sum_{j=1}^{7}(\beta_j + 1)\,D_{jt} = \frac{N_t}{N_t^*} + \sum_{j=1}^{6}\beta_j\left(\frac{D_{jt} - D_{7t}}{N_t^*}\right) \tag{4.2}$$

  Thus, if I_t is an estimate of the irregular component, Census X-11 and its direct variants estimate coefficients $\beta = (\beta_1, \ldots, \beta_6)$ (and $\beta_7 = -\sum_{j=1}^{6}\beta_j$) by adjusting the following model using ordinary least squares:

$$N_t^* I_t - N_t = \sum_{j=1}^{6}\beta_j\,(D_{jt} - D_{7t}) + e_t. \tag{4.3}$$

  This corresponds to the model proposed by Young [67].

- For an **additive decomposition**, logically $\bar{\alpha}N_t^*$ must be subtracted from equation (4.1). We thus have:

$$I_t = \beta_0\,(N_t - N_t^*) + \sum_{j=1}^{6}\beta_j\,(D_{jt} - D_{7t}) + e_t \tag{4.4}$$

  where now $\beta_0 = \bar{\alpha}$ and $\beta_j = \alpha_j - \bar{\alpha}$ for $j = 1, \ldots, 6$.

  In Census X-11 and X-11-ARIMA, the first regressor $N_t - N_t^*$ is omitted, and only six parameters are estimated.

Estimating the parameters

In the case of Census X-11 and X-11-ARIMA, the model can therefore be written as follows, whether the model is additive or multiplicative:

$$\begin{aligned} Y_t &= \sum_{j=1}^{6}\beta_j\,(D_{jt} - D_{7t}) + e_t \\ &= \sum_{j=1}^{6}\beta_j Z_{jt} + e_t, \end{aligned}$$

[14] For a month with 31 days, this common average is 4.428574, for a month with 30 days it is 4.285714, whereas for February, it is 4.035714.

where $Y_t = N_t^* I_t - N_t$ in the multiplicative model, or $Y_t = I_t$ in the additive model.

The solution by ordinary least squares[15] leads to the following results, given $\hat{\sigma}^2 = \hat{e}'\hat{e}/(n-6)$ where the $\hat{e}$ are the residuals of the regression:

Parameter	Variance
$\hat{\beta} = (Z'Z)^{-1} Z'Y$	$\hat{s}^2\left(\hat{\beta}_j\right) = \hat{\sigma}^2 (Z'Z)^{-1}_{jj}$
$\hat{\beta}_7 = -\sum_{j=1}^{6} \hat{\beta}_j$	$\hat{s}^2\left(\hat{\beta}_7\right) = \hat{\sigma}^2 \sum_{i=1}^{6} \sum_{j=1}^{6} (Z'Z)^{-1}_{ij}$

The T-tests (Student's tests) aimed at verifying whether a coefficient is zero, and the F-tests (Fisher's tests) for the presence of an overall effect attributed to trading-day are also available:

Hypothesis	Statistic	Test used
$\beta_j = k$	$t_j = \frac{\hat{\beta}_j - k}{\hat{s}^2(\hat{\beta}_j)}$	t-test, Student's distribution with $n-6$ degrees of freedom.
$\beta_j = 0, j = 1, \ldots, 6$	$F = \frac{\hat{\beta}'Z'Z\hat{\beta}}{6\hat{\sigma}^2}$	F-test, F distribution with 6 and $n-6$ degrees of freedom.

4.1.14 Table B14: Values of the Irregular Component Excluded from the Trading-Day Regression

Description and method of calculation

At the user's request, the program can estimate, within the series, the effect of the daily composition of the month. This calendar effect is identified in the irregular component and evaluated by linear regression. First, X-11 identifies very extreme values of the irregular component and excludes them from the calculations, thus making the regression results more robust. The process involves three steps:

Step 1: Calculating the mean of the irregular by type of month. We have 15 different types of months:

- months of 31 days beginning on a Monday, Tuesday, Wednesday, etc., i.e. seven categories;
- months of 30 days beginning on a Monday, Tuesday, Wednesday, etc., i.e. seven more categories;
- months of February with 28 days (the 29 days February are excluded at this stage).

[15] The use of OLS and associated tests is based on the additional hypothesis that the irregular component to be explained involves little or no autocorrelation.

Values for the irregular component are thus grouped into 15 categories for which the means m_i are calculated. We thus have n^* data points distributed into 15 groups of size $n_i, (i = 1, \ldots, 15)$ with $n^* = \sum_{i=1}^{15} n_i$ and $m_i = \sum_{j=1}^{n_i} I_{ij}/n_i$.

Step 2: First calculation of an overall standard deviation and identification of extreme values.

The squares of the deviations from the average in each category are calculated, and their average provides an estimate of the overall standard deviation: $\sigma^2 = \sum_{i=1}^{15} \sum_{j=1}^{n_i} \left(I_{ij} - m_i\right)^2 / n^*$.

A value I_{ij} of the irregular is considered very extreme if it is too far away from the mean m_i of the category to which it belongs, i.e. if $|I_{ij} - m_i| \geq \lambda \times \sigma$ where σ is the overall standard deviation calculated above and λ is a parameter that can be modified by the user and is set by default at 2.5.

Step 3: Final calculation of the overall standard deviation and identification of extreme values.

Both steps are repeated, excluding the extreme values identified in the first iteration: calculating the means by category and estimating the standard deviation. Extreme values identified in this way and excluded from the regression for trading-day are shown in Table B14.

Comments

- n^*, the number of observations for computing the standard deviation, is smaller than n, the total number of observations in the series, because leap years are not used.

- Theoretically speaking, we are adopting here a single-factor analysis of variance model in an attempt to evaluate the trading-day effect for each of the 15 types of months. It is therefore assumed that, in each category i, the irregular in Table B13 follows a Normal distribution with mean m_i and standard deviation σ. In this case, the standard deviation estimate used in Step 2 is biased for the unknown constant σ. It would be necessary to divide by $n^* - 15$ to obtain an unbiased estimate.

- For leap years' Februarys, the mean of the category is not calculated but rather considered equal to *xbar* (0 for an additive model, 1 for a multiplicative model).

- In Step 3, values considered as extreme in Step 2 are compared to the theoretical mean *xbar*.

- All observations, with the exception of extreme values, are used for the regression.

Length	1^{st} of the month	Months involved	Mean
28		FEB86 FEB87 FEB89 FEB90 FEB91 FEB93 FEB94 FEB95	**99.590**
30	Sunday	JUN86 NOV87 APR90 SEP91 NOV92	98.879
	Monday	SEP86 JUN87 APR91 JUN92 NOV93	100.644
	Tuesday	APR86 SEP87 NOV88 SEP92 JUN93 NOV94	101.967
	Wednesday	APR87 JUN88 NOV89 APR92 SEP93 JUN94	101.436
	Thursday	SEP88 JUN89 NOV90 APR93 SEP94	101.197
	Friday	NOV85 APR88 SEP89 JUN90 NOV91 APR94	98.646
	Saturday	NOV86 APR89 SEP90 JUN91	98.320
31	Sunday	DEC85 MAR87 MAY88 JAN89 OCT89 JUL90 DEC91 MAR92 AUG93 MAY94 JAN95	100.277
	Monday	DEC86 AUG88 MAY89 JAN90 OCT90 JUL91 MAR93 AUG94	101.691
	Tuesday	OCT85 JUL86 DEC87 MAR88 AUG89 MAY90 JAN91 OCT91 DEC92 MAR94	101.493
	Wednesday	JAN86 OCT86 JUL87 MAR89 AUG90 MAY91 JAN92 JUL92 DEC93 MAR95	101.303
	Thursday	MAY86 JAN87 OCT87 DEC88 MAR90 AUG91 OCT92 JUL93 DEC94	99.703
	Friday	AUG86 MAY87 JAN88 JUL88 DEC89 MAR91 MAY92 JAN93 OCT93 JUL94	97.310
	Saturday	MAR86 AUG87 OCT88 JUL89 DEC90 AUG92 MAY93 JAN94 OCT94	97.188

TABLE 4.36. B14a: Distribution of months by type.

Example

Step 1: Calculating the mean of the irregular by type of month.

Months are first arranged within 15 categories according to the number of days and the day corresponding to the first of the month, see Table B14a.

For each group, we calculate the mean of the values for the irregular in Table B13 (means shown in the last column of Table B14a). Thus, for months of February with 28 days, we have:

$$m_{Feb} = \frac{98.213 + 101.153 + 99.859 + 98.700 + 99.489}{8} + \frac{101.568 + 99.149 + 98.592}{8} = 99.590.$$

Step 2: First calculation of an overall standard deviation and identification of extreme values.

We then calculate the absolute deviation of each value of the irregular from the average for the category to which it belongs (see Table B14b). Thus we have: $JUN86 = |98.460 - 98.879| = 0.419$ and $FEB88 = |102.248 - 100| = 2.248$.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	0.958	1.825	2.361
1986	1.825	1.378	1.798	5.391	0.055	**0.419**	1.273	2.572	0.459	0.875	0.341	0.505
1987	3.829	1.563	1.310	0.860	1.582	0.547	0.977	1.101	**1.194**	1.162	0.903	0.171
1988	0.413	**2.248**	0.005	0.444	0.853	0.128	0.150	0.645	0.802	2.134	1.594	1.944
1989	0.751	0.269	2.550	1.593	1.761	1.764	0.781	0.599	0.054	1.131	0.838	0.902
1990	0.234	0.890	0.546	0.148	0.951	0.903	0.634	1.123	1.448	0.188	0.063	0.774
1991	1.567	0.102	0.143	1.355	0.706	0.486	1.837	0.154	0.768	0.152	0.105	1.283
1992	0.416	0.985	1.216	0.002	0.568	0.307	0.070	1.287	0.330	1.582	0.137	1.378
1993	0.752	1.978	0.518	1.808	1.589	0.759	0.445	0.840	1.107	0.605	0.041	0.016
1994	1.528	0.441	1.184	0.319	1.588	0.721	0.074	2.015	0.821	0.623	1.515	0.684
1995	0.015	0.999	1.075	.	.	.	.	.	.	.	.	.

TABLE 4.37. B14b: Deviations from the averages in absolute value.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	0.917	3.332	5.576
1986	3.330	1.898	3.231	**29.061**	0.003	0.175	1.620	6.617	0.211	0.766	0.116	0.255
1987	**14.661**	2.442	1.717	0.740	2.504	0.300	0.955	1.212	1.425	1.350	0.815	0.029
1988	0.171	5.052	0.000	0.197	0.727	0.016	0.022	0.415	0.644	4.555	2.540	3.779
1989	0.564	0.072	6.501	2.536	3.101	3.112	0.609	0.359	0.003	1.280	0.702	0.813
1990	0.055	0.792	0.298	0.022	0.905	0.815	0.403	1.260	2.095	0.035	0.004	0.600
1991	2.456	0.010	0.021	1.835	0.498	0.236	3.374	0.024	0.590	0.023	0.011	1.646
1992	0 173	0.971	1.478	0.000	0.323	0.094	0.005	1.656	0.109	2.504	0.019	1.899
1993	0.565	3.912	0.268	3.267	2.524	0.575	0.198	0.706	1.225	0.366	0.002	0.000
1994	2.334	0.195	1.401	0.102	2.522	0.520	0.005	4.060	0.675	0.389	2.296	0.468
1995	0.000	0.998	1.156	.	.	.	.	.	.	.	.	.

TABLE 4.38. B14c: Squares of the deviations from the averages.

The mean of the square of these values (see Table B14c), except the 29 days February values, provides the first estimate of the standard deviation:

$$\sigma = \left(\frac{1}{n^*} \sum_{i=1}^{15} \sum_{j=1}^{n_i} (I_{ij} - m_i)^2 \right)^{1/2} = 1.2499$$

where $n^* = 114 - 2 = 112$. This standard deviation will be used to determine the limits in relation to which an irregular will be considered extreme. By default, this limit is equal to 2.5 times the value of the standard deviation, i.e. 3.125.

The only two points that, in absolute value, are farther than 3.125 (i.e. 2.5σ) are the values for April 1986 (a month of 30 days beginning on a Tuesday) and for January 1987 (a month of 31 days beginning on a Thursday). Both these points are therefore excluded, and the calculations for the first two steps are redone.

Step 3: Final calculation of the overall standard deviation and identification of extreme values

Only those means for the types of months involving excluded values are affected (see Table B14d), and the calculation of the squares of the deviations from the averages leads to a new estimate of the standard deviation:

$$\sigma = \left(\frac{1}{n^* - 2} \sum_{i=1}^{15} \sum_{j=1}^{n_i} (I_{ij} - m_i)^2 \right)^{1/2} = 1.0600.$$

Length	1^{st} of the month	Months involved	Mean
28		FEB86 FEB87 FEB89 FEB90 FEB91 FEB93 FEB94 FEB95	99.590
30	Sunday	JUN86 NOV87 APR90 SEP91 NOV92	98.879
	Monday	SEP86 JUN87 APR91 JUN92 NOV93	100.644
	Tuesday	SEP87 NOV88 SEP92 JUN93 NOV94	**100.889**
	Wednesday	APR87 JUN88 NOV89 APR92 SEP93 JUN94	101.436
	Thursday	SEP88 JUN89 NOV90 APR93 SEP94	101.197
	Friday	NOV85 APR88 SEP89 JUN90 NOV91 APR94	98.646
	Saturday	NOV86 APR89 SEP90 JUN91	98.320
31	Sunday	DEC85 MAR87 MAY88 JAN89 OCT89 JUL90 DEC91 MAR92 AUG93 MAY94 JAN95	100.277
	Monday	DEC86 AUG88 MAY89 JAN90 OCT90 JUL91 MAR93 AUG94	101.691
	Tuesday	OCT85 JUL86 DEC87 MAR88 AUG89 MAY90 JAN91 OCT91 DEC92 MAR94	101.493
	Wednesday	JAN86 OCT86 JUL87 MAR89 AUG90 MAY91 JAN92 JUL92 DEC93 MAR95	101.303
	Thursday	MAY86 OCT87 DEC88 MAR90 AUG91 OCT92 JUL93 DEC94	**100.182**
	Friday	AUG86 MAY87 JAN88 JUL88 DEC89 MAR91 MAY92 JAN93 OCT93 JUL94	97.310
	Saturday	MAR86 AUG87 OCT88 JUL89 DEC90 AUG92 MAY93 JAN94 OCT94	97.188

TABLE 4.39. B14d: Distribution of months by type, excluding extreme irregular values.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	0.958	1.825	2.361
1986	1.825	1.378	1.798	**7.358**	0.534	0.419	1.273	2.572	0.459	0.875	0.341	0.505
1987	**4.126**	1.563	1.310	0.860	1.582	0.547	0.977	1.101	**0.115**	1.641	0.903	0.171
1988	0.413	2.248	0.005	0.444	0.853	0.128	0.150	0.645	0.802	2.134	0.515	1.465
1989	0.751	0.269	2.550	1.593	1.761	1.764	0.781	0.599	0.054	1.131	0.838	0.902
1990	0.234	0.890	0.067	0.148	0.951	0.903	0.634	1.123	1.448	0.188	0.063	0.774
1991	1.567	0.102	0.143	1.355	0.706	0.486	1.837	0.633	0.768	0.152	0.105	1.283
1992	0.416	0.985	1.216	0.002	0.568	0.307	0.070	1.287	0.748	1.104	0.137	1.378
1993	0.752	1.978	0.518	1.808	1.589	0.320	0.033	0.840	1.107	0.605	0.041	0.016
1994	1.528	0.441	1.184	0.319	1.588	0.721	0.074	2.015	0.821	0.623	0.437	0.205
1995	0.015	0.999	1.075	.	.	.	.	.	.	.	.	.

TABLE 4.40. B14e: Deviations from the averages in absolute value.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	0.917	3.332	5.576
1986	3.330	1.898	3.231	**54.144**	0.285	0.175	1.620	6.617	0.211	0.766	0.116	0.255
1987	**17.022**	2.442	1.717	0.740	2.504	0.300	0.955	1.212	0.013	2.692	0.815	0.029
1988	0.171	5.052	0.000	0.197	0.727	0.016	0.022	0.415	0.644	4.555	0.266	2.147
1989	0.564	0.072	6.501	2.536	3.101	3.112	0.609	0.359	0.003	1.280	0.702	0.813
1990	0.055	0.792	0.004	0.022	0.905	0.815	0.403	1.260	2.095	0.035	0.004	0.600
1991	2.456	0.010	0.021	1.835	0.498	0.236	3.374	0.401	0.590	0.023	0.011	1.646
1992	0.173	0.971	1.478	0.000	0.323	0.094	0.005	1.656	0.560	1.218	0.019	1.899
1993	0.565	3.912	0.268	3.267	2.524	0.102	0.001	0.706	1.225	0.366	0.002	0.000
1994	2.334	0.195	1.401	0.102	2.522	0.520	0.005	4.060	0.675	0.389	0.191	0.042
1995	0.000	0.998	1.156	.	.	.	.	.	.	.	.	.

TABLE 4.41. B14f: Squares of the deviations from the averages.

Then, the new mean for months of 30 days beginning on a Tuesday is 100.889 instead of 101.967 and the squared deviation from this mean for September 1987 now is $(100.774 - 100.889)^2 = (0.115)^2 = 0.01323$ instead of $(100.774 - 101.967)^2 = (-1.193)^2 = 1.423$.

Note that in Table B14e, the extreme values identified in Step 2 are compared to their theoretical mean (here 100 as the irregular values have been multiplied by 100). For example, the absolute deviation for January 1987 is $|95.874 - 100| = 4.126$.

The new limit used to identify an extreme point is $2.5 \times 1.06 = 2.65$, and the only values of the irregular which are too far away from their average are those for April 1986 and January 1987. These values are placed in Table B14.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	107.358	.	.	.	.	.	.	.	.
1987	95.874	.	.	.	.	.	.	.	.	.	.	.
1988	.	.	.	.	.	.	.	.	.	.	.	.
1989	.	.	.	.	.	.	.	.	.	.	.	.
1990	.	.	.	.	.	.	.	.	.	.	.	.
1991	.	.	.	.	.	.	.	.	.	.	.	.
1992	.	.	.	.	.	.	.	.	.	.	.	.
1993	.	.	.	.	.	.	.	.	.	.	.	.
1994	.	.	.	.	.	.	.	.	.	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.42. B14: Extreme irregular values excluded from the trading-day regression.

4.1.15 Table B15: Preliminary Regression for Trading-Day

Description and method of calculation

Daily weights will now be estimated by means of an ordinary least squares regression carried out for data not considered extreme in Table B13 according to the method explained in Section 4.1.13.

Writing the regression model.
According to the model and notation in Section 4.1.13, we have the following:

For a multiplicative model: $N_t^* I_t - N_t = \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t$
For an additive model: $I_t = \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t$.

Deriving the daily weights.
With the additive model, the weights used for the rest of the seasonal adjustment are the β_j calculated below.
With the multiplicative model, we add 1 to the above estimates or we add the prior daily weights, if any were specified in Part A.

Comments

- With the additive model, the X-12-ARIMA software introduces an additional explanatory variable, and the estimated model becomes

$$I_t = \beta_0 (N_t - N_t^*) + \sum_{j=1}^{6} \beta_j (D_{jt} - D_{7t}) + e_t.$$

  Nevertheless, the estimation principle remains the same.

- With the multiplicative model, if prior daily weights were specified, then N_t and N_t^* are equal to the monthly sum of the prior daily weights.

- X-12-ARIMA allows to introduce in the model predefined or user-defined variables. For instance, such variables deal with calendar effects: Easter effect (see chapter 5), Labor Day, or Thanksgiving.

Example

It would be rather tedious to provide here all the intermediate regression calculations leading to Table B15, and we will not do so; we used the ordinary least squares regression methods.
If we set a type I error of 1%, for example, the tests results are interpreted as follows:

	Combined Weights	*A Priori* Weights	Regression Coefficients	Standard Deviation	T-value	Prob > t
Monday	1.081	1.000	0.081	0.093	0.872	0.192
Tuesday	1.273	1.000	0.273	0.091	2.990	0.002
Wednesday	1.047	1.000	0.047	0.095	0.494	0.311
Thursday	1.319	1.000	0.319	0.095	3.362	0.001
Friday	1.066	1.000	0.066	0.092	0.717	0.237
Saturday	0.565	1.000	-0.435	0.091	-4.772	0.000
Sunday	0.649	1.000	-0.351	0.093	-3.760	0.000

	Sum of Squares	Degrees of freedom	Mean Squares	F-Value	Prob > F
Regression	23.436	6	3.906	31.257	0.000
Error	13.246	106	0.125		
Total	36.682	112			

TABLE 4.43. B15: Preliminary trading-day regression.

- The F-test rejects the null hypothesis of the equality of the daily weights. We may therefore accept the presence of an effect due to the daily composition of the month. In fact, the probability of finding a value for the Fisher statistic greater than that found (31.257) is practically zero and therefore smaller than our type I error. We are therefore within the critical region of the test, and we cannot accept the null hypothesis of the equality of the daily weights.

- The T-tests are interpreted similarly, though we must keep in mind that the Student law is symmetric. The Prob ($T > |t|$) value must therefore be compared to half of the type I error, i.e. 0.005. All tests leading to a value smaller than 0.005 prevent us from accepting the null hypothesis for the daily weight. In our case, the weights for Tuesday, Thursday, Saturday and Sunday are considered significantly different from 0.

4.1.16 Table B16: Regression-Derived Trading-Day Adjustment Factors

Description and method of calculation

From the regression estimates, monthly adjustment factors M_t are derived for trading-day, by directly using equations (4.2) and (4.4) in Section 4.1.13:

For a multiplicative model: $M_t = \frac{1}{N_t^*} \sum_{j=1}^{7} (\beta_j + 1) D_{jt}$

For an additive model: $M_t = \sum_{j=1}^{7} \beta_j D_{jt}$

where N_t^* is the number of days in the month if prior daily weights have been provided, and otherwise 31, 30 or 28.25 depending on whether the month has 31 or 30 days or is a month of February.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	102.061	98.772	100.009
1986	101.393	99.115	97.726	**101.067**	99.840	99.099	102.061	97.678	101.180	101.393	97.380	101.294
1987	99.840	99.115	100.009	101.219	97.678	101.180	101.393	97.726	101.067	99.840	99.099	102.061
1988	97.678	102.941	102.061	98.772	100.009	101.219	97.678	101.294	101.283	97.726	101.067	99.840
1989	100.009	99.115	101.393	97.380	101.294	101.283	97.726	102.061	98.772	100.009	101.219	97.678
1990	101.294	99.115	99.840	99.099	102.061	98.772	100.009	101.393	97.380	101.294	101.283	97.726
1991	102.061	99.115	97.678	101.180	101.393	97.380	101.294	99.840	99.099	102.061	98.772	100.009
1992	101.393	101.116	100.009	101.219	97.678	101.180	101.393	97.726	101.067	99.840	99.099	102.061
1993	97.678	99.115	101.294	101.283	97.726	101.067	99.840	100.009	101.219	97.678	101.180	101.393
1994	97.726	99.115	102.061	98.772	100.009	101.219	97.678	101.294	101.283	97.726	101.067	99.840
1995	100.009	99.115	101.393	.	.	.	.	.	.	.	.	.

TABLE 4.44. B16: Trading-day adjustment factors derived from regression coefficients.

The irregular component in Table B13 is then adjusted for such calendar effects, and we get Table B16bis, which unfortunately cannot be printed by the current versions of the softwares: $B16bis = B13 \; op \; B16$.

Example

Consider for example the month of April 1986.

	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Sun.	No. of days
Weight	1.0809	1.2732	1.0469	1.3187	1.0663	0.5653	0.6487	
Number of occurrences	4	5	5	4	4	4	4	30

Thus we have:

$$\begin{aligned} APR86 &= 100 \times \frac{4 \times 1.08089 + 5 \times 1.27322 + 5 \times 1.04691}{30} + \\ & \quad 100 \times \frac{4 \times 1.31870 + 4 \times 1.06625 + 4 \times 0.56534 + 4 \times 0.64868}{30} \\ &= 101.067. \end{aligned}$$

Or again, since the days which appear 5 times are Tuesday and Wednesday:

$$APR86 = 100 \times \frac{28 + 1.27322 + 1.04691}{30} = 101.067.$$

We then get the adjusted values for the irregular component in Table B16bis. For example,

$$APR86 = 100 \times 107.358/101.067 = 106.225.$$

4.1.17 Table B17: Preliminary Weights Used to Adjust the Irregular

Description and method of calculation

On the basis of the estimate of the irregular component of Table B16bis, or of Table B13 if no adjustment is requested for trading-day, an attempt

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	98.505	101.721	97.907
1986	101.710	99.090	97.610	**106.225**	99.808	99.356	100.691	96.989	99.016	100.774	101.316	99.894
1987	96.028	102.056	101.578	101.064	101.243	98.929	100.874	98.323	99.710	98.699	100.689	99.612
1988	99.200	99.326	99.448	99.423	101.121	100.341	99.776	99.756	100.707	97.265	99.314	101.810
1989	101.019	100.751	97.396	102.601	98.654	101.657	100.248	98.856	99.818	99.137	101.042	100.546
1990	100.161	99.582	100.410	99.927	100.376	98.958	100.902	101.018	99.479	100.207	99.977	98.657
1991	100.979	100.377	99.476	100.809	99.215	100.467	102.206	99.709	99.003	99.592	99.766	98.985
1992	99.501	99.870	99.053	100.216	100.204	99.167	99.842	100.766	100.564	101.448	99.916	98.093
1993	98.853	102.475	99.881	98.130	101.075	100.140	100.309	101.108	99.121	100.242	99.430	99.926
1994	101.013	100.034	98.284	99.550	101.856	99.502	99.699	102.382	99.104	100.087	99.392	100.548
1995	100.283	99.472	98.850	.	.	.	.	.	.	.	.	.

TABLE 4.45. B16bis: Irregular component adjusted for regression-derived trading-day effects.

is made to identify and adjust extreme values. To this end, we use the algorithm described for Tables B4 and B9 (see Sections 4.1.4 and 4.1.9 respectively) to detect extreme values and calculate adjustment weights. Since we already have an estimate of the irregular, only Steps 4 and 5 need be applied.

Comments

The comments made for Tables B4 and B9 concerning the calculation of moving standard deviations and adjustment weights are equally valid here.

Example

Calculating a moving standard deviation.

The standard deviation corresponding to the year 1989 is calculated on the basis of data for the years 1987 to 1991 (two years before, two years after) using the formula[16]:

$$\sigma_{89} = \left[\frac{1}{60} \sum_{t=Jan87}^{Dec91} (I_t - 100)^2 \right]^{1/2} = 1.198.$$

Those for the years 1988, 1990, 1991 and 1992 are calculated using the same principle.

For X-11-ARIMA and X-12-ARIMA, the standard deviation for 1987 is calculated on the basis of all available observations from 1985 to 1990, i.e. 63 observations. These first estimates of the standard deviations are shown in Table B17a, in the Standard Deviation 1 column.

These first estimates are used to identify the extreme values. As can be seen from Figure 4.6, the values for April 1986, January 1987, February 1993 and August 1994 are found to be very extreme. In fact, we have, e.g.:

[16] Here the theoretical mean is assumed to be 100 as the irregular values have been multiplied by 100.

Year	Standard Deviation 1	Standard Deviation 2
1985	1.5282	1.2322
1986	1.5282	1.2322
1987	1.5282	1.2322
1988	1.5142	1.1965
1989	1.1979	1.0918
1990	1.0200	1.0200
1991	1.0173	0.9740
1992	0.9484	0.8527
1993	0.9399	0.8479
1994	0.9399	0.8479
1995	0.9399	0.8479

TABLE 4.46. B17a: Five-year moving standard deviations.

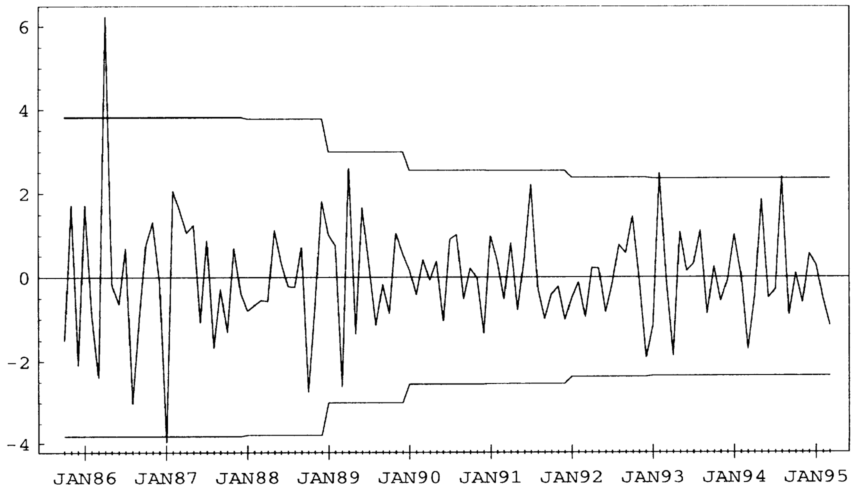


FIGURE 4.6. B17: Deviations of the irregular component from its theoretical average and upper detection thresholds $\pm 2.5\sigma$.

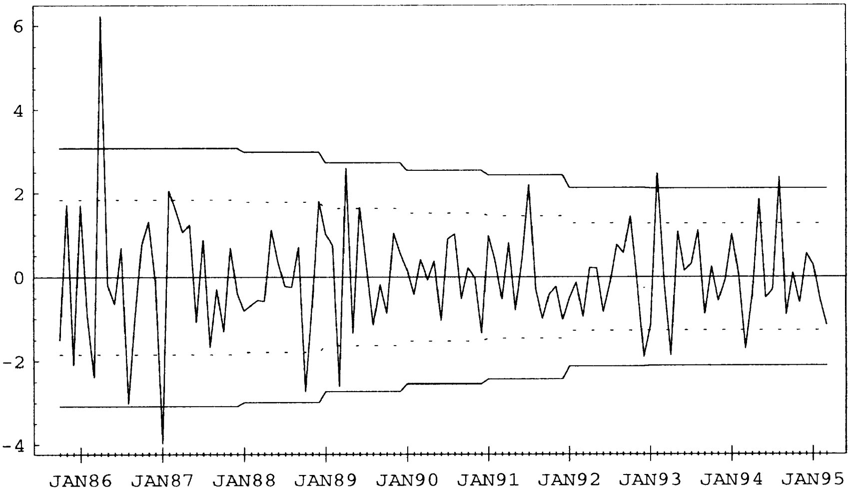


FIGURE 4.7. B17: Deviations of the irregular component from its theoretical average and detection thresholds ($\pm 1.5\sigma$ and $\pm 2.5\sigma$).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	80.138
1986	100.000	100.000	56.025	**0.000**	100.000	100.000	100.000	5.658	100.000	100.000	100.000	100.000
1987	0.000	83.133	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	**21.455**	100.000	98.701
1989	100.000	100.000	11.498	11.770	100.000	98.258	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	100.000	100.000	100.000	100.000	23.528	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	80.157	100.000	26.391
1993	100.000	0.000	100.000	29.466	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	47.606	100.000	31.094	100.000	100.000	0.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

TABLE 4.47. B17: Preliminary weights for the irregular component. Limits used are for 1.5σ to 2.5σ.

$|APR86 - 100| = |106.225 - 100| = 6.225 > 2.5 \times \sigma_{86} = 2.5 \times 1.5282 = 3.8205$, and $|AUG94 - 100| = |102.382 - 100| = 2.382 > 2.5 \times \sigma_{94} = 2.5 \times 0.94 = 2.35$. These points are therefore removed from the second calculation of the moving standard deviations, which leads to the results in the Standard Deviation 2 column in Table B17a.

Detecting and adjusting extreme values.
Values of the irregular are then classified in relation to the upper and lower detection thresholds calculated on the basis of the new standard deviations estimates. All values beyond the lower detection thresholds (see Figure 4.7) are considered extreme, and are therefore adjusted in varying degrees. The weights (multiplied by 100) associated with each of these values are shown in Table B17.

The values previously considered very extreme remain so, and are assigned zero weight. Thus, $|APR86 - 100| = |106.225 - 100| = 6.225 >$

$2.5 \times \sigma_{86} = 2.5 \times 1.2322 = 3.0805$. For October 1988, located between the two detection thresholds and therefore considered moderately extreme, we have $|OCT88-100| = |97.265-100| = 2.735$, and $1.5\times\sigma_{88} = 1.5\times 1.1965 = 1.79475 < 2.735 < 2.5\times\sigma_{88} = 2.5\times 1.1965 = 2.99125$. This value is assigned a weight proportional to the deviation from its average, i.e.:

$$weight(OCT88) = \frac{2.99125 - 2.735}{2.99125 - 1.79475} = 0.214.$$

4.1.18 Table B18: Combined Trading-Day Factors (derived from the prior daily weights and the regression for trading-day)

Description and method of calculation

If prior daily weights are provided for trading-day effects (multiplicative model only), and if a trading-day regression is also requested, Table B18 shows the combined result for both adjustments, a simple addition of the two effects. These combined daily weights will be used to estimate adjustment factors for each month, as was done in Table B16.

For the **additive model**, this table is not produced since, in this case, prior daily weights cannot be specified.

For the **multiplicative model**, we calculate $M_t = \sum_{j=1}^{7} \alpha_j D_{jt}/N_t^*$ where D_{jt} is the number of days j (Monday, Tuesday, Wednesday, ..., Sunday) contained in month t, $(\alpha_1, \ldots, \alpha_7)$ are the combined weights for each day from Table B15, and N_t^* is equal to the number of days in the month if prior daily weights have been provided, and otherwise to 31, 30 or 28.25 depending on whether the month has 31 or 30 days or is a month of February.

Comment

In X-12-ARIMA, if holiday factors (Easter, Labor Day, Thanksgiving, etc.) are also estimated in the irregular component regression, they are included in Table B18.

Example

In our case, Table B18 is identical to Table B16.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	102.061	98.772	100.009
1986	101.393	99.115	97.726	101.067	99.840	99.099	102.061	97.678	101.180	101.393	97.380	101.294
1987	99.840	99.115	100.009	101.219	97.678	101.180	101.393	97.726	101.067	99.840	99.099	102.061
1988	97.678	102.941	102.061	98.772	100.009	101.219	97.678	101.294	101.283	97.726	101.067	99.840
1989	100.009	99.115	101.393	97.380	101.294	101.283	97.726	102.061	98.772	100.009	101.219	97.678
1990	101.294	99.115	99.840	99.099	102.061	98.772	100.009	101.393	97.380	101.294	101.283	97.726
1991	102.061	99.115	97.678	101.180	101.393	97.380	101.294	99.840	99.099	102.061	98.772	100.009
1992	101.393	101.116	100.009	101.219	97.678	101.180	101.393	97.726	101.067	99.840	99.099	102.061
1993	97.678	99.115	101.294	101.283	97.726	101.067	99.840	100.009	101.219	97.678	101.180	101.393
1994	97.726	99.115	102.061	98.772	100.009	101.219	97.678	101.294	101.283	97.726	101.067	99.840
1995	100.009	99.115	101.393	.	.	.	.	.	.	.	.	.

TABLE 4.48. B18: Combined adjustment factors for trading-day effects.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	113.364	111.165	100.591
1986	105.135	99.581	106.318	**108.344**	97.857	104.643	97.687	67.262	103.973	115.491	111.214	103.067
1987	100.661	104.121	112.890	105.810	102.377	107.037	100.401	70.299	107.552	117.088	115.743	107.779
1988	110.260	107.051	116.303	109.444	107.390	113.319	103.605	75.029	113.148	120.644	120.019	114.884
1989	117.889	113.202	118.548	117.786	109.089	118.776	108.058	77.797	115.620	126.689	125.273	115.379
1990	119.553	113.504	123.799	117.156	113.266	118.252	111.790	82.155	117.683	130.314	125.490	113.379
1991	120.810	113.807	122.136	118.007	111.743	119.840	113.827	81.731	117.459	129.727	126.352	115.790
1992	121.803	115.610	123.989	118.555	112.410	117.315	110.560	81.862	118.040	129.207	123.211	111.502
1993	116.403	114.110	121.133	112.753	110.411	115.864	108.274	79.693	113.418	123.876	120.280	113.223
1994	119.007	112.496	121.496	116.835	113.990	119.543	112.103	84.309	119.072	129.342	126.352	120.193
1995	124.089	117.338	128.411	.	.	.	.	.	.	.	.	.

TABLE 4.49. B19: Original series adjusted for trading-day effects.

4.1.19 Table B19: Raw Series Corrected for Trading-Day Effects

Description and method of calculation

The series in Table B1, or in Table A1 if no prior adjustment is requested, is corrected for trading-day effects estimated in Table B18. We thus have: $B19 = B1 \ op \ B18$.

Example

For example, we have:

$$APR86 = 100 \times 109.500/1.011 = 108.344.$$

4.1.20 Table B20: Adjustment Values for Extreme Values of the Irregular

Description and method of calculation

Values of irregular component B16bis, or B13 if no regression is requested for trading-day, identified as extreme when Table B17 was set up, and for which a weight was therefore calculated, are adjusted as follows:

$$\text{Additive model:} \quad B20 = B16bis \times (1 - B17)$$
$$\text{Multiplicative model:} \quad B20 = B16bis/[1 + B17 \times (B16bis - 1)].$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	99.577
1986	100.000	100.000	98.935	106.225	100.000	100.000	100.000	97.155	100.000	100.000	100.000	100.000
1987	96.028	100.341	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	**97.839**	100.000	100.023
1989	100.000	100.000	97.689	102.288	100.000	100.028	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	100.000	100.000	100.000	100.000	101.678	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.284	100.000	98.589
1993	100.000	102.475	100.000	98.674	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	99.093	100.000	101.272	100.000	100.000	102.382	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

TABLE 4.50. B20: Weights for extreme values.

Or, in symbolic notation:

$$B20 = B16bis\ op\ [\ xbar\ + B17 \times (B16bis -\ xbar\ )].$$

Comments

We have here the values that will be used to adjust the initial series. A point considered very extreme will therefore receive a weight of zero and an adjustment value equal to the value of the irregular; in other words, in this case, we will remove from the initial series the irregular corresponding to that date (see Table C1, Section 4.2.1).

Example

The value for October 1988, identified as extreme and assigned a weight equal to 0.21455, will therefore be adjusted as follows, since we have here a multiplicative model:

$$OCT88 = 100 \times \frac{0.97265}{1 + 0.21455 \times (0.97265 - 1)} = 97.839.$$

4.2 PART C: Final Estimation of Extreme Values and Calendar Effects

4.2.1 Table C1: Raw Series Corrected for A Priori Adjustments, Trading-Day Adjustments and Extreme Values

Description and method of calculation

This table shows the raw series adjusted for various effects detected in Part B: for points considered extreme and for trading-day effects, and adjusted *a priori* using elements from Part A. It is therefore calculated on the basis

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	113.364	111.165	101.018
1986	105.135	99.581	107.463	**101.995**	97.857	104.643	97.687	69.231	103.973	115.491	111.214	103.067
1987	104.825	103.768	112.890	105.810	102.377	107.037	100.401	70.299	107.552	117.088	115.743	107.779
1988	110.260	107.051	116.303	109.444	107.390	113.319	103.605	75.029	113.148	123.308	120.019	114.858
1989	117.889	113.202	121.354	115.151	109.089	118.742	108.058	77.797	115.620	126.689	125.273	115.379
1990	119.553	113.504	123.799	117.156	113.266	118.252	111.790	82.155	117.683	130.314	125.490	113.379
1991	120.810	113.807	122.136	118.007	111.743	119.840	111.949	81.731	117.459	129.727	126.352	115.790
1992	121.803	115.610	123.989	118.555	112.410	117.315	110.560	81.862	118.040	128.841	123.211	113.098
1993	116.403	111.354	121.133	114.269	110.411	115.864	108.274	79.693	113.418	123.876	120.280	113.223
1994	119.007	112.496	122.608	116.835	112.558	119.543	112.103	82.348	119.072	129.342	126.352	120.193
1995	124.089	117.338	128.411	.	.	.	.	.	.	.	.	.

TABLE 4.51. C1: Trading-day adjusted series modified by preliminary weights.

of Table B19, which takes into consideration the effects due to trading-day, or of Table B1 if no regression for trading-day is requested, and of Table B20 which indicates adjustments to be made for points considered extreme. We therefore have: $C1 = B19 \ op \ B20$.

Example

For example, the value for the month of April 1986, considered extreme, becomes:

$$APR86 = 100 \times 108.344/106.225 = 101.995.$$

4.2.2 Table C2: Preliminary Estimation of the Trend-Cycle

Description and method of calculation

A new estimate of the trend-cycle component is obtained by applying to the data in Table C1 a centered moving average of order 12, just as for Table B2 (Section 4.1.2).

Comments

- X-11-ARIMA and X-12-ARIMA also include a centered 24-term moving average due to Cholette [13].
- The first six and last six points in the series are not imputed at this stage of the calculations.

Example

For example, the value for April 1986 is obtained from the values in Table C1 from October 1985 to October 1986 (six months before and six months after):

$$APR86 = \frac{113.364}{24} +$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**101.181**	101.272	101.359	101.432	101.593	101.994	102.379	102.726	103.014
1987	103.227	103.385	103.578	103.794	104.049	104.434	104.857	105.220	105.500	105.793	106.153	106.624
1988	107.019	107.350	107.780	108.273	108.710	109.183	109.796	110.370	110.837	111.285	111.594	111.890
1989	112.302	112.603	112.821	113.065	113.425	113.665	113.756	113.838	113.953	114.138	114.396	114.549
1990	114.684	115.021	115.289	115.526	115.686	115.612	115.581	115.646	115.589	115.555	115.527	115.530
1991	115.603	115.592	115.565	115.531	115.542	115.679	115.821	115.937	116.089	116.189	116.240	116.163
1992	115.999	115.947	115.977	115.964	115.796	115.553	115.216	114.814	114.517	114.220	113.958	113.814
1993	113.658	113.473	113.190	112.790	112.461	112.344	112.458	112.614	112.723	112.892	113.088	113.331
1994	113.644	113.914	114.260	114.723	115.204	115.748	116.250	116.663	117.107	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.52. C2: Trend-cycle, centered 12-term moving average.

$$
\begin{aligned}
& \frac{111.165 + 101.018 + 105.135 + 99.581 + 107.463 + 101.995}{12} + \\
& \frac{97.857 + 104.643 + 97.687 + 69.231 + 103.973}{12} + \\
& \frac{115.491}{24} \\
= {} & 101.181.
\end{aligned}
$$

4.2.3 Table C4: Preliminary Estimation of the Modified Seasonal-Irregular Component

Description and method of calculation

The trend-cycle component is removed from the analyzed series, and an estimate of the seasonal-irregular component is obtained. Thus we have: $C4 = C1 \ op\ C2$.

Comments

- Again, there is no estimate for the six values at the beginning and six values at the end of the series.

- This table corresponds to Table B4g (Section 4.1.4); here there is no calculation of replacement values for extreme values of the seasonal-irregular component since such values have already been reweighted using the weights in Table B20.

Example

The value for April 1986 is therefore obtained simply as follows:

$$APR86 = 100 \times 101.995/101.181 = 100.804.$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**100.804**	96.628	103.240	96.308	68.146	101.940	112.807	108.262	100.051
1987	101.548	100.370	108.990	101.943	98.393	102.492	95.751	66.811	101.946	110.676	109.034	101.083
1988	103.028	99.722	107.908	101.082	98.786	103.788	94.362	67.980	102.085	110.804	107.550	102.652
1989	104.976	100.532	107.563	101.846	96.177	104.467	94.991	68.340	101.463	110.996	109.509	100.724
1990	104.246	98.681	107.381	101.411	97.908	102.284	96.720	71.041	101.812	112.772	108.623	98.138
1991	104.505	98.456	105.686	102.143	96.712	103.597	96.657	70.496	101.180	111.651	108.699	99.679
1992	105.003	99.709	106.908	102.234	97.075	101.525	95.959	71.300	103.076	112.801	108.119	99.370
1993	102.414	98.132	107.018	101.311	98.177	103.132	96.279	70.766	100.616	109.730	106.360	99.905
1994	104.719	98.755	107.306	101.840	97.703	103.279	96.433	70.586	101.678	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.53. C4: Modified seasonal-irregular ratios.

4.2.4 Table C5: Estimation of the Seasonal Component

Description and method of calculation

The estimation is done on the basis of the values for the seasonal-irregular component in Table C4. The process consists in three steps similar to those explained for Table B5 (Section 4.1.5):

- Step 1: Estimating the seasonal component using a 3x3 moving average.
- Step 2: Normalizing the seasonal factors using a centered 12-term moving average.
- Step 3: Imputing the missing seasonal factors.

Comment

We can specify the moving average to be used. In such a case, X-11-ARIMA gives the choice between a simple 3-term, a 3×3, a 3×5, a 3×9 and constant seasonality (a simple average). X-12-ARIMA also includes a 3×15.

Example

The estimation is based on the modified seasonal-irregular component in Table C4.

Step 1: Estimating the seasonal component.

The data in Table C4 are smoothed column by column (month by month), using a 3×3 moving average of coefficients $\{1, 2, 3, 2, 1\}/9$, and Table C5a is obtained.

The seasonal factor for the month of April 1988 is therefore estimated as follows:

$$\begin{aligned} APR88 &= \frac{100.804 + 2 \times 101.943 + 3 \times 101.082 + 2 \times 101.846 + 101.411}{9} \\ &= 101.449. \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	101.319	97.746	103.037	95.720	67.571	101.969	111.568	108.445	100.953
1987	102.786	100.136	108.285	**101.414**	97.791	103.241	95.451	67.630	101.927	111.297	108.502	101.182
1988	103.284	99.984	108.040	**101.449**	97.782	103.423	95.288	68.159	101.870	111.260	108.514	101.084
1989	103.947	99.692	107.549	101.623	97.447	103.515	95.505	68.930	101.701	111.385	108.734	100.502
1990	104.414	99.272	107.051	101.725	97.263	103.144	95.975	70.008	101.765	111.913	108.662	99.694
1991	104.378	98.979	106.691	101.875	97.162	102.890	96.289	70.586	101.710	111.870	108.383	99.409
1992	104.201	98.860	106.758	101.874	97.402	102.622	96.322	70.895	101.812	111.704	107.869	99.452
1993	103.915	98.739	106.916	101.780	97.606	102.805	96.278	70.828	101.592	111.337	107.510	99.645
1994	103.833	98.678	107.115	101.698	97.780	102.895	96.282	70.792	101.504	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.54. C5a: Preliminary seasonal factors (3 × 3 ma).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**99.965**	99.965	99.965	99.965	99.965	99.965	99.965	99.971	99.981
1987	99.979	99.970	99.971	99.958	99.949	99.961	99.991	100.005	99.989	99.980	99.981	99.988
1988	99.989	100.004	100.024	100.020	100.019	100.015	100.039	100.054	100.022	100.009	100.002	99.992
1989	100.005	100.046	100.071	100.069	100.084	100.068	100.064	100.066	100.027	100.011	100.007	99.984
1990	99.988	100.053	100.100	**100.125**	100.144	100.107	100.072	100.059	100.031	100.023	100.025	100.010
1991	100.012	100.049	100.071	100.067	100.054	100.030	100.011	99.999	99.997	99.999	100.009	100.008
1992	99.998	100.013	100.030	100.027	99.999	99.979	99.969	99.952	99.953	99.956	99.961	99.977
1993	99.983	99.978	99.966	99.942	99.911	99.905	99.909	99.903	99.909	99.914	99.918	99.929
1994	99.932	99.931	99.926	99.926	99.926	99.926	99.926	99.926	99.926	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.55. C5b: Centered 12-month moving average.

This symmetric moving average can be used to estimate the seasonal factors for the years 1988 to 1992. For the beginning of the series (years 1986 and 1987) and the end of the series (years 1993 and 1994), predefined asymmetric averages are used (see Table 3.11). Thus, for April 1987 for example, one point in the past, the current point and two points in the future are used:

$$\begin{aligned} APR87 &= \frac{100.804 \times 7 + 101.943 \times 10 + 101.082 \times 7 + 101.846 \times 3}{27} \\ &= 101.414. \end{aligned}$$

Step 2: Normalizing the seasonal factors.

On Table C5a, a centered 12-month moving average is used to produce Table C5b. The first computable term is therefore that for October 1986, and the last that for March 1994. Thus, for example,

$$\begin{aligned} APR90 &= \frac{111.385}{24} + \\ &\quad \frac{108.734 + 100.502 + 104.414 + 99.272 + 107.051 + 101.725}{12} + \\ &\quad \frac{97.263 + 103.144 + 95.975 + 70.008 + 101.765}{12} + \\ &\quad \frac{111.913}{24} \\ &= 100.125. \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.607	108.476	100.972
1986	102.808	100.166	108.317	**101.355**	97.780	103.072	95.754	67.595	102.005	111.607	108.476	100.972
1987	102.808	100.166	108.317	101.457	97.841	103.282	95.459	67.626	101.938	111.319	108.522	101.194
1988	103.296	99.980	108.014	101.428	97.763	103.407	95.251	68.122	101.847	111.250	108.512	101.093
1989	103.943	99.646	107.473	101.553	97.366	103.444	95.444	68.885	101.673	111.373	108.726	100.517
1990	104.426	99.220	106.943	101.598	97.123	103.033	95.906	69.967	101.733	111.888	108.635	99.684
1991	104.365	98.930	106.615	101.807	97.110	102.858	96.278	70.587	101.714	111.870	108.373	99.401
1992	104.203	98.848	106.726	101.846	97.403	102.643	96.352	70.929	101.860	111.752	107.912	99.475
1993	103.933	98.760	106.952	101.839	97.692	102.904	96.365	70.896	101.685	111.433	107.599	99.716
1994	103.903	98.746	107.194	101.773	97.853	102.971	96.354	70.844	101.580	111.433	107.599	99.716
1995	103.903	98.746	107.194	.	.	.	.	.	.	.	.	.

TABLE 4.56. C5: Seasonal factors.

The first six values, from April to September 1986, which cannot be computed using this symmetric moving average, are taken to be equal to the first computable value, that for October 1986 (99.965). The same applies to the end of the series: the value calculated for March 1994 (99.926) is repeated for the next six months. Normalized seasonal factors are then obtained by dividing Table C5a by Table C5b, and we get Table C5. For example,

$$APR86 = 100 \times 101.319/99.965 = 101.355.$$

Step 3: Imputing the missing seasonal factors.

Finally, to obtain the seasonal factors for the six missing values at either end of the series (October 1985 to March 1986 and October 1994 to March 1995), caused by the use of the centered 12-term moving average in Table C2, we repeat the nearest available factor for that particular month.

4.2.5 Table C6: Estimation of the Seasonally Adjusted Series

Description and method of calculation

The estimation is done simply by removing from the series in Table C1 the estimate of the seasonal component in Table C5: $C6 = C1 \; op \; C5$.

Example

For example, $APR86 = 100 \times 101.995/101.355 = 100.632$.

4.2.6 Table C7: Estimation of the Trend-Cycle Component

Description and method of calculation

This table shows an estimate of the trend-cycle component obtained from the seasonally adjusted series in Table C6. As for Table B7 (Section 4.1.7), the program uses a Henderson moving average whose order depends on the value of the $\bar{I}/\bar{C}$ ratio.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	101.574	102.479	100.046
1986	102.264	99.416	99.212	**100.632**	100.078	101.524	102.019	102.421	101.929	103.480	102.524	102.075
1987	101.962	103.595	104.222	104.291	104.636	103.635	105.177	103.952	105.507	105.182	106.654	106.507
1988	106.742	107.073	107.674	107.903	109.848	109.586	108.771	110.140	111.096	110.839	110.605	113.616
1989	113.418	113.603	112.915	113.391	112.040	114.789	113.216	112.938	113.717	113.751	115.219	114.785
1990	114.486	114.397	115.761	115.313	116.621	114.771	116.562	117.420	115.679	116.469	115.515	113.738
1991	115.757	115.038	114.558	115.913	115.069	116.509	116.277	115.788	115.480	115.962	116.589	116.487
1992	116.890	116.957	116.175	116.406	115.407	114.294	114.745	115.415	115.885	115.292	114.177	113.695
1993	111.997	112.752	113.259	112.205	113.019	112.594	112.357	112.407	111.539	111.166	111.786	113.545
1994	114.536	113.924	114.379	114.800	115.029	116.094	116.345	116.238	117.220	116.071	117.429	120.535
1995	119.428	118.829	119.793	.	.	.	.	.	.	.	.	.

TABLE 4.57. C6: Seasonally adjusted series.

Step 1: Selecting the moving average, calculating the $\bar{I}/\bar{C}$ ratio.

- If the ratio is smaller than 1, a 9-term Henderson moving average is selected.
- If the ratio is greater than 3.5, a 23-term Henderson moving average is selected.
- Otherwise, a 13-term Henderson moving average is selected.

Step 2: Smoothing the SA series using a Henderson moving average.

Comments

- At this stage, unlike in Part B, the program selects between a 9-term average, a 13-term average and a 23-term average.
- It is possible to specify the length of the Henderson moving average to be used. X-11-ARIMA provides a choice between a 9-term, a 13-term or a 23-term moving average. X-12-ARIMA allows any odd-numbered average less than 101.
- As the series in Table C1 has been corrected for extreme values, the numerator of the $\bar{I}/\bar{C}$ ratio should be smaller than the one calculated in Part B.

Example

Step 1: Selecting the moving average, calculating the $\bar{I}/\bar{C}$ ratio.

First, Table C6 is smoothed using a 13-term Henderson moving average whose coefficients are shown in Table 3.8.

The first computable term is therefore that for April 1986, and we have:

$$\begin{aligned} APR86 \quad = \quad & 101.574 \times (-0.01935) + 102.479 \times (-0.02786) + \\ & 100.046 \times (0.00000) + 102.264 \times (0.06549) + \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**100.198**	100.587	101.166	101.772	102.244	102.449	102.498	102.511	102.620
1987	102.881	103.274	103.699	104.070	104.311	104.422	104.534	104.765	105.141	105.593	106.028	106.428
1988	106.850	107.321	107.802	108.319	108.846	109.307	109.680	110.054	110.537	111.161	111.891	112.540
1989	112.988	113.249	113.355	113.338	113.257	113.240	113.358	113.564	113.811	114.073	114.323	114.591
1990	114.852	115.028	115.234	115.503	115.820	116.153	116.348	116.353	116.189	115.874	115.489	115.137
1991	114.922	114.947	115.150	115.426	115.680	115.834	115.909	115.959	116.015	116.137	116.345	116.579
1992	116.732	116.654	116.328	115.875	115.474	115.233	115.181	115.203	115.108	114.803	114.281	113.645
1993	113.087	112.747	112.638	112.664	112.668	112.524	112.229	111.935	111.829	111.980	112.387	112.970
1994	113.584	114.142	114.586	114.967	115.342	115.681	116.010	116.382	116.837	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.58. C7a: Trend-cycle (13-term Henderson ma).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.433	99.494	100.354	100.243	100.173	99.493	100.958	100.013	99.469
1987	99.107	100.311	100.504	100.212	100.311	99.247	100.615	99.224	100.349	99.610	100.590	100.074
1988	99.899	99.768	99.881	99.616	100.921	100.255	99.170	100.079	100.505	99.710	98.851	100.956
1989	100.381	100.313	99.612	100.046	98.926	101.368	99.875	99.449	99.917	99.718	100.784	100.169
1990	99.681	99.451	100.458	99.836	100.692	98.810	100.184	100.917	99.561	100.513	100.023	98.785
1991	100.727	100.079	99.485	100.422	99.472	100.583	100.317	99.852	99.539	99.849	100.210	99.921
1992	100.135	100.260	99.869	100.458	99.942	99.185	99.621	100.184	100.675	100.426	99.909	100.044
1993	99.036	100.004	100.552	99.593	100.312	100.063	100.115	100.422	99.740	99.274	99.466	100.509
1994	100.839	99.809	99.819	99.855	99.729	100.357	100.289	99.876	100.329	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.59. C7b: Irregular component.

$$\begin{aligned}
& 99.416 \times (0.14736) + 99.212 \times (0.21434) + \\
& 100.632 \times (0.24006) + 100.078 \times (0.21434) + \\
& 101.524 \times (0.14736) + 102.019 \times (0.06549) + \\
& 102.421 \times (0.00000) + 101.929 \times (-0.02786) + \\
& 103.480 \times (-0.01935) \\
= \; & 100.198.
\end{aligned}$$

At this step in the calculation, there is no attempt to estimate the six points that cannot be calculated at the beginning and end of the series. An estimate is derived for the trend-cycle (Table C7a), and one for the irregular component (Table C7b) by division using Table C6.

As this is a multiplicative model, the rates of growth are calculated (see Section 4.1.7).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1985	.	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	0.389	0.575	0.600	0.464	0.200	0.048	0.012	0.106	2.394
1987	0.255	0.382	0.411	0.358	0.232	0.106	0.108	0.221	0.358	0.431	0.412	0.377	3.650
1988	0.397	0.441	0.448	0.479	0.486	0.424	0.342	0.340	0.439	0.564	0.657	0.580	5.598
1989	0.398	0.231	0.094	0.015	0.072	0.015	0.104	0.182	0.218	0.230	0.219	0.235	2.012
1990	0.228	0.153	0.178	0.233	0.274	0.288	0.167	0.004	0.141	0.271	0.333	0.305	2.576
1991	0.187	0.022	0.177	0.239	0.221	0.133	0.064	0.044	0.048	0.106	0.179	0.201	1.620
1992	0.132	0.067	0.280	0.389	0.346	0.209	0.045	0.019	0.083	0.265	0.454	0.557	2.845
1993	0.491	0.301	0.097	0.023	0.004	0.128	0.262	0.261	0.095	0.134	0.364	0.519	2.679
1994	0.543	0.491	0.389	0.332	0.326	0.294	0.285	0.321	0.390	.	.	.	3.372
1995	.	.	.	.	.	.	.	.	.	.	.	.	

TABLE 4.60. C7c: Absolute growth rate (in %) of the trend-cycle.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1985	.	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	0.935	0.865	0.111	0.069	0.679	1.473	0.937	0.544	5.614
1987	0.364	1.215	0.193	0.290	0.098	1.061	1.379	1.383	1.134	0.736	0.984	0.513	9.350
1988	0.176	0.130	0.113	0.265	1.309	0.660	1.082	0.916	0.426	0.791	0.862	2.130	8.859
1989	0.570	0.067	0.699	0.436	1.120	2.469	1.473	0.427	0.471	0.199	1.069	0.610	9.611
1990	0.487	0.230	1.012	0.619	0.858	1.869	1.391	0.732	1.344	0.956	0.488	1.237	11.223
1991	1.966	0.643	0.593	0.941	0.946	1.117	0.264	0.464	0.314	0.311	0.362	0.288	8.210
1992	0.214	0.124	0.390	0.591	0.515	0.757	0.440	0.564	0.491	0.248	0.514	0.135	4.983
1993	1.007	0.978	0.547	0.954	0.722	0.248	0.052	0.306	0.678	0.468	0.193	1.049	7.202
1994	0.328	1.021	0.010	0.035	0.126	0.630	0.068	0.411	0.453	.	.	.	3.083
1995	.	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.61. C7d: Absolute growth rate (in %) of the irregular.

Using the row totals from Tables C7c and C7d, we have:

$$\begin{aligned}\bar{C} &= \frac{2.394 + 3.650 + 5.598 + 2.012 + 2.576}{101} + \\ &\quad \frac{1.620 + 2.845 + 2.679 + 3.372}{101} \\ &= 0.2648,\end{aligned}$$

(in Table B7, this quantity was equal to 0.312), and

$$\begin{aligned}\bar{I} &= \frac{5.614 + 9.350 + 8.859 + 9.611 + 11.223}{101} + \\ &\quad \frac{8.210 + 4.983 + 7.202 + 3.083}{101} \\ &= 0.6746\end{aligned}$$

(in Table B7, this quantity was equal to 2.226), therefore $\bar{I}/\bar{C} = 0.6746/0.2648 = 2.547$ (in Table B7, this quantity was equal to 7.14).

Step 2: Smoothing the SA series using a Henderson moving average.

Since the ratio is greater than 1 and smaller than 3.5, we select a 13-term Henderson moving average whose coefficients and those of the associated asymmetric moving averages are shown in Table 3.8. The trend-cycle for October 1985 is estimated from the seasonally adjusted series in Table C6, using the current point and six points in the future, to which are assigned the coefficients of the $H6\_0$ moving average in Table 3.8.

$$\begin{aligned}OCT85 &= 101.574 \times (0.42113) + 102.479 \times (0.35315) + \\ &\quad 100.046 \times (0.24390) + 102.264 \times (0.11977) + \\ &\quad 99.416 \times (0.01202) + 99.212 \times (-0.05811) + \\ &\quad 100.632 \times (-0.09186) \\ &= 101.801.\end{aligned}$$

This leads to Table C7.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	**101.801**	101.494	101.102
1986	100.683	100.300	100.105	100.198	100.587	101.166	101.772	102.244	102.449	102.498	102.511	102.620
1987	102.881	103.274	103.699	104.070	104.311	104.422	104.534	104.765	105.141	105.593	106.028	106.428
1988	106.850	107.321	107.802	108.319	108.846	109.307	109.680	110.054	110.537	111.161	111.891	112.540
1989	112.988	113.249	113.355	113.338	113.257	113.240	113.358	113.564	113.811	114.073	114.323	114.591
1990	114.852	115.028	115.234	115.503	115.820	116.153	116.348	116.353	116.189	115.874	115.489	115.137
1991	114.922	114.947	115.150	115.426	115.680	115.834	115.909	115.959	116.015	116.137	116.345	116.579
1992	116.732	116.654	116.328	115.875	115.474	115.233	115.181	115.203	115.108	114.803	114.281	113.645
1993	113.087	112.747	112.638	112.664	112.668	112.524	112.229	111.935	111.829	111.980	112.387	112.970
1994	113.584	114.142	114.586	114.967	115.342	115.681	116.010	116.382	116.837	117.399	118.026	118.651
1995	119.188	119.603	119.876	.	.	.	.	.	.	.	.	.

TABLE 4.62. C7: Trend-cycle, I/C ratio is 2.548, a 13-term Henderson moving average has been selected.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.358	109.528	99.917
1986	104.422	99.284	107.350	**101.794**	97.285	103.438	95.986	67.712	101.487	112.676	108.490	100.436
1987	101.889	100.478	108.863	101.672	98.145	102.504	96.046	67.101	102.294	110.886	109.163	101.269
1988	103.191	99.748	107.886	101.039	98.663	103.670	94.461	68.175	102.362	110.928	107.264	102.059
1989	104.338	99.959	107.056	101.600	96.320	104.859	95.325	68.505	101.589	111.060	109.579	100.687
1990	104.093	98.675	107.433	101.432	97.795	101.807	96.083	70.609	101.286	112.462	108.660	98.473
1991	105.124	99.008	106.066	102.236	96.597	103.458	96.584	70.483	101.245	111.701	108.601	99.323
1992	104.344	99.104	106.586	102.313	97.346	101.807	95.988	71.059	102.548	112.228	107.814	99.518
1993	102.932	98.764	107.542	101.425	97.997	102.968	96.476	71.195	101.420	110.624	107.024	100.224
1994	104.774	98.558	107.001	101.625	97.587	103.338	96.632	70.757	101.913	110.173	107.055	101.300
1995	104.112	98.106	107.120	.	.	.	.	.	.	.	.	.

TABLE 4.63. C9: Modified seasonal-irregular ratios.

4.2.7 Table C9: Estimation of the Seasonal-Irregular Component

Description and method of calculation

This table is similar to Table C4: the trend-cycle component is removed from the analysed series by subtraction or division, depending on the decomposition model used, and a new estimate of the seasonal-irregular component is obtained. Thus we have: $C9 = C1 \; op \; C7$.

Comment

Unlike in Table C4, since the points at the beginning and end of the series have been estimated using asymmetric moving averages for the trend-cycle, we have a complete estimate of the seasonal-irregular component.

Example

The value for April 1986 is therefore obtained simply:

$$APR86 = 100 \times 101.995/100.198 = 101.794.$$

4.2.8 Table C10: Estimation of the Seasonal Component

Description and method of calculation

The estimation is done using a procedure similar to the one that produced Table B10 (Section 4.1.10), with values for the seasonal-irregular component in Table C9. The procedure follows two steps: estimating the seasonal component using a 3×5 moving average, then normalizing the seasonal factors.

Comment

We can specify the moving average to be used. In such a case, X-11-ARIMA gives the choice between a simple 3-term, a 3×3, a 3×5, a 3×9 and constant seasonality (a simple average). X-12-ARIMA also includes a 3×15.

Example

Step 1: Estimating the seasonal component.
The data in Table C9 are smoothed column by column (month by month), using a 3×5 moving average (see Table 3.12), and we get Table C10a. The seasonal factor for the month of April 1989 is therefore estimated as follows:

$$\begin{aligned} APR89 &= \frac{101.794 + 101.672 \times 2 + 101.039 \times 3 + 101.600 \times 3}{15} + \\ &\quad \frac{101.432 \times 3 + 102.236 \times 2 + 102.313}{15} \\ &= 101.609. \end{aligned}$$

For the beginning of the series (years 1986 to 1988) and the end of the series (years 1992 to 1994), predefined asymmetric averages are used, e.g.:

$$\begin{aligned} APR87 &= \frac{101.794 \times 15 + 101.672 \times 15 + 101.039 \times 15}{60} + \\ &= \frac{101.600 \times 11 + 101.432 \times 4}{60} \\ &= 101.515 \end{aligned}$$

(one point in the past, the current point and three points in the future).

Step 2: Normalizing the seasonal factors.
A centered 12-month moving average is applied to Table C10a. The first computable term is therefore that for April 1986, and the last that for September 1994. Thus:

$$APR86 = \frac{111.533}{24} +$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.533	108.791	100.768
1986	103.343	99.855	107.886	101.516	97.775	103.452	95.472	67.789	101.979	111.471	108.766	100.829
1987	103.444	99.782	107.814	**101.515**	97.702	103.414	95.505	68.014	101.913	111.482	108.683	100.793
1988	103.591	99.690	107.656	101.543	97.583	103.378	95.578	68.423	101.831	111.464	108.696	100.607
1989	103.844	99.500	107.394	**101.609**	97.497	103.212	95.656	69.054	101.788	111.562	108.556	100.320
1990	104.037	99.325	107.134	101.707	97.353	103.120	95.826	69.704	101.726	111.566	108.457	100.007
1991	104.212	99.074	106.956	101.777	97.340	102.925	96.044	70.319	101.702	111.576	108.183	99.808
1992	104.226	98.877	106.908	101.836	97.352	102.850	96.288	70.688	101.694	111.411	107.930	99.835
1993	104.184	98.721	106.950	101.846	97.462	102.783	96.386	70.882	101.784	111.232	107.627	100.035
1994	104.092	98.627	107.027	101.855	97.486	102.818	96.398	70.925	101.853	111.112	107.493	100.194
1995	104.000	98.570	107.126	.	.	.	.	.	.	.	.	.

TABLE 4.64. C10a: Preliminary seasonal factors (3×5 ma).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.011	100.011	100.011
1986	100.011	100.011	100.011	**100.011**	100.007	100.009	100.015	100.016	100.010	100.007	100.004	99.999
1987	99.999	100.010	100.017	100.014	100.011	100.006	100.011	100.013	100.003	99.998	99.994	99.987
1988	99.989	100.009	100.023	100.018	100.018	100.011	100.014	100.016	99.998	99.990	99.989	99.978
1989	99.974	100.004	100.029	100.031	100.029	100.011	100.007	100.008	99.990	99.983	99.981	99.971
1990	99.975	100.009	100.033	100.031	100.027	100.010	100.004	100.001	99.983	99.979	99.981	99.972
1991	99.973	100.008	100.033	100.032	100.021	100.001	99.994	99.986	99.976	99.976	99.979	99.977
1992	99.984	100.009	100.024	100.017	100.000	99.990	99.989	99.981	99.976	99.979	99.984	99.985
1993	99.987	99.999	100.011	100.007	99.987	99.983	99.987	99.979	99.979	99.982	99.984	99.986
1994	99.988	99.990	99.995	99.993	99.982	99.983	99.986	99.980	**99.982**	99.982	99.982	99.982
1995	99.982	99.982	99.982	.	.	.	.	.	.	.	.	.

TABLE 4.65. C10b: Centered 12-term moving average.

$$
\begin{aligned}
& \frac{108.791 + 100.768 + 103.343 + 99.855 + 107.886}{12} + \\
& \frac{101.516 + 97.775 + 103.452 + 95.472 + 67.789 + 101.979}{12} + \\
& \frac{111.471}{24} \\
= \;& 100.011.
\end{aligned}
$$

The first six values, from October 1985 to March 1986, which cannot be computed using this symmetric moving average, are taken to be equal to the first computable value, i.e. that for April 1986. The same procedure is used for the end of the series: the value calculated for September 1994 (99.982) is repeated for the next six months. The normalized seasonal factors, in Table C10, are then obtained by dividing Table C10a by Table C10b. For example,

$$APR86 = 100 \times 101.516/100.011 = 101.505.$$

4.2.9 *Table C11: Estimation of the Seasonally Adjusted Series*

Description and method of calculation

The estimation is done simply by removing from the starting series, in Table B1 (and not C1), the estimate of the seasonal component in Table C10: $C11 = B1 \; op \; C10$.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.521	108.779	100.758
1986	103.332	99.844	107.875	**101.505**	97.768	103.444	95.457	67.778	101.968	111.463	108.761	100.829
1987	103.444	99.772	107.796	101.500	97.691	103.408	95.495	68.004	101.910	111.485	108.689	100.806
1988	103.603	99.681	107.631	101.524	97.565	103.366	95.565	68.412	101.833	111.475	108.709	100.629
1989	103.870	99.497	107.364	101.578	97.468	103.200	95.649	69.048	101.799	111.580	108.576	100.348
1990	104.063	99.316	107.098	101.675	97.327	103.110	95.822	69.703	101.743	111.590	108.478	100.034
1991	104.240	99.066	106.921	101.744	97.319	102.924	96.050	70.329	101.727	111.603	108.206	99.832
1992	104.243	98.868	106.882	101.818	97.353	102.860	96.298	70.701	101.718	111.435	107.947	99.849
1993	104.198	98.722	106.939	101.839	97.474	102.801	96.399	70.896	101.806	111.252	107.645	100.048
1994	104.105	98.636	107.033	101.862	97.504	102.835	96.411	70.940	101.872	111.133	107.512	100.212
1995	104.019	98.588	107.145	.	.	.	.	.	.	.	.	.

TABLE 4.66. C10: Seasonal factors.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	103.747	100.938	99.844
1986	103.163	98.854	96.315	**107.876**	99.931	100.248	104.445	96.934	103.169	105.058	99.576	103.541
1987	97.154	103.436	104.735	105.517	102.364	104.731	106.603	101.023	106.663	104.858	105.530	109.121
1988	103.955	110.553	110.284	106.477	110.080	110.965	105.897	111.092	112.537	105.763	111.582	113.983
1989	113.507	112.768	111.956	112.919	113.370	116.569	110.404	114.992	112.182	113.550	116.784	112.309
1990	116.371	113.275	115.408	114.187	118.775	113.277	116.674	119.507	112.636	118.290	117.167	110.762
1991	118.284	113.863	111.578	117.353	116.421	113.385	120.042	116.026	114.424	118.635	115.336	115.995
1992	118.474	118.238	116.015	117.857	112.786	115.400	116.410	113.152	117.285	115.762	113.111	113.972
1993	109.120	114.564	114.739	112.138	110.696	113.910	112.138	112.418	112.764	108.762	113.057	114.745
1994	111.714	113.042	115.852	113.290	116.919	117.665	113.576	120.384	118.384	113.738	118.777	119.746
1995	119.305	117.965	121.517	.	.	.	.	.	.	.	.	.

TABLE 4.67. C11: Seasonally adjusted series.

As we use the B1 series, this seasonally adjusted series includes the extreme values that have been previously detected.

Example

The value for April 86 is: $APR86 = 100 \times 109.500/101.505 = 107.876$.

4.2.10 Table C13: Estimation of the Irregular Component

Description and method of calculation

The estimation is done simply by removing from the seasonally adjusted series in Table C11, the estimate of the trend-cycle component in Table C7: $C13 = C11$ *op* $C7$.

Example

The value for April 86 is:

$$APR86 = 100 \times 107.876/100.198 = 107.663.$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	101.911	99.452	98.755
1986	102.463	98.559	96.214	**107.663**	99.347	99.093	102.626	94.806	100.703	102.497	97.137	100.898
1987	94.433	100.157	100.999	101.390	98.133	100.296	101.979	96.428	101.448	99.303	99.531	102.530
1988	97.290	103.011	102.302	98.300	101.134	101.516	96.550	100.944	101.809	95.144	99.724	101.282
1989	100.459	99.576	98.765	99.630	100.100	102.940	97.394	101.258	98.569	99.542	102.153	98.008
1990	101.323	98.475	100.151	98.861	102.552	97.524	100.281	102.711	96.942	102.085	101.453	96.200
1991	102.926	99.057	96.898	101.670	100.640	97.885	103.566	100.058	98.629	102.151	99.133	99.499
1992	101.492	101.357	99.731	101.710	97.672	100.145	101.066	98.220	101.892	100.836	98.976	100.288
1993	96.492	101.612	101.865	99.533	98.250	101.232	99.920	100.431	100.835	97.127	100.597	101.571
1994	98.354	99.036	101.105	98.542	101.367	101.715	97.902	103.439	101.325	96.882	100.637	100.923
1995	100.098	98.631	101.369	.	.	.	.	.	.	.	.	.

TABLE 4.68. C13: Irregular series.

4.2.11 Table C14: Values of the Irregular Component Excluded from the Trading-Day Regression

Description and method of calculation

Tables C14, C15, and C16 deal with the final estimation of the effect of the daily composition of the month. In Table C14, X-11 identifies extreme values of the irregular component and excludes them from the calculations. The process involves two steps:

Step 1: First calculation of an overall standard deviation and identification of extreme values.

Part B provides us with a first estimate of the trading-day effect, presented in Table B16 (or B18 if the user has requested an *a priori* adjustment); see Section 4.1.16. The irregular component in Table C13 is then adjusted for this effect, and a residual is obtained[17]: $Res = C13 - B16$.

An estimate of the variance of this residual is then calculated using the mean of the squares of the residual (and doing so we assume the residuals have a mean equal to 0). We thus have:

$$\begin{aligned} \sigma^2 &= \frac{1}{n}\sum_{t=1}^{n}(C13_t - B16_t)^2 \\ &= \frac{1}{n}\sum_{t=1}^{n}R_t^2. \end{aligned}$$

A value I_t of the irregular is considered extreme if the associated residual R_t is too large, more precisely if $|R_t| \geq \lambda\sigma$ where σ is the overall standard deviation calculated above and λ is a parameter that can be modified by the user and is set by default at 2.5.

Step 2: Final calculation of the overall standard deviation and identification of extreme values.

[17] We are dealing here with a subtraction insofar as the model used for the trading-day effect is a linear regression model.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	0.149	0.680	1.254
1986	1.070	0.556	1.511	**6.596**	0.493	0.006	0.565	2.872	0.477	1.104	0.243	0.396
1987	**5.407**	1.042	0.990	0.171	0.455	0.884	0.586	1.298	0.381	0.536	0.432	0.469
1988	0.388	0.070	0.241	0.472	1.125	0.298	1.128	0.350	0.526	2.581	1.343	1.442
1989	0.450	0.460	2.628	2.250	1.193	1.657	0.331	0.803	0.203	0.467	0.935	0.330
1990	0.029	0.640	0.312	0.238	0.491	1.248	0.272	1.318	0.438	0.791	0.170	1.525
1991	0.865	0.058	0.781	0.489	0.753	0.505	2.272	0.218	0.469	0.090	0.361	0.510
1992	0.098	0.241	0.278	0.492	0.007	1.036	0.327	0.494	0.824	0.996	0.123	1.773
1993	1.187	2.497	0.572	1.750	0.524	0.165	0.080	0.422	0.383	0.552	0.584	0.178
1994	0.629	0.079	0.955	0.230	1.358	0.496	0.223	2.145	0.041	0.844	0.430	1.084
1995	0.089	0.485	0.024	.	.	.	.	.	.	.	.	.

TABLE 4.69. C14a: Deviations from the averages in absolute value.

The above calculations are redone, excluding values identified as extreme. A new standard deviation is then obtained, as well as new extreme values to be excluded from the regression for trading-day; these values are shown in Table C14.

Comments

- The calculation carried out here is markedly different from that performed in Table B14. The use of the adjustment factors in Table B16 makes it possible, insofar as these coefficients were calculated on the basis of the daily structure of each month, to identify all the types of months and not simply the 15 categories used in Part B.

- All observations, with the exception of values identified as extreme, are used for the regression. In X-11-ARIMA, whenever an ARIMA extrapolation is requested, the daily weights are estimated for all available data up to the last month of December[18].

Example

Step 1: First calculation of an overall standard deviation and identification of extreme values.

The irregular component in Table C13 is adjusted for the trading-day factors in Table B16 to provide in absolute values the residuals in Table C14a. Thus we have, for the month of April 1986: $APR86 = |107.663 - 101.067| = 6.596$.

The mean of the squares of the elements in Table C14a provides a first estimate of the standard deviation:

$$\sigma = \left(\frac{1}{n} \sum_{t=1}^{n} R_t^2 \right)^{1/2} = 1.2302.$$

[18]In fact if we also ask for an estimation of the Easter effect, the program will use all the available data.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	107.663	.	.	.	94.806	.	.	.	.
1987	94.433	.	.	.	.	.	.	.	.	.	.	.
1988	.	.	.	.	.	.	.	.	.	95.144	.	.
1989	.	.	98.765	.	.	.	.	.	.	.	.	.
1990	.	.	.	.	.	.	.	.	.	.	.	.
1991	.	.	.	.	.	.	.	.	.	.	.	.
1992	.	.	.	.	.	.	.	.	.	.	.	.
1993	.	101.612	.	.	.	.	.	.	.	.	.	.
1994	.	.	.	.	.	.	.	.	.	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.70. C14: Extreme irregular values excluded from the trading-day regression.

The only two points that, in absolute value, are farther than 3.076 (i.e. 2.5σ) are the values for April 1986 and January 1987, and they are therefore excluded.

Step 2: Final calculation of the overall standard deviation and identification of extreme values

The new standard deviation is calculated simply by removing the residuals concerned, and we have $\sigma = 0.9439$.

This new calculation leads to a limit equal to 2.360 (i.e. 2.5σ) and therefore to the elimination of six points: APR86, AUG86, JAN87, OCT88, MAR89 and FEB93.

4.2.12 Table C15: Final Regression for Trading-Day

Description and method of calculation

Daily weights will now be estimated by means of an ordinary least squares regression carried out for data not considered extreme in Table C13, using a method identical to that used in Table B15 (Section 4.1.15).

Example

If we set a type I error of 1%, for example, the tests are interpreted as follows:

- The F-test rejects the null hypothesis of the equality of the daily weights. We may therefore accept the presence of an effect due to the daily composition of the month. In fact, the probability of finding a value for the Fisher statistic greater than that found (68.245) is practically zero and therefore smaller than our type I error. We are therefore within the critical region of the test, and we cannot accept the null hypothesis of the equality of the daily weights.

- The T-tests are interpreted similarly, though we must keep in mind that the Student law is symmetric. The Prob ($T > |t|$) value must

	Combined Weights	*A Priori* Weights	Regression Coefficients	Standard Deviation	T-value	Prob > t
Monday	1.092	1.000	0.092	0.067	1.373	0.086
Tuesday	1.242	1.000	0.242	0.066	3.649	0.000
Wednesday	1.083	1.000	0.083	0.068	1.210	0.114
Thursday	1.356	1.000	0.356	0.068	5.215	0.000
Friday	1.076	1.000	0.076	0.068	1.126	0.131
Saturday	0.518	1.000	-0.482	0.066	-7.281	0.000
Sunday	0.632	1.000	-0.368	0.067	-5.458	0.000

	Sum of Squares	Degrees of freedom	Mean Squares	F-Value	Prob > F
Regression	26.115	6	4.352	68.245	0.000
Error	6.505	106	0.064		
Total	32.620	112			

TABLE 4.71. C15: Final trading-day regression.

therefore be compared to half of the type I error, i.e. 0.005. All tests leading to a value smaller than 0.005 prevent us from accepting the null hypothesis for the daily weight. In our case, the weights for Tuesday, Thursday, Saturday and Sunday are considered significantly different from 0.

Finally, with respect to the results in Table B15, note that the precision of the estimates has increased as a result of the adjustment for extreme values.

4.2.13 Table C16: Regression-Derived Trading-Day Adjustment Factors

Description and method of calculation

From the regression estimates we derive Table C16, similar to Table B16, the monthly trading-day adjustment factors M_t.

The irregular component in Table C13 is then adjusted for calendar effects, and we get Table C16bis, which unfortunately cannot be printed using current versions of the softwares: $C16bis = C13 \; op \; C16$.

Comments

- X-11-ARIMA also includes a Table C16A which repeats the daily weights obtained by regression.

- X-12-ARIMA and X-11-ARIMA estimate the adjustment factors for the next 12 months and output them in Table C16C.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	102.198	98.646	99.895
1986	101.662	99.115	97.557	**101.084**	99.839	99.083	102.198	97.504	101.116	101.662	97.167	101.347
1987	99.839	99.115	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1988	97.504	102.982	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1989	99.895	**99.115**	101.662	97.167	101.347	101.441	97.557	102.198	98.646	99.895	101.463	97.504
1990	101.347	99.115	99.839	99.083	102.198	98.646	99.895	101.662	97.167	101.347	101.441	97.557
1991	102.198	99.115	97.504	101.116	101.662	97.167	101.347	99.839	99.083	102.198	98.646	99.895
1992	101.662	100.947	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1993	97.504	99.115	101.347	101.441	97.557	101.084	99.839	99.895	101.463	97.504	101.116	101.662
1994	97.557	99.115	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1995	99.895	99.115	101.662	.	.	.	.	.	.	.	.	.

TABLE 4.72. C16: Trading-day adjustment factors derived from regression coefficients.

Example

Consider for example the months of April 1986 and February 1989.

	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Sun.	No. of days
Weight	1.092	1.242	1.083	1.356	1.076	0.518	0.632	
Number of occurrences								
APR86	4	5	5	4	4	4	4	30
FEB89	4	4	4	4	4	4	4	28

The adjustment factors for these months are:

$$\begin{aligned} APR86 &= 100 \times \frac{4 \times 1.09246 + 5 \times 1.24237 + 5 \times 1.08276}{30} + \\ & \quad 100 \times \frac{4 \times 1.35622 + 4 \times 1.07609 + 4 \times 0.51763 + 4 \times 0.63247}{30} \\ &= 101.084 \end{aligned}$$

or more simply, since Tuesday and Wednesday are the only days which occur five times in the month:

$$APR86 = 100 \times \frac{28 + 1.2424 + 1.0828}{30} = 101.084$$

and, for one month of February having 28 days:

$$FEB89 = 100 \times \frac{28}{28.25} = 99.115.$$

We then get the adjusted values for the irregular component in Table C16bis. For example,

$$APR86 = 100 \times 107.663/101.084 = 106.509$$

and

$$FEB89 = 100 \times 99.576/99.115 = 100.465.$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	99.720	100.817	98.860
1986	100.788	99.439	98.624	**106.509**	99.508	100.010	100.419	97.234	99.592	100.822	99.969	99.557
1987	94.585	101.051	101.106	99.928	100.645	99.189	100.312	98.843	100.360	99.464	100.452	100.325
1988	99.781	100.028	100.102	99.649	101.241	100.052	99.022	99.602	100.363	97.527	98.655	101.446
1989	100.565	**100.465**	97.151	102.535	98.770	101.478	99.834	99.080	99.922	99.647	100.680	100.517
1990	99.976	99.355	100.313	99.776	100.346	98.863	100.387	101.033	99.769	100.728	100.012	98.610
1991	100.712	99.941	99.378	100.547	98.995	100.739	102.189	100.220	99.542	99.954	100.494	99.604
1992	99.833	100.406	99.837	100.244	100.172	99.039	99.415	100.680	100.799	100.999	99.892	98.131
1993	98.962	102.519	100.511	98.119	100.710	100.147	100.081	100.537	99.381	99.613	99.486	99.911
1994	100.818	99.920	98.931	99.895	101.474	100.248	100.408	102.064	99.885	99.308	99.558	101.086
1995	100.204	99.511	99.712	.	.	.	.	.	.	.	.	.

TABLE 4.73. C16bis: Irregular component adjusted for regression-derived trading-day effects.

4.2.14 Table C17: Final Weights Used to Adjust the Irregular

Description and method of calculation

On the basis of the estimate of the irregular component of Table C16bis, or of Table C13 if no adjustment is requested for trading-day, an attempt is made to identify and adjust extreme values. To this end, we use the algorithm described for Tables B4 and B9 (Sections 4.1.4 and 4.1.9 respectively) to detect extreme values and calculate adjustment weights. Since we already have an estimate of the irregular, only Steps 4 and 5 need be applied.

Comments

The comments made for Tables B4 and B9 concerning the calculation of moving standard deviations and adjustment weights are equally valid here.

Example

Calculating a moving standard deviation.

The standard deviation corresponding to the year 1989 is calculated on the basis of data for the years 1987 to 1991 (two years before, two years after) using the formula[19]:

$$\sigma_{89} = \left[\frac{1}{60} \sum_{t=Jan87}^{Dec91} (I_t - 100)^2 \right]^{1/2} = 1.4629.$$

Those for the years 1988, 1990, 1991 and 1992 are calculated using the same principle.

For X-11-ARIMA and X-12-ARIMA, the standard deviation for 1987 is calculated on the basis of all available observations from 1985 to 1990, i.e.

[19]Here the theoretical mean is assumed to be 100 as the irregular values have been multiplied by 100.

Year	Standard Deviation 1	Standard Deviation 2
	1.4389	0.9815
1986	1.4389	0.9815
1987	1.4389	0.9815
1988	1.4629	0.9889
1989	1.1712	0.9476
1990	0.9538	0.9538
1991	0.9526	0.9030
1992	0.8592	0.8021
1993	0.8420	0.7861
1994	0.8420	0.7861
1995	0.8420	0.7861

TABLE 4.74. C17a: Five-year moving standard deviations.

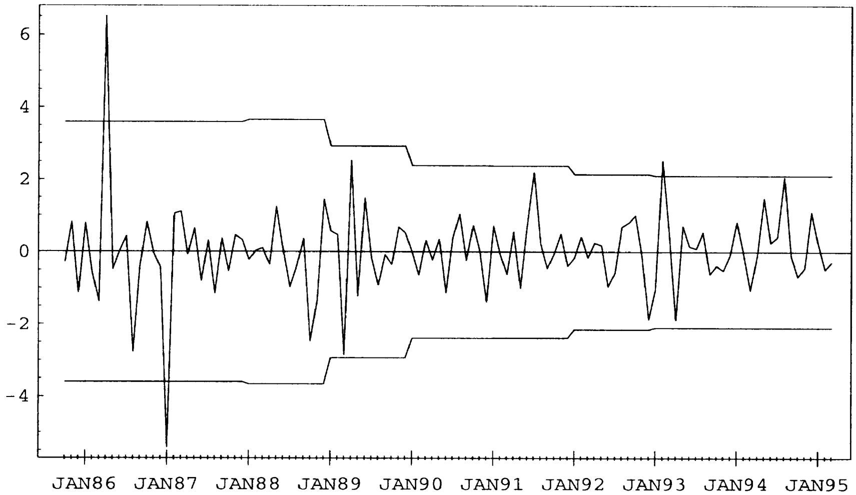


FIGURE 4.8. C17: Deviations of the irregular component from its theoretical average and upper detection thresholds ($\pm 2.5\sigma$).

63 observations. These first estimates of the standard deviation are shown in Table C17a, in the Standard Deviation 1 column.

They are used to identify the possible extreme values. As can be seen from Figure 4.8, the values for April 1986, January 1987, February 1993 and August 1994 are found to be very extreme. In fact, we have, e.g.: $|APR86 - 100| = |106.509 - 100| = 6.509 > 2.5 \times \sigma_{86} = 2.5 \times 1.4389 = 3.597$, $|JAN87 - 100| = |94.585 - 100| = 5.415 > 2.5 \times \sigma_{87} = 2.5 \times 1.4389 = 3.597$, and $|FEB93 - 100| = |102.519 - 100| = 2.519 > 2.5 \times \sigma_{93} = 2.5 \times 0.8420 = 2.105$. These points are therefore removed from the second calculation of the moving standard deviations, which leads to the results in the Standard Deviation 2 column in Table C17a.

Detecting and adjusting extreme values.

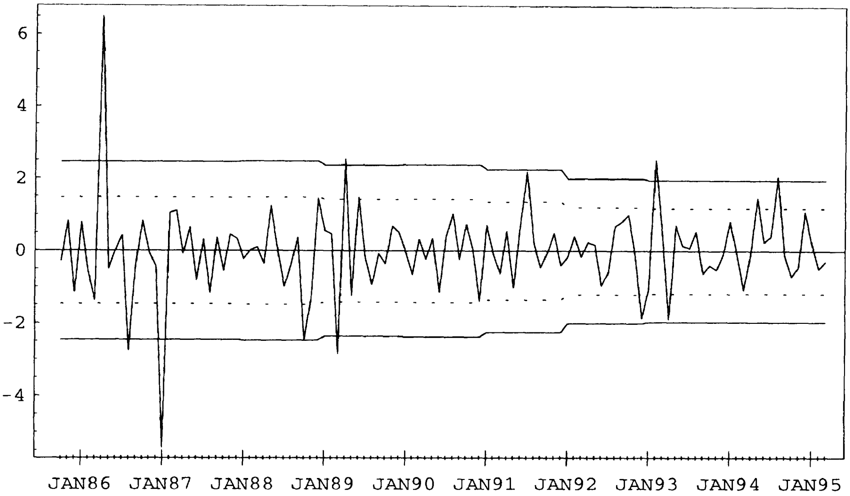


FIGURE 4.9. C17: Deviations of the irregular component from its theoretical average and detection thresholds ($\pm 1.5\sigma$ and $\pm 2.5\sigma$).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	100.000	**0.000**	100.000	100.000	100.000	0.000	100.000	100.000	100.000	100.000
1987	0.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	0.000	100.000	100.000
1989	100.000	100.000	0.000	0.000	100.000	94.034	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	100.000	100.000	100.000	100.000	**7.552**	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	16.963
1993	100.000	0.000	100.000	10.773	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	100.000	100.000	**62.449**	100.000	100.000	0.000	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

TABLE 4.75. C17: Final weights for the irregular component. Limits used are for 1.5σ to 2.5σ.

Values of the irregular are then classified in relation to the upper and lower detection thresholds calculated on the basis of the new estimates of standard deviations. All values beyond the lower detection thresholds (see Figure 4.9) are considered extreme, and are therefore adjusted in varying degrees. The weights (multiplied by 100) associated with each of these values are shown in Table C17.

The values previously considered very extreme remain so, and are assigned zero weight. Thus, $|APR86 - 100| = |106.509 - 100| = 6.509 > 2.5 \times \sigma_{86} = 2.5 \times 0.9815 = 2.454$.

For July 1991, located between the two detection thresholds and therefore considered moderately extreme, we have: $|JUL91 - 100| = |102.189 - 100| = 2.189$, and $1.5 \times \sigma_{91} = 1.5 \times 0.9030 = 1.3545 < 2.189 < 2.5 \times \sigma_{91} = 2.5 \times 0.9030 = 2.2575$. This value, considered moderately extreme, is assigned a

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	102.198	98.646	99.895
1986	101.662	99.115	97.557	**101.084**	99.839	99.083	102.198	97.504	101.116	101.662	97.167	101.347
1987	99.839	99.115	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1988	97.504	102.982	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1989	99.895	99.115	101.662	97.167	101.347	101.441	97.557	102.198	98.646	99.895	101.463	97.504
1990	101.347	99.115	99.839	99.083	102.198	98.646	99.895	101.662	97.167	101.347	101.441	97.557
1991	102.198	99.115	97.504	101.116	101.662	97.167	101.347	99.839	99.083	102.198	98.646	99.895
1992	101.662	100.947	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1993	97.504	99.115	101.347	101.441	97.557	101.084	99.839	99.895	101.463	97.504	101.116	101.662
1994	97.557	99.115	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1995	99.895	99.115	101.662	.	.	.	.	.	.	.	.	.

TABLE 4.76. C18: Combined adjustment factors for trading-day effects.

weight proportional to the deviation from its average, i.e.:

$$weight(JUL91) = \frac{2.2575 - 2.189}{2.2575 - 1.3545} = 0.075.$$

4.2.15 Table C18: Combined Trading-Day Factors (derived from the a priori adjustment and the regression for trading-days)

Description and method of calculation

As in Table B18, if prior daily weights are provided for trading-day effects (multiplicative model only), and if a trading-day regression is also requested, Table C18 shows the combined result for both adjustments, a simple addition of the two effects, where the combined weights are obtained from Table C15.

Comment

In X-12-ARIMA, if holiday factors (Easter, Labor Day, Thanksgiving, etc.) are also estimated in the irregular component regression, they are included in Table C18.

Example

In our case, Table C18 is identical to Table C16.

4.2.16 Table C19: Raw Series Corrected for Trading-Day Effects

Description and method of calculation

The series in Table B1, or in Table A1 if no prior adjustment is requested, is adjusted for the trading-day effects estimated in Table C18. We thus have: $C19 = B1 \ op \ C18$.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	113.212	111.307	100.706
1986	104.858	99.581	106.502	**108.326**	97.858	104.660	97.556	67.382	104.039	115.186	111.458	103.012
1987	100.663	104.121	113.019	105.555	102.560	107.105	100.136	70.421	107.535	117.089	115.761	107.634
1988	110.457	107.009	116.147	109.584	107.513	113.046	103.791	74.990	112.972	120.853	119.999	114.886
1989	118.024	113.202	118.236	118.044	109.031	118.591	108.245	77.692	115.768	126.834	124.971	115.585
1990	119.490	113.504	123.800	117.174	113.114	118.404	111.918	81.939	117.941	130.246	125.294	113.575
1991	120.648	113.807	122.354	118.082	111.448	120.103	113.767	81.732	117.477	129.553	126.513	115.922
1992	121.482	115.803	124.131	118.269	112.611	117.390	110.268	82.004	118.021	129.209	123.230	111.353
1993	116.611	114.110	121.069	112.578	110.602	115.845	108.275	79.784	113.144	124.098	120.357	112.924
1994	119.213	112.496	121.333	116.984	114.120	119.255	112.303	84.265	118.887	129.566	126.331	120.194
1995	124.231	117.338	128.072	.	.	.	.	.	.	.	.	.

TABLE 4.77. C19: Original series adjusted for trading-day effects.

Example

For example, we have:

$$APR86 = 100 \times 109.500/101.084 = 108.326.$$

4.2.17 Table C20: Adjustment Values for Extreme Values of the Irregular

Description and method of calculation

Values of the irregular component in Table C16bis, or C13 if no regression is requested for trading-day, identified as extreme when Table C17 was set up, and for which a weight was therefore calculated, are adjusted as in Table B20 as follows:

$$C20 = C16bis \ op \ [\ xbar \ + C17 \times (C16bis - \ xbar \ )] \, .$$

Example

The value for May 1994, identified as extreme and assigned a weight equal to 0.62449, will therefore be adjusted as follows, since we have here a multiplicative model:

$$MAY94 = 100 \times \frac{1.01474}{1 + 0.62449 \times (1.01474 - 1)} = 100.549.$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.000	100.000	100.000
1986	100.000	100.000	100.000	106.509	100.000	100.000	100.000	97.234	100.000	100.000	100.000	100.000
1987	94.585	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1988	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	97.527	100.000	100.000
1989	100.000	100.000	97.151	102.535	100.000	100.087	100.000	100.000	100.000	100.000	100.000	100.000
1990	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1991	100.000	100.000	100.000	100.000	100.000	100.000	102.021	100.000	100.000	100.000	100.000	100.000
1992	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	98.443
1993	100.000	102.519	100.000	98.319	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
1994	100.000	100.000	100.000	100.000	**100.549**	100.000	100.000	102.064	100.000	100.000	100.000	100.000
1995	100.000	100.000	100.000	.	.	.	.	.	.	.	.	.

TABLE 4.78. C20: Weights for extreme values.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	113.212	111.307	100.706
1986	104.858	99.581	106.502	**101.706**	97.858	104.660	97.556	69.299	104.039	115.186	111.458	103.012
1987	106.425	104.121	113.019	105.555	102.560	107.105	100.136	70.421	107.535	117.089	115.761	107.634
1988	110.457	107.009	116.147	109.584	107.513	113.046	103.791	74.990	112.972	123.917	119.999	114.886
1989	118.024	113.202	121.703	115.126	109.031	118.488	108.245	77.692	115.768	126.834	124.971	115.585
1990	119.490	113.504	123.800	117.174	113.114	118.404	111.918	81.939	117.941	130.246	125.294	113.575
1991	120.648	113.807	122.354	118.082	111.448	120.103	111.514	81.732	117.477	129.553	126.513	115.922
1992	121.482	115.803	124.131	118.269	112.611	117.390	110.268	82.004	118.021	129.209	123.230	113.114
1993	116.611	111.306	121.069	114.503	110.602	115.845	108.275	79.784	113.144	124.098	120.357	112.924
1994	119.213	112.496	121.333	116.984	113.498	119.255	112.303	82.561	118.887	129.566	126.331	120.194
1995	124.231	117.338	128.072	.	.	.	.	.	.	.	.	.

TABLE 4.79. D1: Trading-day adjusted series modified by final weights.

4.3 PART D: Final Estimation of the Different Components

4.3.1 Table D1: Raw Series Corrected for A Priori Adjustments, Trading-Day Adjustments and Extreme Values

Description and method of calculation

This table shows the raw series adjusted for various effects detected in Part C: for points considered extreme and for trading-day effects, and adjusted *a priori* using elements from Part A. It is therefore calculated on the basis of Table C19, which takes into consideration the effects due to trading-days, or of Table B1 if no regression for trading-day is requested, and of Table C20, which indicates adjustments to be made for points considered extreme. We therefore have: $D1 = C19 \; op \; C20$.

Example

For example, the value for the month of April 1986, considered extreme, becomes

$$APR86 = 100 \times 108.326/106.509 = 101.706.$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**101.023**	101.111	101.213	101.375	101.629	102.090	102.522	102.878	103.176
1987	103.385	103.540	103.732	103.957	104.216	104.588	104.948	105.236	105.487	105.785	106.160	106.614
1988	107.013	107.356	107.773	108.284	108.745	109.224	109.841	110.415	110.904	111.366	111.661	111.951
1989	112.363	112.661	112.890	113.128	113.457	113.693	113.784	113.857	113.957	114.130	114.385	114.552
1990	114.702	115.032	115.299	115.532	115.687	115.617	115.582	115.642	115.595	115.572	115.541	115.542
1991	115.596	115.571	115.543	115.495	115.516	115.665	115.798	115.915	116.073	116.154	116.211	116.146
1992	115.981	115.941	115.975	115.983	115.832	115.578	115.258	114.868	114.553	114.268	114.028	113.879
1993	113.732	113.557	113.261	112.845	112.512	112.384	112.485	112.643	112.703	112.818	113.042	113.305
1994	113.615	113.898	114.253	114.720	115.197	115.749	116.261	116.672	117.154	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.80. D2: Trend-cycle, centered 12-term moving average.

4.3.2 Table D2: Preliminary Estimation of the Trend-Cycle

Description and method of calculation

A new estimate of the trend-cycle component is obtained, as in Tables B2 and C2, by applying a centered moving average of order 12 to the data in Table D1.

Comments

- X-11-ARIMA and X-12-ARIMA also include a centered 24-term moving average due to Cholette [13].

- The first six and last six points in the series are not imputed at this stage of the calculations.

Example

For example, the value for April 1986 is obtained from the values in Table D1 between October 1985 and October 1986 (six months before and six months after):

$$\begin{aligned} APR86 &= \frac{113.212}{24} + \\ &\quad \frac{111.307 + 100.706 + 104.858 + 99.581 + 106.502 + 101.706}{12} + \\ &\quad \frac{97.858 + 104.660 + 97.556 + 69.299 + 104.039}{12} + \\ &\quad \frac{115.186}{24} \\ &= 101.023. \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**100.677**	96.783	103.405	96.233	68.188	101.909	112.353	108.339	99.841
1987	102.940	100.562	108.953	101.538	98.411	102.407	95.415	66.917	101.941	110.686	109.045	100.957
1988	103.218	99.677	107.770	101.201	98.867	103.499	94.492	67.917	101.865	111.269	107.468	102.622
1989	105.039	100.480	107.806	101.766	96.099	104.217	95.132	68.237	101.589	111.131	109.255	100.902
1990	104.175	98.672	107.373	101.422	97.776	102.410	96.830	70.855	102.030	112.696	108.442	98.297
1991	104.371	98.474	105.895	102.240	96.478	103.836	96.301	70.510	101.210	111.535	108.865	99.807
1992	104.743	99.881	107.033	101.971	97.219	101.568	95.671	71.390	103.028	113.075	108.070	99.328
1993	102.531	98.018	106.894	101.470	98.303	103.079	96.257	70.829	100.391	109.998	106.471	99.664
1994	104.927	98.769	106.197	101.974	98.525	103.029	96.596	70.764	101.479	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.81. D4: Modified seasonal-irregular ratios.

4.3.3 Table D4: Preliminary Estimation of the Modified Seasonal-Irregular Component

Description and method of calculation

The trend-cycle component is removed from the analyzed series, and an estimate of the seasonal-irregular component is obtained. Thus we have: $D4 = D1$ *op* $D2$.

Comment

Again, there is no estimate for the six values at the beginning and six values at the end of the series.

Example

The value for April 1986 is therefore obtained simply as follows:

$$APR86 = 100 \times 101.706/101.023 = 100.677.$$

4.3.4 Table D5: Estimation of the Seasonal Component

Description and method of calculation

The estimation is done on the basis of the values for the seasonal-irregular component in Table D4. Again, the process includes three steps, as in Tables B5 and C5:

- Step 1: Estimating the seasonal component using a 3x3 moving average.
- Step 2: Normalizing the seasonal factors using a centered 12-term moving average.
- Step 3: Imputing the missing seasonal factors.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	101.124	97.832	103.016	95.577	67.620	101.914	111.473	108.465	100.811
1987	103.442	100.186	108.259	**101.252**	97.850	103.150	95.356	67.652	101.874	111.319	108.476	101.093
1988	103.724	100.003	108.042	**101.367**	97.798	103.284	95.292	68.122	101.844	111.388	108.420	101.080
1989	104.135	99.686	107.617	101.591	97.386	103.413	95.529	68.853	101.745	111.505	108.610	100.590
1990	104.367	99.276	107.147	101.717	97.174	103.156	95.947	69.929	101.842	111.974	108.567	99.807
1991	104.280	99.003	106.800	101.860	97.092	102.973	96.144	70.565	101.747	111.920	108.372	99.471
1992	104.126	98.897	106.694	101.859	97.502	102.664	96.173	70.941	101.755	111.836	107.903	99.425
1993	103.930	98.746	106.638	101.816	97.877	102.758	96.198	70.922	101.448	111.536	107.566	99.553
1994	103.917	98.669	106.636	101.768	98.193	102.779	96.287	70.906	101.323	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.82. D5a: Preliminary seasonal factors (3×3 ma).

Comment

We can specify the moving average to be used. In such a case, X-11-ARIMA gives the choice between a simple 3-term, a 3×3, a 3×5, a 3×9 and constant seasonality (a simple average). X-12-ARIMA also includes a 3×15.

Example

The estimation is based on the modified seasonal-irregular component in Table D4.

Step 1: Estimating the seasonal component.

The data in Table D4 are smoothed column by column (month by month), using a 3×3 moving average, and Table D5a is obtained.

The seasonal factor for the month of April 1988 is therefore estimated as follows:

$$
\begin{aligned}
APR88 &= \frac{100.677 + 2 \times 101.538 + 3 \times 101.201 + 2 \times 101.766 + 101.422}{9} \\
&= 101.367.
\end{aligned}
$$

This symmetric moving average can be used to estimate the seasonal factors for the years 1988 to 1992. For the beginning of the series (years 1986 and 1987) and the end of the series (years 1993 and 1994), predefined asymmetric averages are used (see Table 3.11). Thus, for April 1987 for example, one point in the past, the current point and two points in the future are used:

$$
\begin{aligned}
APR87 &= \frac{100.677 \times 7 + 101.538 \times 10 + 101.201 \times 7 + 101.766 \times 3}{27} \\
&= 101.252.
\end{aligned}
$$

Step 2: Normalizing the seasonal factors.

On Table D5a, a centered 12-month moving average is used to produce Table D5b. The first computable term is therefore that for October 1986,

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	99.982	99.982	99.982	99.982	99.982	99.982	**99.982**	99.988	99.994
1987	99.991	99.983	99.983	99.974	99.968	99.981	100.004	100.008	99.992	99.987	99.990	99.994
1988	99.996	100.013	100.032	100.033	100.034	100.031	100.048	100.051	100.021	100.012	100.004	99.993
1989	100.008	100.048	100.074	100.075	100.088	100.075	100.065	100.057	100.021	100.006	100.003	99.983
1990	99.990	100.052	100.101	**100.125**	100.142	100.108	100.072	100.057	100.031	100.022	100.025	100.014
1991	100.014	100.049	100.072	100.065	100.055	100.033	100.012	100.002	99.993	99.988	100.005	100.010
1992	99.998	100.015	100.031	100.028	100.005	99.983	99.973	99.959	99.950	99.946	99.960	99.979
1993	99.984	99.984	99.971	99.946	99.919	99.910	99.915	99.911	99.908	99.906	99.917	99.931
1994	99.936	99.939	**99.933**	99.933	99.933	99.933	99.933	99.933	99.933	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.83. D5b: Centered 12-month moving average.

and the last that for March 1994. Thus:

$$
\begin{aligned}
APR90 &= \frac{111.505}{24} + \\
&\quad \frac{108.610 + 100.590 + 104.367 + 99.276 + 107.147 + 101.717}{12} + \\
&\quad \frac{97.174 + 103.156 + 95.947 + 69.929 + 101.842}{12} + \\
&\quad \frac{111.974}{24} \\
&= 100.125.
\end{aligned}
$$

The first six values, from April to September 1986, which cannot be computed using this symmetric moving average, are taken to be equal to the first computable value, that for October 1986 (99.982). The same applies to the end of the series: the value calculated for March 1994 (99.933) is repeated for the next six months. Normalized seasonal factors are then obtained by dividing Table D5a by Table D5b, and we get Table D5. For example,

$$APR86 = 100 \times 101.124/99.982 = 101.143.$$

Step 3: Imputing the missing seasonal factors.

To obtain the seasonal factors for the six missing values at either end of the series (October 1985 to March 1986 and October 1994 to March 1995) caused by the use of the centered 12-term moving average in Table D2, we repeat the nearest available factor for that particular month.

4.3.5 Table D6: Estimation of the Seasonally Adjusted Series

Description and method of calculation

The estimation is done simply by removing from the series in Table D1 the estimate of the seasonal component in Table D5: $D6 = D1$ *op* $D5$.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.493	108.478	100.817
1986	103.452	100.203	108.278	**101.143**	97.850	103.034	95.594	67.632	101.932	111.493	108.478	100.817
1987	103.452	100.203	108.278	101.278	97.881	103.170	95.352	67.646	101.882	111.333	108.487	101.100
1988	103.728	99.989	108.008	101.334	97.765	103.252	95.247	68.087	101.823	111.374	108.416	101.088
1989	104.127	99.638	107.537	101.515	97.300	103.335	95.467	68.814	101.724	111.498	108.608	100.607
1990	104.378	99.224	107.039	101.590	97.036	103.045	95.878	69.890	101.811	111.949	108.540	99.793
1991	104.265	98.955	106.723	101.794	97.038	102.940	96.132	70.564	101.755	111.933	108.366	99.461
1992	104.128	98.882	106.661	101.831	97.498	102.681	96.199	70.970	101.806	111.896	107.947	99.445
1993	103.947	98.762	106.669	101.871	97.956	102.851	96.279	70.985	101.541	111.641	107.655	99.622
1994	103.984	98.729	106.707	101.836	98.259	102.848	96.351	70.954	101.391	111.641	107.655	99.622
1995	103.984	98.729	106.707	.	.	.	.	.	.	.	.	.

TABLE 4.84. D5: Seasonal factors.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	101.542	102.608	99.890
1986	101.359	99.379	98.360	**100.557**	100.008	101.577	102.052	102.465	102.067	103.312	102.746	102.178
1987	102.874	103.910	104.379	104.223	104.780	103.814	105.017	104.101	105.548	105.171	106.705	106.463
1988	106.487	107.020	107.536	108.142	109.971	109.485	108.970	110.138	110.950	111.262	110.685	113.649
1989	113.347	113.613	113.173	113.408	112.056	114.664	113.385	112.902	113.806	113.754	115.067	114.888
1990	114.479	114.392	115.659	115.340	116.569	114.905	116.730	117.240	115.843	116.344	115.436	113.811
1991	115.714	115.010	114.646	116.001	114.850	116.673	116.001	115.827	115.451	115.742	116.746	116.550
1992	116.665	117.112	116.379	116.143	115.501	114.325	114.625	115.547	115.927	115.472	114.158	113.745
1993	112.183	112.702	113.499	112.400	112.910	112.634	112.459	112.396	111.427	111.158	111.798	113.352
1994	114.646	113.943	113.707	114.875	115.509	115.953	116.556	116.359	117.256	116.055	117.348	120.650
1995	119.472	118.848	120.022	.	.	.	.	.	.	.	.	.

TABLE 4.85. D6: Seasonally adjusted series.

Example

For example,

$$APR86 = 100 \times 101.706/101.143 = 100.557.$$

4.3.6 Table D7: Estimation of the Trend-Cycle Component

Description and method of calculation

This table shows an estimate of the trend-cycle component obtained from the seasonally adjusted series in Table D6. The method used is the same as that followed to produce Table C7 (Section 4.2.6).

Comment

It is possible to specify the length of the Henderson moving average to be used. X-11-ARIMA provides a choice between a 9-term, a 13-term or a 23-term moving average. X-12-ARIMA allows any odd-numbered average less than 101.

Example

Step 1: Selecting the moving average, calculating the $\bar{I}/\bar{C}$ ratio.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	99.924	100.448	101.139	101.793	102.277	102.510	102.619	102.727	102.919
1987	103.227	103.605	103.961	104.237	104.392	104.458	104.568	104.811	105.176	105.595	105.986	106.358
1988	106.769	107.255	107.783	108.342	108.897	109.373	109.753	110.132	110.620	111.245	111.976	112.620
1989	113.055	113.305	113.399	113.377	113.299	113.273	113.379	113.577	113.823	114.082	114.322	114.577
1990	114.827	115.007	115.221	**115.507**	115.844	116.190	116.387	116.378	116.190	115.852	115.462	115.118
1991	114.922	114.960	115.161	115.427	115.657	115.784	115.845	115.893	115.954	116.098	116.330	116.583
1992	116.752	116.675	116.343	115.875	115.469	115.236	115.197	115.241	115.170	114.874	114.359	113.735
1993	113.185	112.848	112.740	112.756	112.737	112.564	112.240	111.920	111.807	111.956	112.336	112.886
1994	113.478	114.036	114.508	114.946	115.391	115.783	116.130	116.484	116.891	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.86. D7a: Trend-cycle (13-term Henderson ma).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.633	99.562	100.433	100.254	100.183	99.568	100.676	100.018	99.280
1987	99.659	100.295	100.402	99.987	100.372	99.383	100.430	99.323	100.354	99.598	100.678	100.099
1988	99.737	99.781	99.771	99.816	100.986	100.102	99.286	100.005	100.298	100.015	98.847	100.914
1989	100.258	100.272	99.800	100.027	98.903	101.228	100.005	99.406	99.985	99.713	100.652	100.271
1990	99.697	99.466	100.380	99.855	100.626	98.894	100.294	100.740	99.702	100.425	99.978	98.865
1991	100.689	100.043	99.553	100.497	99.301	100.767	100.135	99.943	99.567	99.693	100.358	99.972
1992	99.926	100.375	100.031	100.231	100.028	99.209	99.503	100.265	100.658	100.520	99.825	100.008
1993	99.115	99.871	100.673	99.684	100.153	100.062	100.195	100.426	99.660	99.286	99.521	100.413
1994	101.029	99.919	99.300	99.939	100.102	100.147	100.367	99.893	100.313	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.87. D7b: Irregular component.

First, Table D6 is smoothed using a 13-term Henderson moving average whose coefficients are shown in Table 3.8. Thus, we have for the month of April 1990:

$$
\begin{aligned}
APR90 &= 113.754 \times (-0.01935) + 115.067 \times (-0.02786) + \\
&\quad 114.888 \times (0.00000) + 114.479 \times (0.06549) + \\
&\quad 114.392 \times (0.14736) + 115.659 \times (0.21434) + \\
&\quad 115.340 \times (0.24006) + 116.569 \times (0.21434) + \\
&\quad 114.905 \times (0.14736) + 116.730 \times (0.06549) + \\
&\quad 117.240 \times (0.00000) + 115.843 \times (-0.02786) + \\
&\quad 116.344 \times (-0.01935) \\
&= 115.507.
\end{aligned}
$$

At this step in the calculation, there is no attempt to estimate the six points that cannot be calculated at the beginning and end of the series. An estimate is derived for the trend-cycle (Table D7a), and one for the irregular component (Table D7b) by division using Table D6.

As this is a multiplicative model, the mean growth rates are calculated (see Section 4.1.7).

Using the row totals from Tables D7c and D7d, we have:

$$
\bar{C} = \frac{2.961 + 3.291 + 5.735 + 1.948 + 2.664}{101} + \frac{1.606 + 2.835 + 2.697 + 3.494}{101}
$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1985	.	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	0.524	0.688	0.646	0.476	0.227	0.106	0.106	0.187	2.961
1987	0.299	0.367	0.343	0.265	0.149	0.064	0.105	0.232	0.348	0.399	0.370	0.350	3.291
1988	0.386	0.456	0.492	0.518	0.513	0.437	0.347	0.345	0.443	0.566	0.657	0.575	5.735
1989	0.387	0.221	0.083	0.019	0.069	0.023	0.093	0.175	0.217	0.227	0.210	0.224	1.948
1990	0.218	0.156	0.186	0.249	0.291	0.299	0.169	0.008	0.161	0.291	0.337	0.298	2.664
1991	0.170	0.033	0.175	0.231	0.199	0.110	0.053	0.041	0.053	0.124	0.199	0.217	1.606
1992	0.146	0.066	0.285	0.401	0.351	0.201	0.034	0.038	0.062	0.257	0.449	0.545	2.835
1993	0.484	0.298	0.095	0.014	0.017	0.154	0.288	0.285	0.100	0.133	0.339	0.489	2.697
1994	0.525	0.491	0.414	0.382	0.388	0.340	0.299	0.305	0.350	.	.	.3.494	
1995	.	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.88. D7c: Absolute growth rate (in %) of the trend-cycle.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1985	.	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	1.065	0.875	0.178	0.071	0.615	1.113	0.653	0.739	5.308
1987	0.382	0.638	0.107	0.414	0.385	0.985	1.053	1.102	1.038	0.754	1.085	0.575	8.518
1988	0.362	0.044	0.010	0.045	1.173	0.876	0.815	0.724	0.293	0.283	1.168	2.092	7.884
1989	0.650	0.014	0.471	0.227	1.124	2.351	1.208	0.599	0.582	0.272	0.942	0.378	8.817
1990	0.573	0.232	0.920	0.523	0.772	1.721	1.416	0.445	1.031	0.725	0.445	1.113	9.916
1991	1.845	0.641	0.490	0.948	1.190	1.476	0.628	0.191	0.377	0.127	0.668	0.385	8.965
1992	0.047	0.450	0.343	0.200	0.203	0.819	0.297	0.766	0.392	0.137	0.692	0.184	4.527
1993	0.893	0.762	0.804	0.983	0.471	0.091	0.133	0.230	0.762	0.375	0.236	0.896	6.637
1994	0.613	1.099	0.619	0.643	0.164	0.044	0.220	0.473	0.420	.	.	.4.295	
1995	.	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.89. D7d: Absolute growth rate (in %) of the irregular.

$$
\begin{aligned}
&= 0.2696, \\
\bar{I} &= \frac{5.308 + 8.518 + 7.884 + 8.817 + 9.916}{101} + \\
&\quad \frac{8.965 + 4.527 + 6.637 + 4.295}{101} \\
&= 0.64224, \\
\bar{I}/\bar{C} &= \frac{0.64224}{0.2696} \\
&= 2.3822.
\end{aligned}
$$

Step 2: Smoothing the SA series using a Henderson moving average.

Since the ratio is greater than 1 and smaller than 3.5, we select a 13-term Henderson moving average whose coefficients and those of the associated asymmetric moving averages are shown in Table 3.8. The trend-cycle for October 1985 is estimated, from the seasonally adjusted series in Table D6, using the current point and six points in the future, to which are assigned the coefficients of the $H6\_0$ moving average in Table 3.8.

$$
\begin{aligned}
OCT85 &= 101.542 \times (0.42113) + 102.608 \times (0.35315) + \\
&\quad 99.890 \times (0.24390) + 101.359 \times (0.11977) + \\
&\quad 99.379 \times (0.01202) + 98.360 \times (-0.05811) + \\
&\quad 100.557 \times (-0.09186) \\
&= 101.743.
\end{aligned}
$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	**101.743**	101.327	100.829
1986	100.322	99.896	99.730	99.924	100.448	101.139	101.793	102.277	102.510	102.619	102.727	102.919
1987	103.227	103.605	103.961	104.237	104.392	104.458	104.568	104.811	105.176	105.595	105.986	106.358
1988	106.769	107.255	107.783	108.342	108.897	109.373	109.753	110.132	110.620	111.245	111.976	112.620
1989	113.055	113.305	113.399	113.377	113.299	113.273	113.379	113.577	113.823	114.082	114.322	114.577
1990	114.827	115.007	115.221	115.507	115.844	116.190	116.387	116.378	116.190	115.852	115.462	115.118
1991	114.922	114.960	115.161	115.427	115.657	115.784	115.845	115.893	115.954	116.098	116.330	116.583
1992	116.752	116.675	116.343	115.875	115.469	115.236	115.197	115.241	115.170	114.874	114.359	113.735
1993	113.185	112.848	112.740	112.756	112.737	112.564	112.240	111.920	111.807	111.956	112.336	112.886
1994	113.478	114.036	114.508	114.946	115.391	115.783	116.130	116.484	116.891	117.414	118.042	118.680
1995	119.243	119.691	120.001	.	.	.	.	.	.	.	.	.

TABLE 4.90. D7: Trend-cycle, I/C ratio is 2.382, a 13-term Henderson moving average has been selected.

This leads to Table D7.

4.3.7 Table D8: Estimation of the Unmodified Seasonal-Irregular Component

Description and method of calculation

The trend-cycle component is removed from the analysed series in Table C19, or B1 if the trading-day adjustment option has not been selected, and we obtain an estimate of the seasonal-irregular component. Thus we have: $D8 = C19 \; op \; D7$. Note that this series includes the extreme values.

Several tests, both parametric and non-parametric, are carried out for the presence of seasonality. These tests are further reported in Table F2I.

1. Two **stable seasonality** tests are performed, one parametric and the other non-parametric, as a means of choosing between the two hypotheses:

$$
\begin{aligned}
H_0 &: \; m_1 = m_2 = \ldots = m_k \\
H_1 &: \; m_p \neq m_q \text{ for at least one pair } (p, q)
\end{aligned}
$$

   where $m_1, \ldots, m_k$ are the stable seasonal factors for 12 months or 4 quarters.

   - The first, called the **stable seasonality test**, is a parametric test based on a one-way analysis of variance model similar to that following Table B1 (see Section 4.1.3).
   - The second is a non-parametric Kruskal-Wallis test, for which the data in Table D8 are assumed to be the result of k independent samples $A_1, A_2, \ldots, A_k$ ($k = 4$ if the series is quarterly, $k = 12$ if it is monthly) of sizes $n_1, n_2, \ldots, n_k$ respectively. The test is based on the statistic

$$
W = \frac{12}{n(n+1)} \sum_{j=1}^{k} \frac{S_j^2}{n_j} - 3(n+1)
$$

where S_j is the sum of the ranks of the observations from sample A_j within the whole sample of $n = \sum_{j=1}^{k} n_j$ observations. This quantity, assuming the null hypothesis, follows a chi-square distribution with $k-1$ degrees of freedom.

2. Next we have a so-called **moving seasonality test** test which is based on a two-way analysis of variance model (the month or the quarter, and the year) proposed by Higginson [34]. It uses the following model for the values of the seasonal-irregular component, for **complete years** only, in Table D8:

$$|SI_{ij} - \ xbar \ | = X_{ij} = b_i + m_j + e_{ij}$$

where

- m_j refers to the monthly or quarterly effect j ($j = 1, \ldots, k$);
- b_i refers to the annual effect i ($i = 1, \ldots, N$) where N is the number of complete years;
- e_{ij} represents the residual effect, a collection of independent and identically distributed (*i.i.d.*) random variables, with a zero average.

The test is based on the decomposition $S^2 = S_A^2 + S_B^2 + S_R^2$ where:

- S^2 is the total sum of squares: $S^2 = \sum_{j=1}^{k} \sum_{i=1}^{N} \left(X_{ij} - \bar{X}_{..}\right)^2$ with $\bar{X}_{..} = \sum_{j=1}^{k} \sum_{i=1}^{N} X_{ij}/(kN)$;
- S_A^2 is the "inter-month" sum of squares: $S_A^2 = N \sum_{j=1}^{k} \left(\bar{X}_{.j} - \bar{X}_{..}\right)^2$ with $\bar{X}_{.j} = \sum_{i=1}^{N} X_{ij}/N$;
- S_B^2 is the "inter-year" sum of squares: $S_B^2 = k \sum_{i=1}^{N} \left(\bar{X}_{i.} - \bar{X}_{..}\right)^2$ with $\bar{X}_{i.} = \sum_{j=1}^{k} X_{ij}/k$;
- S_R^2 is the "residual" sum of squares: $S_R^2 = \sum_{i=1}^{N} \sum_{j=1}^{k} \left(X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..}\right)^2$.

The null hypothesis $H_0^* : b_1 = b_2 = \ldots = b_N$, i.e. no change in seasonality over the years, can be tested using the statistic

$$F_M = \frac{S_B^2/(N-1)}{S_R^2/(N-1)(k-1)}$$

which follows a F-distribution with $(N-1)$ and $(k-1)(N-1)$ degrees of freedom under H_0^*.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.272	109.850	99.878
1986	104.521	99.685	106.791	**108.408**	97.421	103.481	95.838	65.882	101.491	112.247	108.498	100.090
1987	97.516	100.498	108.713	101.265	98.245	102.534	95.762	67.188	102.243	110.885	109.223	101.200
1988	103.455	99.770	107.760	101.147	98.729	103.358	94.567	68.091	102.126	108.636	107.165	102.012
1989	104.396	99.909	104.265	104.116	96.233	104.695	95.472	68.405	101.709	111.178	109.316	100.880
1990	104.061	98.694	107.446	101.443	97.643	101.905	96.160	70.407	101.507	112.424	108.516	98.660
1991	104.983	98.997	106.246	102.300	96.361	103.730	98.207	70.524	101.314	111.589	108.754	99.434
1992	104.051	99.253	106.694	102.066	97.525	101.869	95.721	71.158	102.476	112.478	107.757	97.905
1993	103.027	101.119	107.388	99.842	98.106	102.914	96.468	71.287	101.196	110.845	107.140	100.034
1994	105.054	98.649	105.960	101.774	98.899	102.998	96.705	72.340	101.708	110.349	107.022	101.276
1995	104.183	98.034	106.726	.	.	.	.	.	.	.	.	.

TABLE 4.91. D8: Final unmodified seasonal-irregular ratios.

3. Finally, these tests are completed with a test for the **presence of identifiable seasonality**, designed by combining the F-statistic values of parametric tests for the *stable seasonality* (statistic F_S) and for the *moving seasonality* (statistic F_M) mentioned above. This test was developed, on the grounds of theoretical and practical considerations, by Lothian and Morry [48]. The value of this test statistic T is:

$$T = \left(\frac{T_1 + T_2}{2}\right)^{1/2} \quad \text{here } T_1 = \frac{7}{F_S} \quad \text{and} \quad T_2 = \frac{3F_M}{F_S}$$

The final version of the combined test can then be interpreted using the diagram in Figure 4.10.

Example

The value for April 1986 is therefore obtained simply:

$$APR86 = 100 \times 108.326/99.924 = 108.408.$$

This leads to Table D8, whose values are also shown as solid triangles in Figure 4.11 in Section 4.3.10.

The computations leading to the values of different test statistics need not be described in detail here. Suffice it to explain the combined test:

1. Statistic F_S in Table 4.92 is sufficiently large for the equality of the monthly seasonal factors hypothesis to be rejected.

2. Statistic F_M in Table 4.94 shows that the series has no moving seasonality; the test accepts hypothesis H_0^*.

3. We have $T_1 = 7/F_S = 7/498.194 = 0.014$ and $T_2 = 3F_M/F_S = 3 \times 1.724/498.194 = 0.010$ and therefore these two statistics are smaller than 1.

4. The non-parametric Kruskal-Wallis test in Table 4.93 confirms the presence of seasonality.

The combined test therefore determines the presence of identifiable seasonality.

	SS	DF	RMSE	F	PROB>F
Inter-month	11264.919	11	1024.084	498.194	0.000
Residual	209.670	102	2.056		
Total	11474.589	113			

TABLE 4.92. Test for the presence of stable seasonality.

Kruskal-Wallis Statistic	DF	PROB>W
104.780	11	0.000

TABLE 4.93. Non-parametric test for the presence of stable seasonality.

	SS	DF	RMSE	F	PROB>F
Inter-year	20.628	8	2.578	1.724	0.104
Residual	131.614	88	1.496		

TABLE 4.94. Test for the presence of moving seasonality.

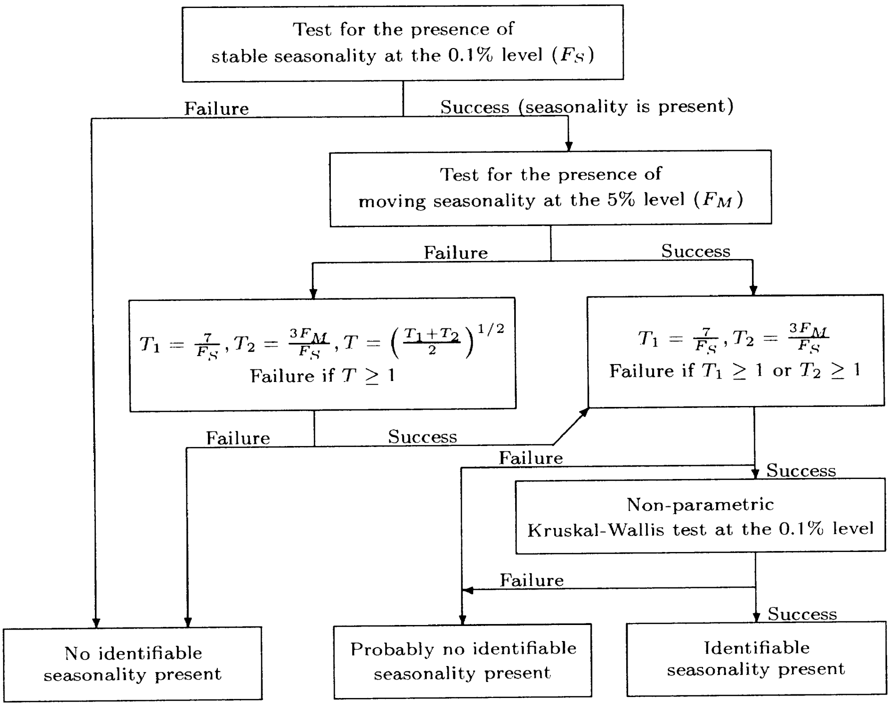


FIGURE 4.10. Test for the presence of identifiable seasonality.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.272	109.850	99.878
1986	104.521	99.685	106.791	**101.783**	97.421	103.481	95.838	67.756	101.491	112.247	108.498	100.090
1987	103.098	100.498	108.713	101.265	98.245	102.534	95.762	67.188	102.243	110.885	109.223	101.200
1988	103.455	99.770	107.760	101.147	98.729	103.358	94.567	68.091	102.126	111.390	107.165	102.012
1989	104.396	99.909	107.322	101.542	96.233	104.604	95.472	68.405	101.709	111.178	109.316	100.880
1990	104.061	98.694	107.446	101.443	97.643	101.905	96.160	70.407	101.507	112.424	108.516	98.660
1991	104.983	98.997	106.246	102.300	96.361	103.730	96.261	70.524	101.314	111.589	108.754	99.434
1992	104.051	99.253	106.694	102.066	97.525	101.869	95.721	71.158	102.476	112.478	107.757	99.453
1993	103.027	98.634	107.388	101.549	98.106	102.914	96.468	71.287	101.196	110.845	107.140	100.034
1994	105.054	98.649	105.960	101.774	98.359	102.998	96.705	70.878	101.708	110.349	107.022	101.276
1995	104.183	98.034	106.726	.	.	.	.	.	.	.	.	.

TABLE 4.95. D9bis: Final values for the modified seasonal-irregular ratios.

4.3.8 Table D9: Replacement Values for Extreme Values of the Seasonal-Irregular Component

Description and method of calculation

Table D8 provides an estimate of the unadjusted seasonal-irregular component, which is the seasonal-irregular factors with the extreme values. This is because it is derived from the series in Table C19, which is the raw series adjusted for trading-day effects (or from the series in Table B1 if no adjustment has been requested), by removing the trend-cycle component from Table D7.

The series in Table D1, which is the raw series adjusted for both extreme values and trading-day effects, can also be used to get another estimate of the seasonal-irregular component by removing the same trend-cycle component from Table D7. A comparison of these two series of seasonal-irregular factors (adjusted or not for extreme values) shows the replacement values that are used for points considered extreme. These values are then kept in Table D9.

Series $C19$ *op* $D7$ and $D1$ *op* $D7$ are therefore compared.

Example

The estimate of the trend-cycle obtained in Table D7 is removed from series D1 to obtain Table D9bis. For example, the value for April 1986 is:

$$APR86 = 100 \times 101.706/99.924 = 101.783.$$

This table is compared to Table D8; values which do not coincide correspond to replacement values for points considered extreme, and are displayed in Table D9. These replacement values are also shown as empty triangles in Figure 4.11 in Section 4.3.10.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**101.783**	.	.	.	67.756	.	.	.	.
1987	103.098	.	.	.	.	.	.	.	.	.	.	.
1988	.	.	.	.	.	.	.	.	.	111.390	.	.
1989	.	.	107.322	101.542	.	104.604	.	.	.	.	.	.
1990	.	.	.	.	.	.	.	.	.	.	.	.
1991	.	.	.	.	.	.	96.261	.	.	.	.	.
1992	.	.	.	.	.	.	.	.	.	.	.	99.453
1993	.	98.634	.	101.549	.	.	.	.	.	.	.	.
1994	.	.	.	.	98.359	.	.	70.878	.	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.96. D9: Final replacement values for extreme seasonal-irregular ratios.

4.3.9 Table D9A: Calculation of Moving Seasonality Ratios (MSR)

Description and method of calculation

On the basis of Table D9bis, which cannot be printed by the current versions of the softwares, we will try to measure the significance of the irregular component in relation to the seasonal component for each month. The modified seasonal-irregular component will have to be smoothed to estimate the seasonal component, and we will use a moving average whose length will depend on the importance of the irregular in the seasonal-irregular component. We have already encountered this concept in the estimation of the trend-cycle where the $\bar{I}/\bar{C}$ ratio was used to select the length of the Henderson moving average. The estimation of the MSRs includes two steps.

Step 1: Estimating the irregular and seasonal components.
An estimate of the seasonal component is obtained by smoothing, month by month and therefore column by column, Table D9bis using a simple 7-term moving average, i.e. of coefficients $\{1,1,1,1,1,1,1\}/7$. In order not to lose three points at the beginning and end of each column, all columns are completed arbitrarily. Let us assume that the column that corresponds to the month is composed of N values $\{x_1,\ldots,x_N\}$. It will be transformed into a series $\{x_{-2},x_{-1},x_0,x_1,\ldots,x_N,x_{N+1},x_{N+2},x_{N+3}\}$ with $x_{-2}=x_{-1}=x_0=(x_1+x_2+x_3)/3$ and $x_{N+1}=x_{N+2}=x_{N+3}=(x_N+x_{N-1}+x_{N-2})/3$. We then have the required estimates $S=M_7(D9bis)$ and $I=D9bis\ op\ S$.

Step 2: Calculating the Moving Seasonality Ratios.
For each month j we will be looking at the mean absolute annual change for each component by calculating:

$$\bar{S}_j = \frac{1}{n_j-1}\sum_{i=2}^{n_j}\left|S_{i,j}\ op\ S_{i-1,j} - xbar\right|, \tag{4.5}$$

$$\bar{I}_j = \frac{1}{n_j-1}\sum_{i=2}^{n_j}\left|I_{i,j}\ op\ I_{i-1,j} - xbar\right|, \tag{4.6}$$

$nyear$	CS	FIS
4	3	$\frac{90}{2\sqrt{842}+21\sqrt{2}}$
5	$\frac{3\sqrt{2}}{1+\sqrt{3}}$	$\frac{60}{\sqrt{894}+2\sqrt{211}}$
6	$\frac{5\sqrt{6}}{8+\sqrt{2}}$	$\frac{25\sqrt{3}}{2\sqrt{298}+\sqrt{67}}$
7+	$\frac{\sqrt{3}\times nyear}{6\sqrt{2}+(nyear-6)\sqrt{3}}$	$\frac{5\sqrt{6}\times nyear}{6\sqrt{149}+5\sqrt{6}(nyear-6)}$

TABLE 4.97. Values for the constants CS_i and FSI_i.

where n_j refers to the number of data points in month j. The moving seasonality ratio of month j is then $MSR_j = \bar{I}_j/\bar{S}_j$. These ratios are shown in Table D9A.

Comments

- These ratios are used to measure the significance of the noise in the seasonal-irregular component. The idea is to obtain, for each month, an indicator capable of selecting the appropriate moving average for the removal of any noise and providing a good estimate of the seasonal factor. The higher the ratio, the more chaotic the series, and the greater the order of the moving average to be used. The program selects the same moving average to smooth each month, but we have here the elements needed to select a moving average for each month.

- The computation of the MSRs is not in fact quite that simple, as an adjustment is made at the monthly level for the computation of $\bar{S}_j$ and $\bar{I}_j$. These quantities are multiplied by parameters CS_j and FIS_j respectively, whose values depend on the number of years available for each month; the theoretical calculation and the justification for these constants can be found in Lothian [47]. The values for these constants are provided in Table 4.97 according to the number of years $nyear$ in each column ($nyear = n_j$ for column j).

Example

On the basis of Table D9bis, we will now provide a detailed calculation for the month of April, as an example. The seasonal-irregular component for that month is found in Column 1 of Table 4.98.

The average of the first three years is $(101.783 + 101.265 + 101.147)/3 = 101.398$. That for the last three years is $(102.066 + 101.549 + 101.774)/3 =$

Year	Column 1	Column 2	MA7
		101.398	
		101.398	
1985		101.398	
1986	101.783	101.783	101.419
1987	101.265	101.265	101.425
1988	101.147	101.147	101.554
1989	101.542	101.542	101.650
1990	101.443	101.443	101.616
1991	102.300	102.300	101.689
1992	102.066	102.066	101.782
1993	101.549	101.549	101.818
1994	101.774	101.774	101.868
1995		101.796	
		101.796	
		101.796	

TABLE 4.98. MSR: April SI and seasonal factors.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.457	108.901	100.621
1986	103.792	99.974	107.693	**101.419**	97.861	103.335	95.401	67.782	101.919	111.415	108.919	100.691
1987	103.845	99.789	107.649	101.425	97.791	103.161	95.511	68.172	101.855	111.552	108.823	100.444
1988	104.029	99.648	107.433	101.554	97.538	103.248	95.636	68.579	101.763	111.569	108.760	100.308
1989	104.081	99.544	107.282	101.650	97.451	103.068	95.683	69.076	101.838	111.742	108.461	100.247
1990	103.867	99.394	107.367	101.616	97.549	102.988	95.773	69.580	101.796	111.541	108.267	100.239
1991	104.147	99.130	106.974	101.689	97.565	103.054	95.908	70.107	101.719	111.465	107.953	100.250
1992	104.251	98.882	106.826	101.782	97.461	102.945	96.155	70.538	101.672	111.441	107.973	99.999
1993	104.207	98.672	106.736	101.818	97.713	102.658	96.273	70.924	101.684	111.448	107.686	99.909
1994	104.210	98.635	106.628	101.868	97.763	102.756	96.293	71.024	101.724	111.276	107.513	100.137
1995	104.083	98.555	106.692	.	.	.	.	.	.	.	.	.

TABLE 4.99. D9A1: Seasonal factors.

101.796. The series to be dealt with is therefore that found in column 2 of Table 4.98. By smoothing using a moving average of order 7, we have, e.g. for April of 1986:

$$\begin{aligned} APR86 &= (101.398 + 101.398 + 101.398 + 101.783 + \\ &\quad 101.265 + 101.147 + 101.542)/7 \\ &= 101.419. \end{aligned}$$

These calculations provide the estimate of seasonal factors found in Table D9A1. By removing this estimate from Table D9bis, we obtain an estimate of the irregular component in Table D9A2. Thus:

$$APR86 = 100 \times 101.783/101.419 = 100.359.$$

Column by column, we then calculate the mean annual variations of the irregular and of the seasonal component using the formulae (4.5) and (4.6) corresponding to the multiplicative model. The absolute annual growths, in percentages, in the seasonal and irregular components are in Tables D9A3 and D9A4. For example, the April 1988 value in Table D9A3 is

$$\begin{aligned} APR88 &= 100 \times |101.554 - 101.425| / 101.425 \\ &= 0.127, \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	99.834	100.871	99.261
1986	100.703	99.711	99.162	**100.359**	99.551	100.141	100.458	99.961	99.581	100.746	99.614	99.403
1987	99.281	100.711	100.989	99.842	100.465	99.391	100.263	98.557	100.381	99.402	100.368	100.753
1988	99.448	100.122	100.304	99.599	101.221	100.106	98.883	99.289	100.357	99.840	98.534	101.699
1989	100.303	100.367	100.038	99.895	98.750	101.490	99.779	99.029	99.873	99.496	100.788	100.631
1990	100.187	99.296	100.073	99.830	100.097	98.949	100.404	101.189	99.717	100.792	100.230	98.425
1991	100.803	99.867	99.320	100.601	98.765	100.656	100.369	100.594	99.601	100.111	100.742	99.186
1992	99.808	100.375	99.877	100.279	100.066	98.955	99.549	100.879	100.791	100.931	99.800	99.455
1993	98.868	99.962	100.611	99.736	100.403	100.250	100.202	100.511	99.520	99.459	99.493	100.125
1994	100.809	100.014	99.374	99.907	100.610	100.236	100.428	99.794	99.983	99.167	99.543	101.137
1995	100.097	99.471	100.032	.	.	.	.	.	.	.	.	.

TABLE 4.100. D9A2: Irregular component.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	0.	.
1986	.	.	.	.	.	.	.	.	.	0.0371	0.0165	0.0696
1987	0.0509	0.1844	0.0410	0.0063	0.0713	0.1685	0.1155	0.5752	0.0625	0.1226	0.0884	0.2454
1988	0.1776	0.1413	0.2002	**0.1270**	0.2588	0.0839	0.1305	0.5963	0.0898	0.0155	0.0572	0.1360
1989	0.0493	0.1049	0.1410	0.0939	0.0889	0.1737	0.0496	0.7249	0.0733	0.1544	0.2748	0.0604
1990	0.2051	0.1508	0.0795	0.0329	0.1004	0.0785	0.0941	0.7302	0.0415	0.1792	0.1790	0.0081
1991	0.2689	0.2658	0.3663	0.0715	0.0167	0.0645	0.1407	0.7575	0.0751	0.0686	0.2904	0.0108
1992	0.0999	0.2502	0.1381	0.0912	0.1072	0.1059	0.2577	0.6148	0.0468	0.0213	0.0186	0.2505
1993	0.0422	0.2124	0.0844	0.0356	0.2585	0.2789	0.1227	0.5473	0.0119	0.0059	0.2659	0.0893
1994	0.0037	0.0369	0.1010	0.0495	0.0517	0.0959	0.0204	0.1411	0.0401	0.1538	0.1605	0.2280
1995	0.1226	0.0809	0.0596	.	.	.	.	.	.	.	.	.

TABLE 4.101. D9A3: Seasonal factors absolute annual growth rates (in %).

and the April 1988 value in Table D9A4 is

$$\begin{aligned} APR88 &= 100 \times |99.599 - 99.842| / 99.842 \\ &= 0.243. \end{aligned}$$

To calculate the Moving Seasonality Ratio for the month of April, for example, we have, with eight years of observations:

$$\begin{aligned} CS_{APR} &= \frac{8 \times 1.732051}{8.485281 + (8-6) \times 1.732051} = 1.1596 \\ FSI_{APR} &= \frac{8 \times 12.247449}{73.239334 + (8-6) \times 12.247449} = 1.0025 \\ \bar{I}_{APR} &= \left(\frac{0.5154 + 0.2434 + 0.2969 + 0.0648}{8} + \right. \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	.	.	.	.	.	0.9131	1.2463	0.1430
1987	1.4115	1.0026	1.8419	0.5154	0.9177	0.7482	0.1942	1.4050	0.8034	1.3343	0.7568	1.3578
1988	0.1679	0.5841	0.6776	**0.2434**	0.7531	0.7193	1.3761	0.7425	0.0241	0.4405	1.8274	0.9392
1989	0.8598	0.2446	0.2659	0.2969	2.4415	1.3817	0.9062	0.2614	0.4821	0.3446	2.2876	1.0502
1990	0.1157	1.0674	0.0357	0.0648	1.3638	2.5036	0.6266	2.1809	0.1564	1.3025	0.5536	2.1928
1991	0.6148	0.5749	0.7530	0.7724	1.3301	1.7250	0.0354	0.5876	0.1158	0.6751	0.5114	0.7735
1992	0.9865	0.5093	0.5606	0.3196	1.3171	1.6898	0.8170	0.2832	1.1944	0.8188	0.9352	0.2711
1993	0.9421	0.4118	0.7348	0.5417	0.3364	1.3088	0.6563	0.3646	1.2607	1.4584	0.3080	0.6735
1994	1.9633	0.0523	1.2294	0.1714	0.2061	0.0140	0.2256	0.7141	0.4654	0.2936	0.0508	1.0115
1995	0.7066	0.5430	0.6625	.	.	.	.	.	.	.	.	.

TABLE 4.102. D9A4: Irregular component absolute annual growth rates (in %).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$\bar{I}$	0.865	0.556	0.753	**0.367**	1.086	1.264	0.606	0.819	0.564	0.844	0.944	0.937
$\bar{S}$	0.129	0.181	0.153	**0.074**	0.138	0.152	0.135	0.679	0.064	0.096	0.171	0.139
Ratio	6.697	3.075	4.911	**4.979**	7.858	8.310	4.491	1.206	8.826	8.790	5.518	6.739

TABLE 4.103. D9A: Year to year change in irregular and seasonal components and moving seasonality ratio.

$$
\begin{aligned}
& \left.\frac{0.7724 + 0.3196 + 0.5417 + 0.1714}{8}\right) \times FSI_{APR} \\
&= 0.3657 \times 1.0025 = 0.3666 \\
\bar{S}_{APR} &= \left(\frac{0.0063 + 0.1270 + 0.0939 + 0.0329}{8} + \right. \\
& \left.\frac{0.0715 + 0.0912 + 0.0356 + 0.0495}{8}\right) \times CS_{APR} \\
&= 0.0635 \times 1.1596 = 0.0736
\end{aligned}
$$

which yields finally:

$$MSR_{APR} = 0.3666/0.0736 = 4.98.$$

The complete results are shown in Table D9A.

4.3.10 Table D10: Final Estimation of the Seasonal Factors

Description and method of calculation

Step 1: Calculating the overall Moving Seasonality Ratio.
To calculate an overall Moving Seasonality Ratio, the program uses all available data up to the last complete year, so that there is no need, during the year, to change the filter. The program calculates, for each month, the mean values for changes in seasonal and irregular components using the method described previously in Table D9A. The overall ratio is derived by weighting these quantities using the number of months for each type:

$$MSR = \frac{\sum_j n_j \bar{I}_j}{\sum_j n_j \bar{S}_j}.$$

Step 2: Selecting a moving average and estimating the seasonal component.
Depending on the value of the ratio, the program automatically selects a moving average that is applied, column by column (i.e. month by month) to the seasonal-irregular component in Table D8 modified, for extreme values, using values in Table D9.

A $MSR < 2.5$	B $2.5 < MSR < 3.5$	C $3.5 < MSR < 5.5$	D $5.5 < MSR < 6.5$	E $6.5 < MSR$
3×3		3×5		3×9

TABLE 4.104. Selection criteria for the seasonal moving average.

In X-11-ARIMA, the selection of a moving average is based on the following strategy (refer to Table 4.104):

- If the overall MSR occurs within zone A ($MSR < 2.5$), a 3×3 moving average is used; if it occurs within zone C ($3.5 < MSR < 5.5$), a 3×5 moving average is selected; if it occurs within zone E ($MSR > 6.5$), a 3×9 moving average is used.

- If the MSR occurs within zone B or D, one year of observations is removed from the end of the series, and the MSR is re-calculated. If the ratio again occurs within zones B or D, we start over again, removing a maximum of five years of observations. If this does not work, i.e. if we are again within zones B or D, a 3×5 moving average is selected.

The chosen symmetric moving average corresponds, as the case may be, to 5 (3×3), 7 (3×5) or 11 (3×9) terms, and therefore does not provide an estimate for the values of seasonal factors in the first 2 (or 3 or 5) and the last 2 (or 3 or 5) years. These are then calculated using the associated asymmetric moving averages. The result is a series of preliminary seasonal factors *fspro*.

Step 3: Normalizing the seasonal factors.
These preliminary seasonal factors are then normalized using a centered 12-term moving average.

Step 4: Forecast of seasonal factors.
Seasonal factors are then forecast over one year by simple linear projection on the basis of the last two seasonal factors for a given month. If n_j is the last available year for a given month j, we have:

$$\begin{aligned} S_{n_j+1,j} &= S_{n_j,j} + (S_{n_j,j} - S_{n_j-1,j})/2 \\ &= (3S_{n_j,j} - S_{n_j-1,j})/2. \end{aligned}$$

These forecasts are then shown in Table D10A.

Comments

- A problem could arise in the application of these asymmetric moving averages if the number of years of observations is not sufficient.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$\bar{I}$	0.883	0.544	0.765	0.367	1.086	1.264	0.606	0.819	0.564	0.844	0.944	0.937
$\bar{S}$	0.128	0.168	0.168	0.074	0.138	0.152	0.135	0.679	0.064	0.096	0.171	0.139
Ratio	6.894	3.248	4.549	4.979	7.858	8.310	4.491	1.206	8.826	8.790	5.518	6.739

TABLE 4.105. D10msr: Year to year change in irregular and seasonal components and moving seasonality ratios.

Let us assume that, for a given month, we have only five years of observations for a 3×5 moving average. The central point cannot be estimated using the symmetric moving average because there are neither three points in the future, nor three points in the past. In such a case, the estimation is based on a simple average of the five available observations. Each time we encounter such a problem, for example with a 3×9 and less than eleven years of observations, a simple average of the available observations is used to estimate the seasonal factors.

- The procedure described is that used by default, but we can select the moving average to be used. In such a case, X-11-ARIMA gives the choice between a simple 3-term, a 3×3, a 3×5, a 3×9 and constant seasonality (a simple average). X-12-ARIMA also includes a 3×15.

Example

The counterpart of Table D9A calculated for the data available up to December 1994 (the last complete year available) is given in Table D10msr.

As can be seen, only the values for the first three months, those for which the last values have been excluded, are modified. The overall ratio is easily derived as follows:

$$\begin{aligned}
\sum_j n_j \bar{I}_j &= 8(0.883 + 0.544 + 0.765 + 0.367 + 1.086 + 1.264) + \\
&\quad 8(0.606 + 0.819 + 0.564) + 9(0.844 + 0.944 + 0.937) \\
&= 79.709, \\
\sum_j n_j \bar{S}_i &= 8(0.128 + 0.168 + 0.168 + 0.074 + 0.138 + 0.152) + \\
&\quad 8(0.135 + 0.679 + 0.064) + 9(0.096 + 0.171 + 0.139) \\
&\quad 17.302, \\
MSR &= \frac{\sum_j n_j \bar{I}_j}{\sum_j n_j \bar{S}_j} \\
&= \frac{79.709}{17.302}
\end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.456	108.887	100.633
1986	103.797	99.973	107.690	101.420	97.847	103.346	95.401	67.787	101.917	111.434	108.827	100.720
1987	103.845	99.885	107.655	**101.428**	97.751	103.314	95.456	67.994	101.879	111.489	108.679	100.725
1988	103.924	99.750	107.584	101.470	97.585	103.299	95.538	68.384	101.829	111.504	108.650	100.606
1989	104.031	99.537	107.399	**101.558**	97.465	103.165	95.614	69.003	101.807	111.643	108.480	100.364
1990	104.097	99.332	107.204	101.673	97.305	103.108	95.766	69.666	101.749	111.669	108.398	100.072
1991	104.176	99.089	106.953	101.769	97.357	102.927	95.972	70.308	101.702	111.684	108.145	99.850
1992	104.199	98.886	106.790	101.843	97.453	102.855	96.189	70.723	101.651	111.518	107.934	99.838
1993	104.191	98.734	106.681	101.865	97.673	102.756	96.282	70.954	101.686	111.371	107.652	99.998
1994	104.141	98.626	106.662	101.872	97.751	102.764	96.292	71.020	101.721	111.279	107.524	100.131
1995	104.082	98.561	106.692	.	.	.	.	.	.	.	.	.

TABLE 4.106. D10bis: Preliminary seasonal factors.

$$= \quad 4.607.$$

The overall Moving Seasonality Ratio is calculated to be 4.607. It therefore falls directly within zone C, and a 3×5 moving average is selected (see Table 3.12) to smooth the modified seasonal-irregular component in Table D9bis (Section 4.3.8) and thus provide Table D10bis. For example, the seasonal factor for the month of April 1989 is estimated as follows:

$$\begin{aligned} APR89 &= \frac{101.783 + 101.265 \times 2 + 101.147 \times 3 + 101.542 \times 3}{15} + \\ &\quad \frac{101.443 \times 3 + 102.300 \times 2 + 102.066}{15} \\ &= 101.558. \end{aligned}$$

For the beginning of the series (years 1986 to 1988) and the end of the series (years 1992 to 1994), the associated asymmetric averages are used, e.g.:

$$\begin{aligned} APR87 &= \frac{101.783 \times 15 + 101.265 \times 15 + 101.147 \times 15}{60} + \\ &\quad \frac{101.542 \times 11 + 101.443 \times 4}{60} \\ &= 101.428 \end{aligned}$$

(one point in the past, the current point and three points in the future).

These factors are then normalized by applying a centered moving average of order 12 which leads to Table D10ter. We have, for example:

$$\begin{aligned} APR86 &= \frac{111.456}{24} + \\ &\quad \frac{108.887 + 100.633 + 103.797 + 99.973 + 107.690}{12} + \\ &\quad \frac{101.420 + 97.847 + 103.346 + 95.401 + 67.787 + 101.917}{12} + \\ &\quad \frac{111.434}{24} \\ &= 100.012. \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.012	100.012	100.012
1986	100.012	100.012	100.012	**100.012**	100.009	100.010	100.015	100.014	100.009	100.007	100.004	99.998
1987	99.999	100.010	100.017	100.018	100.014	100.008	100.012	100.009	100.001	100.000	99.994	99.987
1988	99.990	100.009	100.023	100.022	100.021	100.015	100.015	100.010	99.994	99.990	99.988	99.978
1989	99.975	100.004	100.029	100.034	100.033	100.016	100.008	100.003	99.986	99.983	99.981	99.972
1990	99.976	100.009	100.035	100.033	100.031	100.015	100.006	100.000	99.979	99.973	99.979	99.973
1991	99.974	100.010	100.035	100.033	100.023	100.004	99.995	99.988	99.973	99.969	99.976	99.977
1992	99.983	100.009	100.025	100.016	100.000	99.991	99.990	99.983	99.972	99.968	99.979	99.984
1993	99.983	99.997	100.008	100.003	99.985	99.980	99.985	99.978	99.973	99.972	99.976	99.980
1994	99.980	99.984	99.988	99.985	99.976	99.976	99.979	99.974	99.973	99.973	99.973	99.973
1995	99.973	99.973	99.973	.	.	.	.	.	.	.	.	.

TABLE 4.107. D10ter: Centered 12-term moving average.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	111.443	108.874	100.621
1986	103.785	99.961	107.677	**101.408**	97.839	103.336	95.387	67.778	101.908	111.426	108.823	100.721
1987	103.846	99.874	107.636	101.410	97.738	103.306	95.445	67.987	101.878	111.490	108.686	100.739
1988	103.935	99.741	107.558	101.448	97.564	103.283	95.524	68.377	101.836	111.516	108.662	100.628
1989	104.057	99.532	107.368	101.524	97.433	103.149	95.606	69.001	101.821	111.663	108.501	100.393
1990	104.122	99.323	107.167	101.639	97.275	103.092	95.760	69.666	101.770	111.700	108.421	100.099
1991	104.202	99.080	106.916	101.735	97.334	102.923	95.976	70.317	101.730	111.719	108.171	99.873
1992	104.217	98.876	106.764	101.827	97.454	102.865	96.199	70.735	101.679	111.554	107.957	99.854
1993	104.208	98.737	106.672	101.862	97.688	102.776	96.297	70.969	101.714	111.402	107.678	100.018
1994	104.161	98.642	106.675	101.887	97.775	102.789	96.312	71.038	101.749	111.309	107.553	100.158
1995	104.111	98.588	106.721	.	.	.	.	.	.	.	.	.

TABLE 4.108. D10: Final seasonal factors. A 3×5 moving average has been selected; I/S ratio = 4.949.

There now remains to adjust Table D10bis using Table D10ter to obtain a final estimate of the seasonal factors in Table D10. For example,

$$APR86 = 100 \times 101.420/100.012 = 101.408.$$

The final seasonal factors for each month are represented in Figure 4.11 by the solid line going through the monthly seasonal-irregular components. In those graphs, the extreme values in the seasonal-irregular component are replaced by empty triangles. The final seasonal component is also graphed in the third panel of Figure 4.1 at the beginning of Chapter 4. Finally, we can now forecast the seasonal factor for the month of April 1995 as follows:

$$\begin{aligned} APR95 &= APR94 + \frac{APR94 - APR93}{2} \\ &= \frac{3 \times 101.887 - 101.862}{2} \\ &= 101.899. \end{aligned}$$

This leads to Table D10A.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1995	.	.	.	**101.899**	97.818	102.795	96.320	71.073	101.766	111.262	107.490	100.229
1996	104.085	98.561	106.743	.	.	.	.	.	.	.	.	.

TABLE 4.109. D10A: Seasonal factors 12 months ahead.

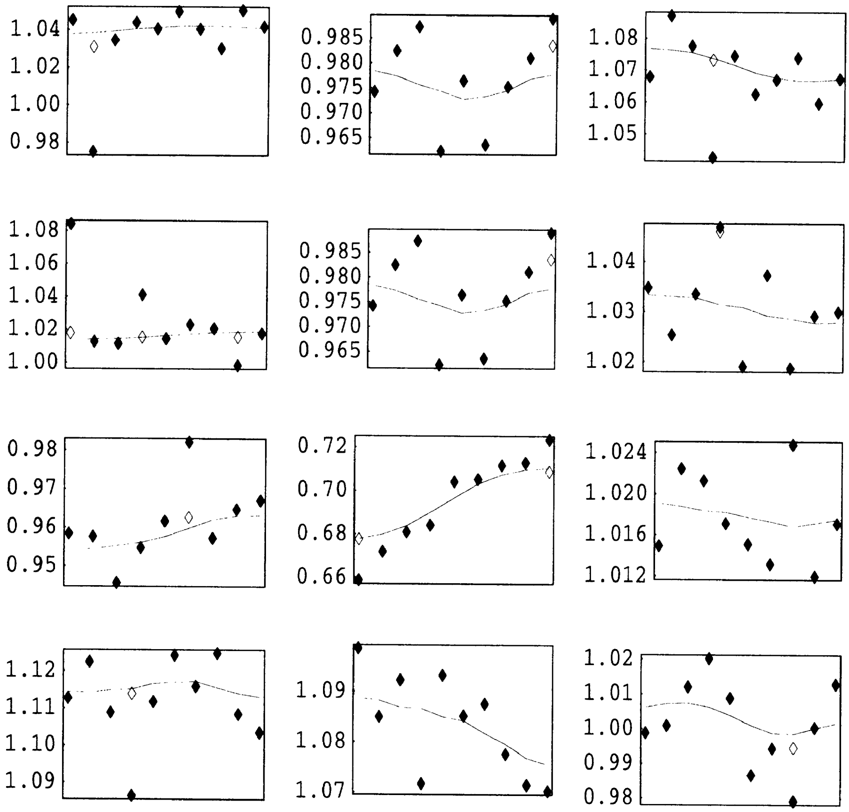


FIGURE 4.11. Monthly sub-plots of the unmodified seasonal-irregulars from Table D8 (solid triangles), their replacement values from Table D9 (empty triangles), and the final seasonal factors from Table D10 (solid line), from January to December, row-wise. Note the change in the values of the y-axis for the various sub-plots.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	101.587	102.236	100.085
1986	101.034	99.620	98.909	**106.822**	100.020	101.281	102.274	99.415	102.091	103.374	102.421	102.275
1987	96.935	104.252	105.001	104.088	104.934	103.677	104.916	103.579	105.552	105.022	106.510	106.845
1988	106.275	107.287	107.985	108.020	110.198	109.453	108.654	109.671	110.936	108.373	110.433	114.168
1989	113.423	113.734	110.122	116.272	111.904	114.971	113.220	112.595	113.697	113.586	115.180	115.133
1990	114.759	114.279	115.521	115.285	116.283	114.852	116.874	117.617	115.890	116.603	115.563	113.463
1991	115.783	114.864	114.439	116.068	114.501	116.692	118.537	116.234	115.479	115.963	116.956	116.069
1992	116.566	117.119	116.267	116.147	115.553	114.120	114.624	115.931	116.072	115.827	114.147	111.515
1993	111.902	115.569	113.496	110.520	113.221	112.715	112.439	112.420	111.238	111.397	111.774	112.903
1994	114.450	114.044	113.741	114.818	116.718	116.020	116.603	118.619	116.844	116.402	117.460	120.004
1995	119.326	119.019	120.007	.	.	.	.	.	.	.	.	.

TABLE 4.110. D11: Final seasonally adjusted series (including trading-day and Easter effects when applicable).

4.3.11 Table D11: Final Trading-Day and Seasonally Adjusted Series

Description and method of calculation

The starting series, or the series in Table C19 if a trading-day adjustment has been requested, is seasonally adjusted using Table D10, and we obtain the final seasonally adjusted series (Table D11): $D11 = C19 \; op \; D10$.

An F-test, similar to those used in Tables B3 and D8, is used to check that there is no residual seasonality in series D11. The trend is first removed by differencing the D11 series: first differences $(X_t - X_{t-1})$ are used for a quarterly series and 3-months differences $(X_t - X_{t-3})$ for a monthly series. The test is then performed first on the complete differenced series, and next only on its last three years (the 36 last observations in the monthly case and the 12 last ones in the quarterly case).

Only the F-statistics and the results of the tests at the 1% and 5% levels are printed. A message cautions the user against misinterpretations of the results when the series level fluctuates a lot in the last years.

Example

The value for April 86 is:

$$APR86 = 100 \times 108.326/101.408 = 106.822.$$

This leads to the final seasonally adjusted series in Table D11, whose values are also graphed with the original series in the top panel of Figure 4.1.

The seasonally adjusted series is then detrended by 3-month differencing; thus, we have for instance

$$APR86 = 106.822 - 101.034 = 5.788.$$

This results in Table 4.111.

At last, the results of the F-tests are printed (Table 4.112).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	-0.553	-2.616	-1.176	**5.788**	0.400	2.372	-4.548	-0.604	0.810	1.100	3.005	0.184
1987	-6.440	1.832	2.726	7.154	0.682	-1.324	0.827	-1.355	1.875	0.107	2.931	1.293
1988	1.253	0.776	1.140	1.745	2.911	1.467	0.634	-0.527	1.483	-0.281	0.763	3.233
1989	5.050	3.300	-4.046	2.849	-1.830	4.849	-3.052	0.691	-1.274	0.366	2.585	1.436
1990	1.173	-0.901	0.388	0.525	2.004	-0.669	1.589	1.334	1.038	-0.271	-2.053	-2.427
1991	-0.820	-0.699	0.976	0.286	-0.364	2.252	2.468	1.733	-1.212	-2.574	0.722	0.590
1992	0.603	0.163	0.198	-0.419	-1.566	-2.147	-1.523	0.377	1.951	1.202	-1.784	-4.556
1993	-3.925	1.423	1.981	-1.381	-2.349	-0.781	1.919	-0.800	-1.477	-1.042	-0.646	1.665
1994	3.054	2.270	0.837	0.368	2.673	2.279	1.785	1.901	0.824	-0.201	-1.159	3.160
1995	2.924	1.559	0.003	.	.	.	.	.	.	.	.	.

TABLE 4.111. 3-month differenced seasonally adjusted series.

NO EVIDENCE OF RESIDUAL SEASONALITY IN THE ENTIRE SERIES F=0.52
AT THE 1 PER CENT LEVEL

NO EVIDENCE OF RESIDUAL SEASONALITY IN THE LAST 3 YEARS F=0.38
AT THE 1 PER CENT LEVEL
NO EVIDENCE OF RESIDUAL SEASONALITY IN THE LAST 3 YEARS
AT THE 5 PER CENT LEVEL

NOTE: SUDDEN LARGE CHANGES IN THE LEVEL OF THE SEASONALLY ADJUSTED SERIES WILL INVALIDATE THE RESULTS OF THIS TEST FOR THE LAST THREE YEAR PERIOD.

TABLE 4.112. Test for the presence of residual seasonality.

4.3.12 Table D11A: Final Seasonally Adjusted Series with Revised Annual Totals

Description and method of calculation

The user may request that the seasonally adjusted series in Table D11 be adjusted so that its annual totals equal those of the raw series (eventually preadjusted), resulting in Table D11A. In such a case, the annual differences observed between the series in Tables D11 and A1 are distributed among the values of the seasonally adjusted series in such a way that the monthly (or quarterly) changes are roughly the same for series D11 and D11A.

Complete theoretical details for the adjustment may be found in Cholette [12] or in Cholette and Dagum [14].

Let us assume, for simplicity, that our series - original x_t, seasonally adjusted a_t, and sum-adjusted $\tilde{a}_t$ - include $n = Nk$ points distributed over N complete years of k periods ($k = 12$ months or $k = 4$ quarters). Changes in the unknown series $\tilde{a}_t$ must parallel as much as possible changes in the known series a_t, provided its annual sums are equal to those of series x_t.

Letting ∇ denote the first difference operator[20], the problem is to minimize, with respect to the unknown $\tilde{a}$'s, the quantity $\sum_{t=2}^{n} [\nabla(\tilde{a}_t - a_t)]^2$ subject to the constraint $\sum_{j=1}^{k} \tilde{a}_{ij} = x_{i.}, i = 1, \ldots N$, where $\tilde{a}_{i1}, \ldots, \tilde{a}_{ik}$ are the k unknown benchmarked seasonally adjusted observations in year

[20] $\nabla x_t = x_t - x_{t-1}$

i, and $x_{i.}$ is the annual total of the original observations in year i. Or again, by defining the following matrices:

$$\Delta_{(n-1)\times n} = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix},$$

$$\Upsilon_{n\times k} = \begin{bmatrix} 1_k & 0 & \dots & 0 & 0 \\ 0 & 1_k & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1_k & 0 \\ 0 & 0 & \dots & 0 & 1_k \end{bmatrix},$$

where 1_k is a vector of ones of dimension $k \times 1$. The problem is then to minimize, with respect to $\tilde{A}$, the quantity $(\tilde{A} - A)'\Delta'\Delta(\tilde{A} - A) = (\tilde{A} - A)'\Gamma(\tilde{A} - A)$ subject to the constraint $X - \Upsilon'\tilde{A} = 0$. If we introduce a vector Λ of N Lagrange multipliers, the solution can be written as follows:

$$\begin{aligned} \begin{bmatrix} \tilde{A} \\ \Lambda \end{bmatrix} &= \begin{bmatrix} \Gamma & \Upsilon \\ \Upsilon' & 0 \end{bmatrix}^{-1} \begin{bmatrix} \Gamma & 0 \\ \Upsilon' & I_N \end{bmatrix} \begin{bmatrix} A \\ X - \Upsilon'A \end{bmatrix} \\ &= W \begin{bmatrix} A \\ X - \Upsilon'A \end{bmatrix} \\ &= W \begin{bmatrix} A \\ R \end{bmatrix} \end{aligned}$$

where I_N denotes the identity matrix of order N. As can be seen, matrix W, of size $(n+N) \times (n+N)$, does not depend on the data, X and A, and it can be shown to always have the form

$$W = \begin{bmatrix} I_n & W_{\tilde{A}} \\ 0 & W_\Lambda \end{bmatrix}$$

where $W_{\tilde{A}}$ is a matrix of size $n \times N$.

Finally, we have:

$$\tilde{a}_t = a_t + \sum_{i=1}^{N} w_{\tilde{A}_{(t,i)}} r_i, (t = 1, \dots, n).$$

In other words, the adjustment factor applied to series a_t (our series D11) is a weighted average of the differences between the annual totals for series x_t (our series A1) and a_t over the N years.

Comments

- The above calculation is rendered somewhat complex because matrix $\Gamma = \Delta'\Delta$ is not invertible. A solution, based on a generalized least

	1	2	3	4	5
Jan	0.10539496	-0.02790048	0.00737759	-0.00191944	0.00038069
Feb	0.10446930	-0.02672983	0.00706804	-0.00183890	0.00036472
Mar	0.10261797	-0.02438853	0.00644894	-0.00167783	0.00033277
Apr	0.09984099	-0.02087658	0.00552030	-0.00143622	0.00028485
May	0.09613833	-0.01619398	0.00428210	-0.00111408	0.00022096
June	0.09151002	-0.01034074	0.00273435	-0.00071140	0.00014110
July	0.08595605	-0.00331684	0.00087706	-0.00022818	0.00004526
Aug	0.07947641	0.00487771	-0.00128979	0.00033557	-0.00006655
Sep	0.07207110	0.01424290	-0.00376618	0.00097985	-0.00019434
Oct	0.06374014	0.02477874	-0.00655213	0.00170467	-0.00033810
Nov	0.05448351	0.03648524	-0.00964762	0.00251003	-0.00049783
Dec	0.04430122	0.04936238	-0.01305266	0.00339593	-0.00067353
Jan	0.03319327	0.06341017	-0.01676725	0.00436236	-0.00086521
Feb	0.02325596	0.07505210	-0.01892111	0.00492273	-0.00097635
Mar	0.01448931	0.08428817	-0.01951423	0.00507704	-0.00100696
Apr	0.00689330	0.09111838	-0.01854662	0.00482530	-0.00095703
May	0.00046795	0.09554273	-0.01601827	0.00416749	-0.00082656
June	-0.00478676	0.09756121	-0.01192919	0.00310363	-0.00061556
July	-0.00887082	0.09717384	-0.00627938	0.00163371	-0.00032402
Aug	-0.01178422	0.09438060	0.00093117	-0.00024226	0.00004805
Sep	-0.01352698	0.08918150	0.00970246	-0.00252430	0.00050066
Oct	-0.01409909	0.08157654	0.02003448	-0.00521239	0.00103380
Nov	-0.01350055	0.07156572	0.03192723	-0.00830655	0.00164748
Dec	-0.01173136	0.05914904	0.04538072	-0.01180676	0.00234170
Jan	-0.00879152	0.04432649	0.06039494	-0.01571303	0.00311645
Feb	-0.00616123	0.03106468	0.07290679	-0.01805856	0.00358166
Mar	-0.00384049	0.01936360	0.08291627	-0.01884336	0.00373731
Apr	-0.00182930	0.00922326	0.09042338	-0.01806743	0.00358341
May	-0.00012766	0.00064366	0.09542812	-0.01573076	0.00311997
June	0.00126443	-0.00637522	0.09793049	-0.01183335	0.00234698
July	0.00234698	-0.01183335	0.09793049	-0.00637522	0.00126443
Aug	0.00311997	-0.01573076	0.09542812	0.00064366	-0.00012766
Sep	0.00358341	-0.01806743	0.09042338	0.00922326	-0.00182930
Oct	0.00373731	-0.01884336	0.08291627	0.01936361	-0.00384049
Nov	0.00358166	-0.01805856	0.07290679	0.03106468	-0.00616123
Dec	0.00311645	-0.01571303	0.06039494	0.04432649	-0.00879152
Jan	0.00234170	-0.01180676	0.04538072	0.05914904	-0.01173136
Feb	0.00164748	-0.00830655	0.03192723	0.07156572	-0.01350055
Mar	0.00103380	-0.00521239	0.02003448	0.08157654	-0.01409909
Apr	0.00050066	-0.00252430	0.00970246	0.08918150	-0.01352698
May	0.00004805	-0.00024226	0.00093117	0.09438060	-0.01178422
June	-0.00032402	0.00163371	-0.00627938	0.09717384	-0.00887082
July	-0.00061556	0.00310363	-0.01192919	0.09756121	-0.00478676
Aug	-0.00082656	0.00416749	-0.01601827	0.09554273	0.00046795
Sep	-0.00095703	0.00482530	-0.01854662	0.09111838	0.00689330
Oct	-0.00100696	0.00507704	-0.01951423	0.08428817	0.01448931
Nov	-0.00097635	0.00492273	-0.01892111	0.07505210	0.02325596
Dec	-0.00086521	0.00436236	-0.01676725	0.06341017	0.03319327
Jan	-0.00067353	0.00339593	-0.01305266	0.04936238	0.04430122
Feb	-0.00049783	0.00251003	-0.00964762	0.03648524	0.05448351
Mar	-0.00033810	0.00170467	-0.00655213	0.02477874	0.06374014
Apr	-0.00019434	0.00097985	-0.00376618	0.01424290	0.07207110
May	-0.00006655	0.00033557	-0.00128979	0.00487771	0.07947641
June	0.00004526	-0.00022818	0.00087706	-0.00331684	0.08595605
July	0.00014110	-0.00071140	0.00273435	-0.01034074	0.09151002
Aug	0.00022096	-0.00111408	0.00428210	-0.01619398	0.09613833
Sep	0.00028485	-0.00143622	0.00552030	-0.02087658	0.09984099
Oct	0.00033277	-0.00167783	0.00644894	-0.02438853	0.10261797
Nov	0.00036472	-0.00183890	0.00706804	-0.02672983	0.10446930
Dec	0.00038069	-0.00191944	0.00737759	-0.02790048	0.10539496

TABLE 4.113. Monthly benchmarking weights matrix used by X-11-ARIMA and X-12-ARIMA for Table D11A.

squares regression, can be found in Cholette and Dagum [14] and in Bournay and Laroque [8].

- Since the adjustment factors depend on all the differences in the observed annual totals, any new year of complete data leads to a revision of the entire series D11A. This is not desirable, and therefore programs X-11-ARIMA and X-12-ARIMA only incorporate values for the matrix of weights $W_{\tilde{A}}$ for 5 years of observations: a matrix of size 60×5 for the monthly case and 20×5 for the quarterly. This is the matrix used, regardless of the length of the series, to calculate the adjusted series D11A. Then (see Table 4.113), the first complete year of data is associated with the first 12 lines (for the monthly case) of matrix $W_{\tilde{A}}$, and the second year is associated with the next 12 lines. The last two complete years are likewise associated with the last 24 lines of the matrix. All the other years, from the third to the antepenultimate, are associated with the 12 central lines of matrix $W_{\tilde{A}}$.
- If the last year is not complete, the estimates in Table D11A for the months concerned are prepared using the last calculated adjustment factor (i.e. that for the month of December of the last complete year).
- X-12-ARIMA and X-11-ARIMA consider only cases of additive adjustment that are not necessarily compatible with a multiplicative decomposition model. Thus, for this type of model, it is possible for Table D11A to include negative values. Cholette and Dagum [14] provide a solution for the multiplicative case, but that solution is not available in the current versions of the softwares.
- X-12-ARIMA can constrain over any period of 12 months, such as April to March (financial year).
- The weights for the quarterly case are shown in Table 4.114.

Example

Series D11 and B1 must be reconciled. The values for the annual totals of these two series and for their differences (residuals R) are provided in Table 4.115. As can be seen, the differences between annual totals are fairly small. This is due to the fact that, during the seasonal adjustment process, we have regularly normalized[21] the estimates of seasonal factors.

For the data of the first complete year (1986), the weight and residual matrices used are provided in Table 4.116.

[21] I.e. ensured that their sum is roughly equal to 0 or 12, depending on the model, for any one-year period.

	1	2	3	4	5
Q1	0.31010142	-0.07454766	0.01790831	-0.00424890	0.00078683
Q2	0.28606085	-0.04472860	0.01074499	-0.00254934	0.00047210
Q3	0.23797972	0.01490953	-0.00358166	0.00084978	-0.00015737
Q4	0.16585801	0.10436673	-0.02507163	0.00594845	-0.00110156
Q1	0.06969574	0.22364300	-0.05372493	0.01274669	-0.00236050
Q2	0.00335254	0.28189630	-0.04369627	0.01036731	-0.00191987
Q3	-0.03317160	0.27912665	0.00501433	-0.00118969	0.00022031
Q4	-0.03987667	0.21533404	0.09240688	-0.02192430	0.00406006
Q1	-0.01676268	0.09051848	0.21848137	-0.05183653	0.00959936
Q2	-0.00081201	0.00438486	0.28151862	-0.04306681	0.00797533
Q3	0.00797533	-0.04306681	0.28151862	0.00438486	-0.00081201
Q4	0.00959936	-0.05183653	0.21848137	0.09051848	-0.01676268
Q1	0.00406006	-0.02192430	0.09240688	0.21533404	-0.03987667
Q2	0.00022031	-0.00118969	0.00501433	0.27912665	-0.03317160
Q3	-0.00191987	0.01036731	-0.04369627	0.28189630	0.00335254
Q4	-0.00236050	0.01274669	-0.05372493	0.22364300	0.06969574
Q1	-0.00110156	0.00594845	-0.02507163	0.10436673	0.16585801
Q2	-0.00015737	0.00084978	-0.00358166	0.01490953	0.23797972
Q3	0.00047210	-0.00254934	0.01074499	-0.04472860	0.28606085
Q4	0.00078683	-0.00424890	0.01790831	-0.07454766	0.31010142

TABLE 4.114. Quarterly benchmarking weights matrix used by X-11-ARIMA and X-12-ARIMA for Table D11A.

Year	B1	D11	R
1986	1220.5	1219.53575	0.96425
1987	1252.8	1251.31218	1.48782
1988	1312.5	1311.45328	1.04672
1989	1361.2	1363.83663	-2.63663
1990	1385.3	1386.98833	-1.68833
1991	1391.1	1391.58606	-0.48606
1992	1389.2	1383.88885	5.31115
1993	1348.8	1349.59492	-0.79492
1994	1391.8	1395.72244	-3.92244

TABLE 4.115. B1 and D11 annual totals.

	Weights						
Month	1	2	3	4	5	Res.	Adj.
Jan	0.10539496	-0.02790048	0.00737759	-0.00191944	0.00038069		0.0723
Feb	0.10446930	-0.02672983	0.00706804	-0.00183890	0.00036472	0.96425	0.0726
Mar	0.10261797	-0.02438853	0.00644894	-0.00167783	0.00033277	1.48782	0.0733
Apr	0.09984099	-0.02087658	0.00552030	-0.00143622	0.00028485	1.04672	0.0743
May	0.09613833	-0.01619398	0.00428210	-0.00111408	0.00022096	-2.63663	0.0757
Jun	0.09151002	-0.01034074	0.00273435	-0.00071140	0.00014110	-1.68833	0.0774
Jul	0.08595605	-0.00331684	0.00087706	-0.00022818	0.00004526		0.0794
Aug	0.07947641	0.00487771	-0.00128979	0.00033557	-0.00006655		0.0818
Sep	0.07207110	0.01424290	-0.00376618	0.00097985	-0.00019434		0.0845
Oct	0.06374014	0.02477874	-0.00655213	0.00170467	-0.00033810		0.0875
Nov	0.05448351	0.03648524	-0.00964762	0.00251003	-0.00049783		0.0909
Dec	0.04430122	0.04936238	-0.01305266	0.00339593	-0.00067353		0.0947

TABLE 4.116. First year benchmarking weights, residuals and adjustments.

Month	Weights 1	2	3	4	5	Res.	Adj.
Jan	-0.00879152	0.04432649	0.06039494	-0.01571303	0.00311645		-0.1009
Feb	-0.00616123	0.03106468	0.07290679	-0.01805856	0.00358166		-0.1401
Mar	-0.00384049	0.01936360	0.08291627	-0.01884336	0.00373731	1.48782	-0.1741
Apr	-0.00182930	0.00922326	0.09042338	-0.01806743	0.00358341	1.04672	-0.2027
May	-0.00012766	0.00064366	0.09542812	-0.01573076	0.00311997	-2.63663	-0.2261
Jun	0.00126443	-0.00637522	0.09793049	-0.01183335	0.00234698	-1.68833	-0.2442
Jul	0.00234698	-0.01183335	0.09793049	-0.00637522	0.00126443	-0.48606	-0.2570
Aug	0.00311997	-0.01573076	0.09542812	0.00064366	-0.00012766		-0.2645
Sep	0.00358341	-0.01806743	0.09042338	0.00922326	-0.00182930		-0.2667
Oct	0.00373731	-0.01884336	0.08291627	0.01936361	-0.00384049		-0.2636
Nov	0.00358166	-0.01805856	0.07290679	0.03106468	-0.00616123		-0.2553
Dec	0.00311645	-0.01571303	0.06039494	0.04432649	-0.00879152		-0.2416

TABLE 4.117. Fourth year benchmarking weights, residuals and adjustments.

The adjustments made are calculated using the product of these two matrices, and we have, for example, for the month of February 1986:

$$\begin{aligned} FEB86 &= (0.10446930 \times 0.96425) + (-0.02672983 \times 1.48782) + \\ &\quad (0.00706804 \times 1.04672) + (-0.00183890 \times -2.63663) + \\ &\quad (0.00036472 \times -1.68833) \\ &= 0.0726. \end{aligned}$$

This leads to the adjusted value:

$$\begin{aligned} D11A(FEB86) &= D11(FEB86) + 0.0726 \\ &= 99.620 + 0.0726 \\ &= 99.693. \end{aligned}$$

For the fourth year of complete data (1989), the weight and residual matrices used are provided in Table 4.117.

For example, we have for the month of August 1989:

$$\begin{aligned} AUG89 &= (0.00311997 \times 1.48782) + (-0.01573076 \times 1.04672) + \\ &\quad (0.09542812 \times -2.63663) + (0.00064366 \times -1.68833) + \\ &\quad (-0.00012766 \times -0.48606) \\ &= -0.2645. \end{aligned}$$

This leads to the adjusted value:

$$\begin{aligned} D11A(AUG89) &= D11(AUG89) - 0.2645 \\ &= 112.595 - 0.2645 \\ &= 112.331. \end{aligned}$$

For the last year of complete data (1994), the weight and residual matrices used are provided in Table 4.118.

For example, we have for the month of December 1994:

$$DEC94 = (0.00038069 \times -1.68833) + (-0.00191944 \times -0.48606) +$$

Month	Weights 1	2	3	4	5	Res.	Adj.
Jan	-0.00067353	0.00339593	-0.01305266	0.04936238	0.04430122		-0.2828
Feb	-0.00049783	0.00251003	-0.00964762	0.03648524	0.05448351		-0.2943
Mar	-0.00033810	0.00170467	-0.00655213	0.02477874	0.06374014	-1.68833	-0.3048
Apr	-0.00019434	0.00097985	-0.00376618	0.01424290	0.07207110	-0.48606	-0.3142
May	-0.00006655	0.00033557	-0.00128979	0.00487771	0.07947641	5.31115	-0.3225
Jun	0.00004526	-0.00022818	0.00087706	-0.00331684	0.08595605	-0.79492	-0.3298
Jul	0.00014110	-0.00071140	0.00273435	-0.01034074	0.09151002	-3.92244	-0.3361
Aug	0.00022096	-0.00111408	0.00428210	-0.01619398	0.09613833		-0.3413
Sep	0.00028485	-0.00143622	0.00552030	-0.02087658	0.09984099		-0.3455
Oct	0.00033277	-0.00167783	0.00644894	-0.02438853	0.10261797		-0.3486
Nov	0.00036472	-0.00183890	0.00706804	-0.02672983	0.10446930		-0.3507
Dec	0.00038069	-0.00191944	0.00737759	-0.02790048	0.10539496		-0.3518

TABLE 4.118. Last year benchmarking weights, residuals and adjustments.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	101.106	**99.693**	98.982	106.896	100.095	101.358	102.353	99.497	102.175	103.462	102.512	102.369
1987	97.033	104.355	105.108	104.200	105.050	103.798	105.041	103.710	105.688	105.163	106.656	106.996
1988	106.432	107.445	108.141	108.168	110.335	109.574	108.756	109.749	110.986	108.391	110.415	114.109
1989	113.322	113.594	109.948	116.070	111.678	114.727	112.963	**112.331**	113.430	113.323	114.925	114.891
1990	114.556	114.095	115.355	115.133	116.144	114.723	116.752	117.499	115.774	116.487	115.444	113.337
1991	115.616	114.694	114.273	115.915	114.368	116.587	118.468	116.210	115.507	116.051	117.112	116.301
1992	116.880	117.502	116.705	116.626	116.060	114.642	115.146	116.440	116.554	116.270	114.536	111.838
1993	112.143	115.736	113.594	110.554	113.197	112.639	112.316	112.257	111.040	111.168	111.522	112.633
1994	114.167	113.750	113.436	114.504	116.395	115.690	116.267	118.278	116.498	116.053	117.109	**119.652**
1995	**118.974**	118.667	119.655	.	.	.	.	.	.	.	.	.

TABLE 4.119. D11A: Final seasonally adjusted series with revised yearly totals.

$$\begin{aligned} & (0.00737759 \times 5.31115) + (-0.02790048 \times -0.79492) + \\ & (0.10539496 \times -3.92244) \\ = \; & -0.3518. \end{aligned}$$

This leads to the following adjusted value:

$$\begin{aligned} D11A(DEC94) &= D11(DEC94) - 0.3518 \\ &= 120.004 - 0.3518 \\ &= 119.652. \end{aligned}$$

This last adjustment factor will help us estimate the values for the last incomplete year 1995. Thus we have:

$$\begin{aligned} D11A(JAN95) &= D11(JAN95) - 0.3518 \\ &= 119.326 - 0.3518 \\ &= 118.974. \end{aligned}$$

This leads to the final Table D11A.

4.3.13 Table D12: Final Estimation of the Trend-Cycle Component

Description and method of calculation

Series D1, from which extreme values and trading-day effects have been removed, is adjusted using the seasonal factors in Table D10, and we ob-

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	101.587	102.236	100.085
1986	101.034	99.620	98.909	**100.294**	100.020	101.281	102.274	102.244	102.091	103.374	102.421	102.275
1987	102.484	104.252	105.001	104.088	104.934	103.677	104.916	103.579	105.552	105.022	106.510	106.845
1988	106.275	107.287	107.985	108.020	110.198	109.453	108.654	109.671	110.936	111.120	110.433	114.168
1989	113.423	113.734	113.351	113.398	111.904	114.871	113.220	112.595	113.697	113.586	115.180	115.133
1990	114.759	114.279	115.521	115.285	116.283	114.852	116.874	117.617	115.890	116.603	115.563	113.463
1991	115.783	114.864	114.439	116.068	114.501	116.692	116.189	116.234	115.479	115.963	116.956	116.069
1992	116.566	117.119	116.267	116.147	115.553	114.120	114.624	115.931	116.072	115.827	114.147	113.279
1993	111.902	112.730	113.496	112.410	113.221	112.715	112.439	112.420	111.238	111.397	111.774	112.903
1994	114.450	114.044	113.741	114.818	116.081	116.020	116.603	116.220	116.844	116.402	117.460	120.004
1995	119.326	119.019	120.007	.	.	.	.	.	.	.	.	.

TABLE 4.120. D11bis: Modified seasonally adjusted series.

tain a modified seasonally adjusted series (Table D11bis). This series is smoothed at this stage, using a Henderson moving average, to provide a final estimate of the trend-cycle. The method is the same as that used for Table C7 (Section 4.2.6). We thus have $D11bis = D1\ op\ D10$.

- Step 1: Selecting the moving average, calculating the $\bar{I}/\bar{C}$ ratio.
- Step 2: Smoothing the SA series using a Henderson moving average.

Comment

It is possible to specify the length of the Henderson moving average to be used. X-11-ARIMA provides a choice between a 9-term, a 13-term or a 23-term moving average. X-12-ARIMA allows any odd-numbered average less than 101.

Example

First we calculate a modified SA series displayed in Table D11bis. For example

$$APR86 = 100 \times 101.706/101.408 = 100.294.$$

Step 1: Selecting the moving average, calculating the $\bar{I}/\bar{C}$ ratio.

Then Table D11bis is smoothed using a 13-term Henderson moving average whose coefficients are shown in Table 3.8. The first computable term is therefore that for April 1986, and we have:

$$\begin{aligned} APR86 \quad = \quad & 101.587 \times (-0.01935) + 102.236 \times (-0.02786) + \\ & 100.085 \times (0.00000) + 101.034 \times (0.06549) + \\ & 99.620 \times (0.14736) + 98.909 \times (0.21434) + \\ & 100.294 \times (0.24006) + 100.020 \times (0.21434) + \\ & 101.281 \times (0.14736) + 102.274 \times (0.06549) + \\ & 102.244 \times (0.00000) + 102.091 \times (-0.02786) + \end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	**99.974**	100.452	101.097	101.732	102.206	102.428	102.530	102.646	102.889
1987	103.273	103.736	104.129	104.379	104.447	104.388	104.399	104.597	104.981	105.466	105.942	106.409
1988	106.900	107.438	107.964	108.469	108.927	109.284	109.565	109.900	110.422	111.138	111.983	112.724
1989	113.206	113.457	113.517	113.439	113.287	113.193	113.251	113.435	113.720	114.050	114.367	114.672
1990	114.915	115.024	115.159	115.400	115.752	116.179	116.475	116.536	116.363	115.983	115.517	115.089
1991	114.825	114.818	115.012	115.309	115.604	115.821	115.971	116.074	116.128	116.209	116.346	116.516
1992	116.647	116.562	116.246	115.807	115.447	115.279	115.308	115.399	115.318	114.956	114.338	113.620
1993	113.033	112.734	112.717	112.815	112.839	112.665	112.313	111.950	111.784	111.883	112.219	112.753
1994	113.367	113.993	114.565	115.078	115.541	115.902	116.190	116.476	116.818	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.121. D12a: Trend-cycle (13-term Henderson ma).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	100.320	99.569	100.182	100.533	100.037	99.670	100.824	99.781	99.403
1987	99.235	100.497	100.837	99.721	100.466	99.319	100.495	99.027	100.544	99.579	100.536	100.410
1988	99.416	99.859	100.020	99.587	101.167	100.154	99.169	99.792	100.465	99.984	98.616	101.281
1989	100.192	100.244	99.854	99.964	98.779	101.483	99.973	99.260	99.980	99.593	100.711	100.402
1990	99.865	99.352	100.315	99.900	100.459	98.858	100.342	100.927	99.593	100.535	100.040	98.587
1991	100.834	100.041	99.502	100.659	99.046	100.752	100.188	100.138	99.441	99.789	100.525	99.617
1992	99.931	100.478	100.018	100.294	100.092	98.995	99.407	100.461	100.653	100.758	99.833	99.700
1993	98.999	99.996	100.691	99.641	100.338	100.045	100.112	100.420	99.512	99.565	99.603	100.134
1994	100.955	100.045	99.280	99.774	100.467	100.102	100.356	99.781	100.022	.	.	.
1995	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.122. D12b: Irregular component.

$$
\begin{aligned}
& 103.374 \times (-0.01935) \\
= \; & 99.974.
\end{aligned}
$$

At this step in the calculation, there is no attempt to estimate the six points that cannot be calculated at the beginning and end of the series. An estimate is derived for the trend-cycle (Table D12a), and one for the irregular component (Table D12b) by division using Table D11bis.

As this is a multiplicative model, the rates of growth are calculated (see Section 4.1.7).

$$
\bar{C} = \frac{2.881 + 3.484 + 5.780 + 2.288 + 2.862}{101} + \frac{1.705 + 2.944 + 2.706 + 3.550}{101}
$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1985	.	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	0.479	0.642	0.628	0.466	0.217	0.099	0.113	0.236	2.881
1987	0.374	0.448	0.379	0.240	0.065	0.057	0.010	0.190	0.367	0.462	0.452	0.440	3.484
1988	0.461	0.504	0.489	0.467	0.423	0.327	0.257	0.306	0.475	0.648	0.760	0.662	5.780
1989	0.428	0.221	0.053	0.069	0.133	0.084	0.051	0.163	0.251	0.290	0.278	0.267	2.288
1990	0.211	0.095	0.117	0.210	0.304	0.369	0.255	0.052	0.148	0.327	0.401	0.371	2.862
1991	0.229	0.007	0.169	0.259	0.255	0.188	0.130	0.089	0.046	0.070	0.118	0.146	1.705
1992	0.112	0.073	0.271	0.378	0.311	0.146	0.025	0.079	0.070	0.315	0.538	0.627	2.944
1993	0.517	0.264	0.015	0.087	0.021	0.154	0.312	0.323	0.148	0.088	0.301	0.475	2.706
1994	0.545	0.552	0.502	0.447	0.403	0.312	0.249	0.246	0.294	.	.	.	3.550
1995	.	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.123. D12c: Absolute growth rate (in %) of the trend-cycle.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1985	.	.	.	.	.	.	.	.	.	.	.	.	.
1986	.	.	.	.	0.749	0.615	0.350	0.493	0.366	1.157	1.035	0.378	5.144
1987	0.169	1.272	0.338	1.107	0.747	1.142	1.184	1.461	1.532	0.959	0.961	0.126	10.998
1988	0.990	0.446	0.161	0.433	1.587	1.001	0.984	0.628	0.675	0.478	1.368	2.702	11.453
1989	1.076	0.053	0.389	0.110	1.186	2.738	1.488	0.713	0.725	0.386	1.122	0.307	10.293
1990	0.535	0.514	0.969	0.414	0.560	1.594	1.502	0.583	1.322	0.945	0.492	1.452	10.882
1991	2.279	0.786	0.538	1.162	1.602	1.722	0.560	0.050	0.695	0.349	0.738	0.903	11.385
1992	0.315	0.547	0.458	0.276	0.201	1.096	0.416	1.060	0.192	0.104	0.918	0.134	5.716
1993	0.703	1.007	0.695	1.043	0.700	0.293	0.068	0.308	0.905	0.054	0.038	0.532	6.345
1994	0.821	0.902	0.764	0.498	0.694	0.364	0.254	0.573	0.241	.	.	.	5.110
1995	.	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4.124. D12d: Absolute growth rate (in %) of the irregular.

$$
\begin{aligned}
&= 0.2792, \\
\bar{I} &= \frac{5.144 + 10.998 + 11.453 + 10.293 + 10.882}{101} + \\
&\quad \frac{11.385 + 5.716 + 6.345 + 5.110}{101} \\
&= 0.7656, \\
\bar{I}/\bar{C} &= \frac{0.7656}{0.2792} \\
&= 2.742.
\end{aligned}
$$

Step 2: Smoothing the SA series using a Henderson moving average.

Since the ratio is greater than 1 and smaller than 3.5, we select a 13-term Henderson moving average whose coefficients and those of the associated asymmetric moving averages are shown in Table 3.8. The trend-cycle for October 1985 is estimated, from the modified seasonally adjusted series in Table D11bis, using the current point and six points in the future, to which are assigned the coefficients of the $H6\_0$ moving average in Table 3.8.

$$
\begin{aligned}
OCT85 &= 101.587 \times (0.42113) + 102.236 \times (0.35315) + \\
&\quad 100.085 \times (0.24390) + 101.034 \times (0.11977) + \\
&\quad 99.620 \times (0.01202) + 98.909 \times (-0.05811) + \\
&\quad 100.294 \times (-0.09186) \\
&= 101.634.
\end{aligned}
$$

This leads to Table D12, whose values are graphed in the second panel of Figure 4.1 with the values of the seasonally adjusted series from Table D11.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	**101.634**	101.254	100.809
1986	100.356	99.967	99.809	99.974	100.452	101.097	101.732	102.206	102.428	102.530	102.646	102.889
1987	103.273	103.736	104.129	104.379	104.447	104.388	104.399	104.597	104.981	105.466	105.942	106.409
1988	106.900	107.438	107.964	108.469	108.927	109.284	109.565	109.900	110.422	111.138	111.983	112.724
1989	113.206	113.457	113.517	113.439	113.287	113.193	113.251	113.435	113.720	114.050	114.367	114.672
1990	114.915	115.024	115.159	115.400	115.752	116.179	116.475	116.536	116.363	115.983	115.517	115.089
1991	114.825	114.818	115.012	115.309	115.604	115.821	115.971	116.074	116.128	116.209	116.346	116.516
1992	116.647	116.562	116.246	115.807	115.447	115.279	115.308	115.399	115.318	114.956	114.338	113.620
1993	113.033	112.734	112.717	112.815	112.839	112.665	112.313	111.950	111.784	111.883	112.219	112.753
1994	113.367	113.993	114.565	115.078	115.541	115.902	116.190	116.476	116.818	117.300	117.921	118.567
1995	119.144	119.619	119.961	.	.	.	.	.	.	.	.	.

TABLE 4.125. D12: Final trend-cycle, I/C ratio is 2.7420, a 13-term Henderson moving average has been selected.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	99.954	100.970	99.281
1986	100.676	99.653	99.099	**106.850**	99.569	100.182	100.533	97.269	99.670	100.824	99.781	99.403
1987	93.862	100.497	100.837	99.721	100.466	99.319	100.495	99.027	100.544	99.579	100.536	100.410
1988	99.416	99.859	100.020	99.587	101.167	100.154	99.169	99.792	100.465	97.512	98.616	101.281
1989	100.192	100.244	97.009	102.498	98.779	101.571	99.973	99.260	99.980	99.593	100.711	100.402
1990	99.865	99.352	100.315	99.900	100.459	98.858	100.342	100.927	99.593	100.535	100.040	98.587
1991	100.834	100.041	99.502	100.659	99.046	100.752	102.212	100.138	99.441	99.789	100.525	99.617
1992	99.931	100.478	100.018	100.294	100.092	98.995	99.407	100.461	100.653	100.758	99.833	98.148
1993	98.999	102.515	100.691	97.966	100.338	100.045	100.112	100.420	99.512	99.565	99.603	100.134
1994	100.955	100.045	99.280	99.774	101.018	100.102	100.356	101.840	100.022	99.235	99.609	101.212
1995	100.153	99.499	100.038	.	.	.	.	.	.	.	.	.

TABLE 4.126. D13: Final irregular series.

4.3.14 Table D13: Final Estimation of the Irregular Component

Description and method of calculation

This final estimate of the irregular component is obtained by removing the trend-cycle component (Table D12) from the estimate of the seasonally adjusted series in Table D11: $D13 = D11 \ op \ D12$.

Example

The value for April 86 is:

$$APR86 = 100 \times 106.822/99.974 = 106.850.$$

The final values of the irregular component are graphed in the bottom panel of Figure 4.1.

4.3.15 Table D16: Estimation of Various Seasonal and Calendar Effects

Description and method of calculation

This final estimate of the seasonal component and of calendar effects is obtained by removing the seasonally adjusted series (Table D11) from the

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	113.892	107.399	100.515
1986	105.509	99.077	105.046	**102.507**	97.681	102.388	97.483	66.086	103.046	113.278	105.740	102.078
1987	103.678	98.991	107.523	102.894	95.298	104.459	97.030	66.326	102.982	111.310	107.689	102.953
1988	101.341	102.715	109.922	100.074	97.461	104.794	93.140	69.298	103.303	108.791	109.840	100.466
1989	103.947	98.651	109.152	98.648	98.746	104.635	93.270	70.518	100.443	111.545	110.089	97.887
1990	105.525	98.444	106.993	100.707	99.413	101.696	95.659	70.823	98.887	113.205	109.983	97.653
1991	106.493	98.203	104.248	102.870	98.951	100.007	97.269	70.203	100.797	114.174	106.706	99.768
1992	105.949	99.813	106.651	103.317	95.021	104.013	97.798	69.007	102.781	111.373	106.968	102.049
1993	101.607	97.863	108.109	103.330	95.301	103.890	96.141	70.895	103.202	108.621	108.880	101.680
1994	101.616	97.769	109.020	100.507	97.672	104.293	93.908	71.995	103.215	108.589	108.718	99.997
1995	104.001	97.715	108.494	.	.	.	.	.	.	.	.	.

TABLE 4.127. D16: Combined seasonal, trading-day and Easter effects adjustment factors.

raw data (Table A1 or A3 if permanent adjustments have been requested): $D16 = A1 \; op \; D11$.

Example

The value for April 86 is:

$$APR86 = 100 \times 109.500/106.822 = 102.507.$$

4.3.16 Table D18: Combined Calendar Effects Factors

Description and method of calculation

Table D18 is available only in X-12-ARIMA, which has a more sophisticated calendar effects estimation procedure than X-11-ARIMA. These calendar effects can be estimated either from the irregular component via the "X11Regression" specification, or from the original series with the "Regression" specification prior to seasonal decomposition.

Table D18 shows the combined trading-day and holiday factors that are used in the final seasonal adjustment, as estimated by the "Regression" or "X11Regression" specifications.

Example

In our case, Table D18 is identical to Table C18 (Section 4.2.15). The final values of the trading-day factors are graphed in the fourth panel of Figure 4.1.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	102.198	98.646	99.895
1986	101.662	99.115	97.557	101.084	99.839	99.083	102.198	97.504	101.116	101.662	97.167	101.347
1987	99.839	99.115	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1988	97.504	102.982	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1989	99.895	99.115	101.662	97.167	101.347	101.441	97.557	102.198	98.646	99.895	101.463	97.504
1990	101.347	99.115	99.839	99.083	102.198	98.646	99.895	101.662	97.167	101.347	101.441	97.557
1991	102.198	99.115	97.504	101.116	101.662	97.167	101.347	99.839	99.083	102.198	98.646	99.895
1992	101.662	100.947	99.895	101.463	97.504	101.116	101.662	97.557	101.084	99.839	99.083	102.198
1993	97.504	99.115	101.347	101.441	97.557	101.084	99.839	99.895	101.463	97.504	101.116	101.662
1994	97.557	99.115	102.198	98.646	99.895	101.463	97.504	101.347	101.441	97.557	101.084	99.839
1995	99.895	99.115	101.662	.	.	.	.	.	.	.	.	.

TABLE 4.128. D18: Combined calendar effects factors.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	115.700	109.800	100.600
1986	106.600	98.700	103.900	**102.480**	97.700	103.700	99.700	67.544	105.200	117.100	108.300	104.400
1987	107.072	103.200	112.900	107.100	100.000	108.300	101.800	68.700	108.700	116.900	114.700	110.000
1988	107.700	110.200	118.700	108.100	107.400	114.700	101.200	76.000	114.600	120.908	121.300	114.700
1989	117.900	112.200	123.906	111.905	110.500	120.300	105.600	79.400	114.200	126.700	126.800	112.700
1990	121.100	112.500	123.600	116.100	115.600	116.800	111.800	83.300	114.600	132.000	127.100	110.800
1991	123.300	112.800	119.300	119.400	113.300	116.700	115.300	81.600	116.400	132.400	124.800	115.800
1992	123.500	116.900	124.000	120.000	109.800	118.700	112.100	80.000	119.300	129.000	122.100	113.800
1993	113.700	110.326	122.700	114.200	107.900	117.100	108.100	79.700	114.800	121.000	121.700	114.800
1994	116.300	111.500	124.000	115.400	114.000	121.000	109.500	83.857	120.600	126.400	127.700	120.000
1995	124.100	116.300	130.200	.	.	.	.	.	.	.	.	.

TABLE 4.129. E1: Original series modified for extremes with zero final weights.

4.4 PART E: Components Modified for Large Extreme Values

4.4.1 Table E1: Raw Series Modified for Large Extreme Values

Description and method of calculation

This table shows the raw series adjusted for extreme values assigned zero weight during Step C17. If a value is considered very extreme, the value of the raw series is replaced by a composite of the trend-cycle component (Table D12), the seasonality (Table D10) and, where applicable, the *a priori* adjustments and corrections for calendar effects. We therefore have: $E1 = D12 \textit{ invop } D10 \textit{ invop } C16$.

Example

For example, the value for the month of April 1986, considered very extreme, becomes

$$APR86 = 99.974 \times 1.014 \times 1.011 = 102.480.$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	101.587	102.236	100.085
1986	101.034	99.620	98.909	**99.974**	100.020	101.281	102.274	102.206	102.091	103.374	102.421	102.275
1987	103.273	104.252	105.001	104.088	104.934	103.677	104.916	103.579	105.552	105.022	106.510	106.845
1988	106.275	107.287	107.985	108.020	110.198	109.453	108.654	109.671	110.936	111.138	110.433	114.168
1989	113.423	113.734	113.517	113.439	111.904	114.971	113.220	112.595	113.697	113.586	115.180	115.133
1990	114.759	114.279	115.521	115.285	116.283	114.852	116.874	117.617	115.890	116.603	115.563	113.463
1991	115.783	114.864	114.439	116.068	114.501	116.692	118.537	116.234	115.479	115.963	116.956	116.069
1992	116.566	117.119	116.267	116.147	115.553	114.120	114.624	115.931	116.072	115.827	114.147	111.515
1993	111.902	112.734	113.496	110.520	113.221	112.715	112.439	112.420	111.238	111.397	111.774	112.903
1994	114.450	114.044	113.741	114.818	116.718	116.020	116.603	116.476	116.844	116.402	117.460	120.004
1995	119.326	119.019	120.007	.	.	.	.	.	.	.	.	.

TABLE 4.130. E2: Final seasonally adjusted series modified for extremes with zero final weights.

4.4.2 Table E2: Seasonally Adjusted Series Modified for Large Extreme Values

Description and method of calculation

This table shows the seasonally adjusted series in Table D11 corrected for extreme values assigned zero weight during Step C17. If a value is considered very extreme, the value of the seasonally adjusted series is replaced by the value of the trend-cycle component (Table D12).

Example

A very extreme point was detected in April 1986. The value for this date in Table D11 (106.822) is replaced by the corresponding value in Table D12 (99.974).

4.4.3 Table E3: Final Irregular Component Adjusted for Large Extreme Values

Description and method of calculation

This table shows the irregular component in Table D13 adjusted for extreme values assigned zero weight during Step C17. If a value is considered very extreme, the value for the component is replaced by its theoretical average (1 or 0 depending on the model).

Example

For example, the value for the month of April 1986, considered highly extreme, becomes

$$APR86 = 100.000 \text{ (i.e. } 100 \times 1.0).$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	99.954	100.970	99.281
1986	100.676	99.653	99.099	**100.000**	99.569	100.182	100.533	100.000	99.670	100.824	99.781	99.403
1987	100.000	100.497	100.837	99.721	100.466	99.319	100.495	99.027	100.544	99.579	100.536	100.410
1988	99.416	99.859	100.020	99.587	101.167	100.154	99.169	99.792	100.465	100.000	98.616	101.281
1989	100.192	100.244	100.000	100.000	98.779	101.571	99.973	99.260	99.980	99.593	100.711	100.402
1990	99.865	99.352	100.315	99.900	100.459	98.858	100.342	100.927	99.593	100.535	100.040	98.587
1991	100.834	100.041	99.502	100.659	99.046	100.752	102.212	100.138	99.441	99.789	100.525	99.617
1992	99.931	100.478	100.018	100.294	100.092	98.995	99.407	100.461	100.653	100.758	99.833	98.148
1993	98.999	100.000	100.691	97.966	100.338	100.045	100.112	100.420	99.512	99.565	99.603	100.134
1994	100.955	100.045	99.280	99.774	101.018	100.102	100.356	100.000	100.022	99.235	99.609	101.212
1995	100.153	99.499	100.038	.	.	.	.	.	.	.	.	.

TABLE 4.131. E3: Modified irregular series.

Year	A1 and D11	E1 and E2
1986	100.079	99.987
1987	**100.119**	**100.137**
1988	100.080	100.098
1989	99.807	99.832
1990	99.878	99.878
1991	99.965	99.965
1992	100.384	100.384
1993	99.941	99.945
1994	99.719	99.762

TABLE 4.132. E4: Ratios of annual totals original and adjusted series.

4.4.4 Table E4: Comparing the Annual Totals of Raw and Seasonally Adjusted Series

Description and method of calculation

This table compares the annual totals of the raw and seasonally adjusted series, using two pairs of series, i.e. the raw series A1 and the final seasonally adjusted series D11 on the one hand, and the corresponding series adjusted for extreme values (E1 and E2) on the other hand. For each of these pairs, and for each year i, we calculate for example $A1_i$ *op* $D11_i$, where, for example, $A1_i$ denotes the annual total of the A1-series in year i.

Example

For example, for the year 1987, we have: $A1_{1987} = 1252.8$, $D11_{1987} = 1251.312$, $E1_{1987} = 1259.372$, $E2_{1987} = 1257.651$, hence

$$\begin{aligned}\frac{A1_{1987}}{D11_{1987}} &= 100 \times \frac{1252.8}{1251.312} \\ &= 100.119, \\ \frac{E1_{1987}}{E2_{1987}} &= 100 \times \frac{1259.378}{1257.651} \\ &= 100.137.\end{aligned}$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	-5.099	-8.379
1986	5.964	-7.411	5.268	**5.390**	-10.776	6.141	-3.857	-34.102	60.122	11.312	-7.515	-3.601
1987	-3.736	2.687	9.399	-5.137	-6.629	8.300	-6.002	-32.515	58.224	7.544	-1.882	-4.098
1988	-2.091	2.321	7.713	-8.930	-0.648	6.797	-11.770	-24.901	50.789	2.880	2.884	-5.441
1989	2.790	-4.835	7.130	-4.576	-3.662	8.869	-12.219	-24.811	43.829	10.946	0.079	-11.120
1990	7.453	-7.102	9.867	-6.068	-0.431	1.038	-4.281	-25.492	37.575	15.183	-3.712	-12.825
1991	11.282	-8.516	5.762	0.084	-5.109	3.001	-1.200	-29.228	42.647	13.746	-5.740	-7.212
1992	6.649	-5.344	6.074	-3.226	-8.500	8.106	-5.560	-28.635	49.125	8.131	-5.349	-6.798
1993	-0.088	-0.528	8.488	-6.927	-5.517	8.526	-7.686	-26.272	44.040	5.401	0.579	-5.670
1994	1.307	-4.127	11.211	-6.935	-1.213	6.140	-9.504	-22.009	41.218	4.809	1.028	-6.030
1995	3.417	-6.285	11.952	.	.	.	.	.	.	.	.	.

TABLE 4.133. E5: Month-to-month changes in the original series.

4.4.5 Table E5: Changes in the Raw Series

Description and method of calculation

This table shows the changes in the raw series. For a monthly series, for example, the table provides the monthly growth for an additive decomposition model and the monthly rate of growth for a multiplicative model. For a given date t we therefore have: $E5_t = A1_t \; op \; A1_{t-1} - \; xbar$.

Example

Thus, the rate of growth from March 1986 to April 1986 is:

$$APR86 = 100 \times (109.5/103.9 - 1) = 5.390.$$

4.4.6 Table E6: Changes in the Final Seasonally Adjusted Series

Description and method of calculation

This table shows the changes in the final seasonally adjusted series (Table D11) calculated as described above. For a given date t, we therefore have: $E6_t = D11_t \; op \; D11_{t-1} - \; xbar$.

Example

Thus, the rate of growth from March 1986 to April 1986 is:

$$APR86 = 100 \times (106.822/98.909 - 1) = 8.000.$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	0.638	-2.104
1986	0.948	-1.400	-0.714	**8.000**	-6.368	1.261	0.980	-2.795	2.691	1.257	-0.923	-0.143
1987	-5.221	7.549	0.718	-0.869	0.813	-1.198	1.194	-1.274	1.905	-0.502	1.417	0.314
1988	-0.533	0.952	0.651	0.032	2.016	-0.677	-0.729	0.936	1.153	-2.310	1.901	3.382
1989	-0.653	0.274	-3.176	5.585	-3.757	2.741	-1.523	-0.552	0.978	-0.097	1.403	-0.041
1990	-0.324	-0.419	1.087	-0.205	0.866	-1.230	1.760	0.635	-1.468	0.615	-0.892	-1.818
1991	2.045	-0.793	-0.370	1.424	-1.351	1.913	1.581	-1.943	-0.649	0.419	0.857	-0.759
1992	0.428	0.474	-0.727	-0.103	-0.511	-1.240	0.442	1.140	0.122	-0.211	-1.450	-2.305
1993	0.346	3.278	-1.794	-2.622	2.443	-0.446	-0.245	-0.016	-1.052	0.142	0.339	1.010
1994	1.370	-0.355	-0.266	0.947	1.655	-0.598	0.503	1.728	-1.497	-0.378	0.909	2.166
1995	-0.565	-0.257	0.830	.	.	.	.	.	.	.	.	.

TABLE 4.134. E6: Month-to-month changes in the final seasonally adjusted series.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	-0.375	-0.439
1986	-0.450	-0.388	-0.158	**0.166**	0.479	0.642	0.628	0.466	0.217	0.099	0.113	0.236
1987	0.374	0.448	0.379	0.240	0.065	-0.057	0.010	0.190	0.367	0.462	0.452	0.440
1988	0.461	0.504	0.489	0.467	0.423	0.327	0.257	0.306	0.475	0.648	0.760	0.662
1989	0.428	0.221	0.053	-0.069	-0.133	-0.084	0.051	0.163	0.251	0.290	0.278	0.267
1990	0.211	0.095	0.117	0.210	0.304	0.369	0.255	0.052	-0.148	-0.327	-0.401	-0.371
1991	-0.229	-0.007	0.169	0.259	0.255	0.188	0.130	0.089	0.046	0.070	0.118	0.146
1992	0.112	-0.073	-0.271	-0.378	-0.311	-0.146	0.025	0.079	-0.070	-0.315	-0.538	-0.627
1993	-0.517	-0.264	-0.015	0.087	0.021	-0.154	-0.312	-0.323	-0.148	0.088	0.301	0.475
1994	0.545	0.552	0.502	0.447	0.403	0.312	0.249	0.246	0.294	0.412	0.529	0.548
1995	0.487	0.398	0.286	.	.	.	.	.	.	.	.	.

TABLE 4.135. E7: Month-to-month changes in the final trend-cycle series.

4.4.7 Table E7: Changes in the Final Trend-Cycle

Description and method of calculation

This table shows the changes in the final trend-cycle component (Table D12) calculated as described above. For a given date t, we therefore have: $E7_t = D12_t \; op \; D12_{t-1} - \; xbar$.

Comment

This table is produced by X-12-ARIMA only.

Example

Thus, the rate of growth from March 1986 to April 1986 is:

$$APR86 = 100 \times (99.974/99.809 - 1) = 0.166.$$

4.4.8 Table E11: Robust Estimation of the Final Seasonally Adjusted Series

Description and method of calculation

This table shows a robust estimate of the final seasonally adjusted series. It is equivalent to Table E2, except for those points considered very extreme,

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	101.587	102.236	100.085
1986	101.034	99.620	98.909	**106.994**	100.020	101.281	102.274	100.362	102.091	103.374	102.421	102.275
1987	96.701	104.252	105.001	104.088	104.934	103.677	104.916	103.579	105.552	105.022	106.510	106.845
1988	106.275	107.287	107.985	108.020	110.198	109.453	108.654	109.671	110.936	108.130	110.433	114.168
1989	113.423	113.734	109.811	116.234	111.904	114.971	113.220	112.595	113.697	113.586	115.180	115.133
1990	114.759	114.279	115.521	115.285	116.283	114.852	116.874	117.617	115.890	116.603	115.563	113.463
1991	115.783	114.864	114.439	116.068	114.501	116.692	118.537	116.234	115.479	115.963	116.956	116.069
1992	116.566	117.119	116.267	116.147	115.553	114.120	114.624	115.931	116.072	115.827	114.147	111.515
1993	111.902	115.509	113.496	110.520	113.221	112.715	112.439	112.420	111.238	111.397	111.774	112.903
1994	114.450	114.044	113.741	114.818	116.718	116.020	116.603	118.019	116.844	116.402	117.460	120.004
1995	119.326	119.019	120.007	.	.	.	.	.	.	.	.	.

TABLE 4.136. E11: Robust estimate of the seasonally adjusted series.

i.e. those which have been assigned zero weight in Table C17. The values of E2 for the corresponding time points are replaced by $D12 + (A1 - E1)$.

Comments

- This table is produced by X-12-ARIMA only.
- Table E11 is always identical to Table D11 when an additive model is used.

Example

Thus, for April 1986, we have:

$$APR86 = 99.974 + (109.500 - 102.480) = 106.994.$$

4.5 PART F: Seasonal Adjustment Quality Measures

4.5.1 Table F1: Smoothing the Seasonally Adjusted Series Using an MCD Moving Average

Description and method of calculation

This table shows a final seasonally adjusted series (Table D11) smoothed by means of a simple moving average whose order is based on the MCD (Month for Cyclical Dominance); the way MCD is computed will be explained in Table F2E.

Comments

- If the calculated MCD value is even, a centered moving average is used, i.e. a $2 \times$ MCD.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	100.912	100.912	**100.912**
1986	100.377	101.294	101.281	101.330	101.861	101.962	101.016	101.687	101.915	101.915	101.419	101.851
1987	102.177	102.510	103.042	104.391	104.523	104.239	104.532	104.549	105.116	105.502	106.041	106.388
1988	106.981	107.283	107.953	108.589	108.862	109.199	109.782	109.417	109.613	110.716	111.467	112.026
1989	112.376	113.544	113.091	113.401	113.298	113.792	113.277	113.614	113.656	114.038	114.471	114.587
1990	114.974	114.995	115.225	115.244	115.763	116.182	116.303	116.367	116.509	115.827	115.460	115.255
1991	114.822	114.924	115.131	115.313	116.047	116.406	116.289	116.581	116.634	116.140	116.207	116.535
1992	116.596	116.434	116.331	115.841	115.343	115.275	115.260	115.315	115.320	114.698	113.892	113.792
1993	113.326	112.601	112.942	113.104	112.478	112.263	112.407	112.042	111.854	111.947	112.353	112.914
1994	113.383	113.991	114.754	115.068	115.580	116.556	116.961	116.898	117.186	117.866	118.007	118.442
1995	119.163	119.163	119.163	.	.	.	.	.	.	.	.	.

TABLE 4.137. F1: Seasonally adjusted series smoothed by means of an MCD moving average, MCD=5.

- Regardless of the order of the moving average, there is no attempt to estimate the extremes of the series.
- If the value of MCD is greater than 6, it is assumed to be equal to 6. Therefore, the order of the moving average cannot be greater than 7.

Example

The MCD parameter is equal here to 5 (see Table F2E). Table F1 is therefore obtained simply, on the basis of Table D11, using a simple moving average of order 5. The first computable term is that for December 1985, and we have:

$$\begin{aligned} DEC85 &= \frac{101.587 + 102.236 + 100.085 + 101.034 + 99.620}{5} \\ &= 100.912. \end{aligned}$$

4.5.2 Table F2A: Changes, in Absolute Value, of the Principal Components

Description and method of calculation

This table shows the average absolute changes, over several periods (from 1 to 12 months for a monthly series, from 1 to 4 quarters for a quarterly series), in a few series.

Let us consider the example of the raw series in Table A1. For a given time lag d (e.g. from 1 to 12 months), we calculate:

$$\bar{O}_d = \frac{1}{n-d} \sum_{t=d+1}^{n} |A1_t \ op \ A1_{t-d} - \ xbar \,| ,$$

i.e., for a multiplicative model, the average of the absolute growth rates over d months. This calculation is carried out, for each time lag d, for the 10 following series:

Time Lag	A1 O	D11 A	D13 I	D12 C	D10 S	A2 P	C18 D	F1 MCD	E1 O^M	E2 A^M	E3 I^M
1	**11.03**	1.34	1.29	0.29	10.73	0.00	2.46	0.34	11.02	0.90	0.86
2	11.84	1.43	1.26	0.57	11.25	0.00	2.16	0.58	11.76	1.06	0.83
3	11.54	1.55	1.21	0.83	11.47	0.00	1.26	0.78	11.46	1.23	0.79
4	11.95	1.70	1.19	1.07	11.37	0.00	2.45	1.00	11.99	1.43	0.78
5	11.22	1.72	1.08	1.30	10.69	0.00	1.93	1.23	11.37	1.57	0.74
6	12.04	1.91	1.14	1.50	12.03	0.00	1.51	1.44	12.34	1.71	0.66
7	11.74	2.07	1.12	1.70	10.91	0.00	2.35	1.64	11.93	1.90	0.75
8	12.05	2.21	1.22	1.89	11.39	0.00	1.86	1.85	12.00	2.06	0.82
9	11.85	2.44	1.17	2.07	10.68	0.00	1.17	2.03	11.81	2.22	0.74
10	12.09	2.52	1.14	2.26	10.92	0.00	2.53	2.22	12.08	2.40	0.75
11	11.04	2.65	1.10	2.44	10.32	0.00	1.84	2.40	11.24	2.60	0.75
12	3.35	2.96	1.25	2.60	0.14	0.00	1.50	2.58	3.23	2.85	0.88

TABLE 4.138. F2A: Average percent change without regard to sign over the indicated span.

Table	Symbol	Series
A1 or B1	$\bar{O}_d$	Original series
D11	$\overline{A}_d$	Final seasonally adjusted series
D13	$\bar{I}_d$	Final irregular component
D12	$\bar{C}_d$	Final trend-cycle
D10	$\bar{S}_d$	Final seasonal factors
A2	$\bar{P}_d$	Preliminary adjustment factors
C18	$\overline{D}_d$	Final trading-day factors
F1	$\overline{MCD}_d$	Final seasonally adjusted series smoothed by means of an MCD moving average
E1	$\bar{O}_d^M$	Original series adjusted for extreme values
E2	$\overline{A}_d^M$	Final seasonally adjusted series corrected for extreme values
E3	$\bar{I}_d^M$	Final irregular component adjusted for extreme values

Example

As the model here is multiplicative, rates of growth are used for each time lag. The value 11.03 obtained for the mean of the absolute monthly rates of growth for the original series can be found by taking the mean of the absolute values of the data in Table E5.

4.5.3 Table F2B: Relative Contribution of Components to Changes in the Raw Series

Description and method of calculation

This table shows, for a given time lag d, the relative contribution of each component to the changes in the raw series. Assuming that the components

Time Lag	D13 I	D12 C	D10 S	A2 P	C18 T	Total	Ratio $\times 100$
1	1.36	0.07	**93.65**	0.00	4.92	100.00	**101.17**
2	1.20	0.24	95.04	0.00	3.52	100.00	94.97
3	1.08	0.51	97.23	0.00	1.17	100.00	101.72
4	1.02	0.83	93.81	0.00	4.34	100.00	96.57
5	0.97	1.39	94.56	0.00	3.08	100.00	96.01
6	0.86	1.50	96.13	0.00	1.51	100.00	103.85
7	0.97	2.23	92.49	0.00	4.30	100.00	93.30
8	1.07	2.58	93.84	0.00	2.51	100.00	95.20
9	1.12	3.54	94.20	0.00	1.14	100.00	86.32
10	0.99	3.86	90.30	0.00	4.86	100.00	90.30
11	1.03	5.08	91.00	0.00	2.89	100.00	95.97
12	14.74	63.79	0.18	0.00	21.28	100.00	94.63

TABLE 4.139. F2B: Relative contributions to the variance of the percent change in the components of the original series.

are independent, we have, at least approximately:

$$\bar{O}_d^2 \approx \bar{C}_d^2 + \bar{S}_d^2 + \bar{P}_d^2 + \overline{D}_d^2 + \bar{I}_d^2.$$

Since the two members of this equation are not exactly equal, we will in fact be calculating:

$$\overline{O'}_d^2 = \bar{C}_d^2 + \bar{S}_d^2 + \bar{P}_d^2 + \overline{D}_d^2 + \bar{I}_d^2.$$

We then calculate the ratios $100 \times \bar{C}_d^2/\overline{O'}_d^2, \ldots, 100 \times \bar{I}_d^2/\overline{O'}_d^2$ to obtain the relative contribution of each component. Finally, we calculate the ratio $100 \times \overline{O'}_d^2/\bar{O}_d^2$ to measure the quality of the approximation.

Comments

- X-12-ARIMA uses for this calculation Table E3, instead of Table D13 (used in X-11-ARIMA), as an estimate of the irregular component.
- X-12-ARIMA calculates $\overline{O'}_d^2$ from Table E1 instead of Table A1 used in X-11-ARIMA.

Example

Thus, for time lag 1, we have:

$$\begin{aligned}\overline{O'}_d^2 &= \bar{C}_d^2 + \bar{S}_d^2 + \bar{P}_d^2 + \overline{D}_d^2 + \bar{I}_d^2 \\ &= (0.29)^2 + (10.73)^2 + (0.00)^2 + (2.46)^2 + (1.29)^2 \\ &= 122.933.\end{aligned}$$

Also,

$$100 \times \frac{\overline{O'}_d^2}{\bar{O}_d^2} = 100 \times \frac{122.933}{(11.03)^2}$$

Time Lag	A1 O		D13 I		D12 C		D10 S		D11 A		F1 MCD	
	Avg	STD	Avg	STD	Avg	STD	Avg	STD	Avg	STD	Avg	STD
1	**1.38**	16.84	0.02	1.90	0.15	0.31	1.15	16.20	0.17	1.92	0.15	0.40
2	1.96	20.42	0.01	1.90	0.30	0.60	1.61	20.04	0.30	2.00	0.31	0.62
3	1.94	19.26	0.01	1.77	0.45	0.86	1.45	18.95	0.47	1.96	0.46	0.82
4	1.95	17.71	0.02	1.85	0.61	1.09	1.25	16.95	0.63	2.14	0.61	1.04
5	2.31	18.40	0.01	1.59	0.77	1.28	1.52	18.23	0.79	1.99	0.75	1.25
6	2.27	17.14	0.02	1.79	0.93	1.45	1.28	16.60	0.95	2.22	0.90	1.42
7	2.75	19.93	-0.05	1.66	1.09	1.62	1.61	18.99	1.04	2.28	1.04	1.60
8	2.35	17.87	-0.02	1.73	1.25	1.78	1.04	17.07	1.22	2.46	1.20	1.79
9	2.45	17.28	-0.02	1.86	1.40	1.95	1.02	16.52	1.37	2.69	1.34	1.97
10	2.96	19.42	-0.03	1.64	1.54	2.13	1.32	18.20	1.51	2.69	1.48	2.15
11	2.88	15.66	0.01	1.60	1.68	2.31	1.12	14.89	1.69	2.81	1.62	2.32
12	1.86	3.62	0.01	1.71	1.82	2.49	0.02	0.22	1.82	3.03	1.76	2.49

TABLE 4.140. F2C: Average percent change with regard to sign and standard deviation over the indicated span.

$$= \quad 101.17$$

and, for example, the contribution of the seasonal component is equal to: $100 \times (10.73)^2/122.933 = 93.65$.

4.5.4 Table F2C: Averages and Standard Deviations of Changes as a Function of the Time Lag

Description and method of calculation

For each time lag, the average and standard deviation of changes are calculated, this time taking into consideration the sign, for the raw series and its components, as well as for the MCD smoothed series of Table F1.

Example

Thus, the average of the monthly growth rates for the raw series (1.38) corresponds to the average of Table E5.

4.5.5 Table F2D: Average Duration of Run

Description and method of calculation

For a few series, the average duration of growth and decline phases is calculated. To do so, we consider the series of changes (growth rate for a multiplicative model, growth values for an additive model). The duration of a growth (decline) phase is the number of successive positive (negative) terms. If a term is zero, it is counted in the current phase.

The average duration of run is calculated for the following series:

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	.	-0.375	-0.439
1986	-0.450	-0.388	-0.158	0.166	0.479	0.642	0.628	0.466	0.217	0.099	0.113	0.236
1987	0.374	0.448	0.379	0.240	0.065	-0.057	0.010	0.190	0.367	0.462	0.452	0.440
1988	0.461	0.504	0.489	0.467	0.423	0.327	0.257	0.306	0.475	0.648	0.760	0.662
1989	0.428	0.221	0.053	-0.069	-0.133	-0.084	0.051	0.163	0.251	0.290	0.278	0.267
1990	0.211	0.095	0.117	0.210	0.304	0.369	0.255	0.052	-0.148	-0.327	-0.401	-0.371
1991	-0.229	-0.007	0.169	0.259	0.255	0.188	0.130	0.089	0.046	0.070	0.118	0.146
1992	0.112	-0.073	-0.271	-0.378	-0.311	-0.146	0.025	0.079	-0.070	-0.315	-0.538	-0.627
1993	-0.517	-0.264	-0.015	0.087	0.021	-0.154	-0.312	-0.323	-0.148	0.088	0.301	0.475
1994	0.545	0.552	0.502	0.447	0.403	0.312	0.249	0.246	0.294	0.412	0.529	0.548
1995	0.487	0.398	0.286	.	.	.	.	.	.	.	.	.

TABLE 4.141. F2D1: Monthly growth rates of the trend-cycle from Table D12.

D11 (A)	D13 (I)	D12 (C)	F1 (MCD)
1.6377	1.5067	8.071	3.2059

TABLE 4.142. F2D: Average duration of run.

Table	Symbol	Series
D11	A	Final seasonally adjusted series
D13	I	Final irregular component
D12	C	Final trend-cycle
F1	MCD	Final seasonally adjusted series smoothed by means of an MCD moving average

Example

Consider, for example, the trend-cycle in Table D12. The series of monthly growth rates is provided in Table F2D1. The series of growth and decline phase durations is as follows: $\{5, 14, 1, 21, 3, 14, 6, 11, 5, 2, 7, 2, 4, 18\}$ or an average duration of $113/14 = 8.071$ months. We thus get Table F2D.

4.5.6 Table F2E: Calculation of the MCD (Months for Cyclical Dominance) Ratio

Description and method of calculation

On the basis of the data in Table F2A, we calculate the $\bar{I}_d/\bar{C}_d$ ratios for each value of time lag d. The MCD value, used for example in Table F1, is the first value of the lag d after which all the ratios, including the ratio for the lag d, are smaller than 1 (i.e. $\bar{I}_d/\bar{C}_d < 1$).

Example

The different values of the ratio are easily calculated on the basis of columns D13 and D12 in Table F2A. Thus, for time lag 1, we have $\bar{I}_1/\bar{C}_1 = 1.29/0.29 = 4.46$. We then get Table F2E, and the MCD value is therefore 5.

1	2	3	4	5	6	7	8	9	10	11	12
4.46	2.22	1.45	1.11	**0.84**	0.76	0.66	0.65	0.56	0.51	0.45	0.48

TABLE 4.143. F2E: $\bar{I}/\bar{C}$ ratio for months span and MCD.

Comment

If, for instance, the value for lag 6 had been greater than 1, the MCD would have been equal to 7.

4.5.7 Table F2F: Relative Contribution of Components to the Variance of the Stationary Part of the Original Series

Description and method of calculation

To make the original series stationary, we extract a linear trend if the model is additive or an exponential trend if the model is multiplicative. This trend is not estimated from the raw series (Table A1) but from the trend-cycle component in Table D12. The latter is also made stationary by removing the same trend. Finally, the relative contribution of each component irregular D13, trend-cycle made stationary, seasonality D10, *a priori* adjustment factors, factors for trading days C18 is calculated.

Comments

- In calculating the variance of components, we use the theoretical averages (*xbar*) for the seasonal and irregular components and the trading-day factors, and empirical averages for the raw series, the trend-cycle component and *a priori* adjustment factors.

- As an estimate of the irregular component, X-12-ARIMA uses Table E3 and X-11-ARIMA uses Table D13.

Example

It would be tedious to describe each step of this fairly long calculation. Suffice it to summarize each step of the calculation. As this is a multiplicative model, the method used is the most complex.

- We fit a straight line, written *ols*, using ordinary least squares, to the logarithm of the trend-cycle component in Table D12, which corresponds to fitting an exponential trend to series D12.

- Original series $A1$ and trend-cycle $D12$ are then made stationary by removing this exponential trend from them: $A1bis = A1\ op\ \exp(ols)$ and $D12bis = D12\ op\ \exp(ols)$.

I	C	S	P	D	Total
1.09	5.36	91.50	0.00	1.91	99.86

TABLE 4.144. F2F: Relative contribution of the components to the stationary portion of the variance in the original series.

1	2	3	4	5	6	7	8	9	10	11	12	13	14
-0.15	-0.15	0.00	-0.10	0.21	0.00	0.00	-0.07	-0.26	0.05	0.08	-0.05	0.02	-0.08

TABLE 4.145. F2G: Autocorrelations of the irregular component.

- We now have a series $A1bis$ decomposed here into 4 independent components: $D12bis$, $D13$, $D10$ and $C18$. We will be transforming these variables logarithmically to revert to the sum of independent random variables, making it possible to use the approximate equality:

$$\begin{aligned}\operatorname{Var}[\log(A1bis)] &\approx \operatorname{Var}[\log(D12bis)] + \operatorname{Var}[\log(D13)] + \\ &\quad \operatorname{Var}[\log(D10)] + \operatorname{Var}[\log(C18)].\end{aligned}$$

- Each of these variances is then calculated using the empirical averages for $\log(A1bis)$ and $\log(D12bis)$, and the theoretical average (0 here since we are dealing with a logarithmic transformation) for $\log(D10), \log(C18)$ and $\log(D13)$.

- Finally, the contribution of the variance of each component to the variance of the original series made stationary is printed in Table F2F.

4.5.8 Table F2G: Autocorrelations of the Irregular Component

Description and method of calculation

This table shows the autocorrelations of the irregular component in Table D13 calculated for time lags between 1 and the number of seasonal periods +2 (i.e. 14 for a monthly series and 6 for a quarterly series). Variance and autocorrelations are calculated using the theoretical average $xbar$. We thus have, with n observations in series I:

$$\operatorname{Corr}_k(I) = \frac{\sum_{t=k+1}^{n} (I_t - xbar)(I_{t-k} - xbar)}{\sum_{t=1}^{n} (I_t - xbar)^2}$$

Example

Here, the theoretical average is 1. The autocorrelations up to order 14 are provided in Table F2G.

Final $\bar{I}/\bar{C}$ ratio from D12	2.74
Final $\bar{I}/\bar{S}$ Ratio from D10	4.60

TABLE 4.146. F2H: Final $\bar{I}/\bar{C}$ and $\bar{I}/\bar{S}$ ratios.

	Statistic	PROB > Stat (%)
F test for stable seasonality (B1)	183.698	0.000
F test for trading-day effects (C15)	68.245	0.000
F test for stable seasonality (D8)	498.194	0.000
Kruskal-Wallis test for stable seasonality (D8)	104.780	0.000
F test for moving seasonality (D8)	1.724	10.386

TABLE 4.147. F2I: Tests for the presence of seasonality.

4.5.9 Table F2H: $\bar{I}/\bar{C}$ and $\bar{I}/\bar{S}$ Ratios

Description and method of calculation

This table repeats the values for $\bar{I}/\bar{C}$ and $\bar{I}/\bar{S}$ ratios calculated in Tables D12 and D10 respectively.

4.5.10 Table F2I: Tests for the Presence of Seasonality

Description and method of calculation

This table repeats the values of tests for the presence of seasonality carried out in Tables B1 (F-test for stable seasonality) and D8 (F-test for seasonality, Kruskal-Wallis test for seasonality, F-test for moving seasonality).

Comment

X-11-ARIMA also prints the results of the test for the presence of a trading-day effect (Table C15).

4.5.11 Table F3: Monitoring and Quality Assessment Statistics

Description and method of calculation

This table provides 11 statistics used to judge the quality of seasonal adjustment. Their calculation and justification are described in detail in Lothian and Morry [49]. These statistics vary between 0 and 3, and only values smaller than 1 are considered acceptable. A composite indicator of the quality of seasonal adjustment is set up as a linear combination of these 11 statistics[22].

[22] On the interpretation of these statistics, one can refer to Baxter [5].

Statistic M1: If the part played by the irregular component in changes to the series is large, it will be difficult to identify and extract a seasonal component. Statistic M1 measures the contribution of the irregular to the total variance on the basis of results in Table F2B. Lothian and Morry [49] have shown that, for a monthly series, it is time lag 3 which provides the best comparison of the respective contributions of the irregular and seasonal components. Statistic M1 is defined, with Table F2B notations, by :

$$M1 = 10 \times \frac{\bar{I}_3^2/\bar{O}_3^2}{1 - \bar{P}_3^2/\bar{O}_3^2}.$$

This contribution, which corresponds to cell $(3, 1)$ in Table F2B, is considered acceptable if it does not exceed 10%. The computed contribution is "normalized" by dividing by 10 (or by dividing the percentage in Table F2B by 10).

Statistic M2: This statistic, similar to statistic M1, is calculated on the basis of the contribution of the irregular component to the variance of the raw series made stationary. This contribution is shown as a percentage in the first column of Table F2F. Statistic M2 is defined, from Table F2F data, by:

$$M2 = 10 \times \frac{\text{Contribution}(I)}{1 - \text{Contribution}(P)}.$$

It is considered acceptable, again, if it does not exceed 10%. Statistic M2 is normalized by multiplying by 10 (or by dividing the percentage in Table F2F by 10).

Statistic M3: To provide a correct seasonal adjustment, X-11 estimates each of the components successively. Specifically, in extracting the trend-cycle component, it is desirable that the contribution of the irregular to changes in the preliminary estimate of the seasonally adjusted series not be too large. Otherwise, it will be difficult to separate the two components. Statistic M3 measures this contribution on the basis of the $\bar{I}/\bar{C}$ ratio in Table D12, also shown in Table F2H. We have:

$$M3 = \frac{1}{2} \times \left(\frac{\bar{I}}{\bar{C}} - 1\right).$$

Statistic M4: One of the basic hypotheses which determine the validity of F-tests carried out during X-11 processing is the random nature of the irregular component. Statistic M4 tests for the presence of autocorrelation on the basis of the average duration of runs (ADR) in the irregular, shown in Table F2D, using the formula:

$$M4 = \frac{\left|\frac{N-1}{ADR} - \frac{2(N-1)}{3}\right|}{2.577 \times \sqrt{\frac{16N-29}{90}}}.$$

Statistic M5: The MCD value calculated in Table F2E shows the number of months needed for the absolute variations of the trend-cycle component to override those of the irregular component. This value, like M3, can therefore be used to compare the significance of changes in trend with that of the irregular.

The MCD is the number of months k such that: $\bar{I}_j/\bar{C}_j \leq 1$ for any $j \geq k$. The value that is used here is based on a linear interpolation:

$$MCD' = (k-1) + \frac{\frac{\bar{I}_{k-1}}{\bar{C}_{k-1}} - 1}{\frac{\bar{I}_{k-1}}{\bar{C}_{k-1}} - \frac{\bar{I}_k}{\bar{C}_k}}.$$

It is generally accepted that, for a monthly series, this value must not exceed 6, so that statistic M5 can be defined as follows:

$$M5 = \frac{MCD' - 0.5}{5}.$$

Statistic M6: To extract the seasonal component, X-11 smoothes an estimate of the seasonal-irregular component, e.g. using a 3×5 moving average. Experience has shown that when annual changes in the irregular component are too small in relation to annual changes in the seasonal component (weak I/S ratio), the 3×5 average is not flexible enough to follow the seasonal movement. Lothian [47] has shown that this 3×5 average works well for values of the I/S ratio between 1.5 and 6.5. Statistic M6 is derived from these values and from the I/S ratio in Table D10, shown again in Table F2H. We have:

$$M6 = \frac{1}{2.5} \times \left| \frac{\bar{I}}{\bar{S}} - 4 \right|.$$

Statistic M7: M7 is the combined test for the presence of identifiable seasonality discussed in Section 4.3.7. It compares, on the basis of the F-tests in Table D8, the relative contribution of stable (statistic F_S) and moving (statistic F_M) seasonality. It is used to determine whether seasonality can or cannot be identified by X-11. We have:

$$M7 = \sqrt{\frac{1}{2}\left(\frac{7}{F_S} + \frac{3F_M}{F_S}\right)}.$$

Statistics M8 to M11: The seasonal filters used by X-11 work best for constant seasonalities. If the seasonal movement changes over the years, the seasonal factors estimates could be erroneous. Two types of movement are considered: movement due to short-term quasi-random variations and movement due to more long-term changes. The significance of the first type of movement can be measured using the average of absolute annual changes (statistics M8 and M10). The simple average of annual changes, on the other hand, provides some idea of the significance of a systematic (linear) movement (statistics M9 and M11).

These last 4 statistics are calculated on the basis of normalized seasonal factors. To the seasonal component in Table D10 we apply the transformation: $S_t = (S_t - \bar{S})/\sigma(S)$, where the average used is the theoretical average *xbar*.

Using a notation similar to that of Equation (4.5) in Section 4.3.9, let $S_{i,j}$, $i = 1, \ldots, n_j$, $j = 1, \ldots, k$ denote the transformed seasonal factor for the i^{th} observation in season j, where the number of season is $k = 4$ in the quarterly case and $k = 12$ in the monthly case.

- **Statistic M8:** The magnitude of variations in the seasonal component is measured by the mean absolute variation:

$$|\overline{\Delta S}| = \frac{1}{\sum_{j=1}^{k}(n_j - 1)} \sum_{j=1}^{k} \sum_{i=2}^{n_j} |S_{i,j} - S_{i-1,j}|.$$

We then have, if the tolerance limit is 10%:

$$M8 = 100 \times |\overline{\Delta S}| \times \frac{1}{10}.$$

- **Statistic M9:** We have, if the tolerance limit is 10%:

$$M9 = \frac{10}{\sum_{j=1}^{k}(n_j - 1)} \sum_{j=1}^{k} \left|S_{n_j,j} - S_{1,j}\right|.$$

- **Statistic M10:** It is the equivalent of statistic M8 measured for the last 3 years:

$$|\overline{\Delta S}|_R = \frac{1}{3k} \sum_{j=1}^{k} \sum_{i=n_j-2}^{n_j} |S_{i,j} - S_{i-1,j}|.$$

and

$$M10 = 100 \times |\overline{\Delta S}|_R \times \frac{1}{10}.$$

- **Statistic M11:** It is the equivalent of statistic M9 measured for the last 3 years:

$$M11 = \frac{10}{3k} \sum_{j=1}^{k} \left|S_{n_j,j} - S_{n_j-2,j}\right|.$$

Statistic Q: Finally, an overall quality statistic is calculated as a linear combination of statistics M1 to M11. We have:

$$\begin{aligned} Q \;=\; & \frac{10 \times M1 + 11 \times M2 + 10 \times M3 + 8 \times M4 + 11 \times M5 + 10 \times M6}{100} + \\ & \frac{18 \times M7 + 7 \times M8 + 7 \times M9 + 4 \times M10 + 4 \times M11}{100}. \end{aligned}$$

Statistic	Value
M1	0.108
M2	0.109
M3	0.871
M4	0.029
M5	0.779
M6	0.241
M7	0.111
M8	0.126
M9	0.099
M10	0.163
M11	0.151
Q	0.270

TABLE 4.148. F3: Monitoring and quality assessment statistics.

Comments

- In computing the statistics M1 and M2, X-12-ARIMA uses Table E3 as the irregular estimate instead of Table D13 used by X-11-ARIMA.
- Statistic M6 makes sense only if a 3×5 moving average has been used. Otherwise, its weight in statistic Q is zero.
- Statistics M8 to M11 can be calculated only if the series covers at least six years. Otherwise, the weight vector used to calculate statistic Q is (14, 15, 10, 8, 11, 10, 32, 0, 0, 0, 0).
- X-12-ARIMA also produces a Q2 statistic which is the value of Q without the M2 statistic.
- For some M statistics, the value can exceed 3. In this case, the program reduces the value to 3 to prevent it from having too big an effect on statistic Q.

Example

The monitoring and quality assessment statistics for our example are provided in Table 4.148 which is an edited version of softwares' Table F3.

We have:

$$
\begin{aligned}
M1 &= \frac{1.08}{10} = 0.108, \\
M2 &= \frac{1.09}{10} = 0.109, \\
M3 &= \frac{1}{2}(2.742 - 1) = 0.871, \\
M4 &= \frac{\left| \frac{114-1}{1.5067} - \frac{2(114-1)}{3} \right|}{2.577 \times \sqrt{\frac{16\times 114-29}{90}}} = \frac{0.3350}{11.509} = 0.029,
\end{aligned}
$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1985	.	.	.	.	.	.	.	.	.	1.143	0.887	0.062
1986	0.378	-0.004	0.767	0.141	-0.216	0.333	-0.461	-3.219	0.191	1.142	0.882	0.072
1987	0.384	-0.012	0.763	0.141	-0.226	0.330	-0.455	-3.198	0.188	1.148	0.868	0.074
1988	0.393	-0.026	0.755	0.145	-0.243	0.328	-0.447	-3.160	0.183	1.151	0.866	0.063
1989	0.405	-0.047	0.736	0.152	-0.256	0.315	-0.439	-3.097	0.182	1.165	0.849	0.039
1990	0.412	-0.068	0.716	0.164	-0.272	0.309	-0.424	-3.031	0.177	1.169	0.841	0.010
1991	0.420	-0.092	0.691	0.173	-0.266	0.292	-0.402	-2.966	0.173	1.171	0.816	-0.013
1992	0.421	-0.112	0.676	0.182	-0.254	0.286	-0.380	-2.924	0.168	1.154	0.795	-0.015
1993	0.420	-0.126	0.667	0.186	-0.231	0.277	-0.370	-2.901	0.171	1.139	0.767	0.002
1994	0.416	-0.136	0.667	0.188	-0.222	0.279	-0.368	-2.894	0.175	1.130	0.755	0.016
1995	0.411	-0.141	0.671	.	.	.	.	.	.	.	.	.

TABLE 4.149. Standardized seasonal factors.

$$M5 = \frac{1}{5}\left[4 + \frac{1.14-1}{1.14-0.81} - 0.5\right] = 0.779,$$

$$M6 = \frac{1}{2.5} \times |4.602 - 4| = 0.241,$$

$$M7 = \sqrt{\frac{1}{2}\left(\frac{7 + 3 \times 1.724}{498.194}\right)} = 0.111.$$

For the other statistics to be calculated, the seasonal factors in Table D10 must be standardized. The standard deviation of this table is calculated using the formula

$$\sigma(S) = \left(\frac{1}{n}\sum_{t=1}^{n}(S_t - 1)^2\right)^{1/2} = 0.1001.$$

Table D10 is then standardized using the formula $S_t = (S_t - 1)/\sigma(S)$. Thus, the value for April 1986 is:

$$APR86 = (1.01408 - 1)/0.1001 = 0.1407.$$

The result is Table 4.149 of standardized seasonal factors.

Annual changes are derived by subtracting from line t line $t-1$ in Table 4.149. Thus, the changes in the seasonal factors between the months of April 1986 and April 1987 is close to 0. This results in Table 4.150 of the annual changes, in %, in standardized seasonal factors.

Statistic M8 is derived using the average of the absolute values in Table 4.150; it is in fact 10 times this average, i.e. 0.126.

Likewise, statistic M10 is derived using the average of the absolute values for the data in this table, from April 1992 to March 1995. M10 is equal to 10 times this average, i.e. 0.163.

To calculate M9, we could for example compute the average, multiplied by 10, of the absolute value of the monthly averages in this table. The average for the month of April would then be:

$$April = \frac{-0.00018 - 0.00380 - 0.00761 - 0.01153}{8} -$$

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1986	.	.	.	.	.	.	.	.	.	0.168	0.501	-1.003
1987	-0.612	0.867	0.406	-0.018	1.011	0.302	-0.577	-2.089	0.299	-0.636	1.377	-0.173
1988	-0.890	1.333	0.779	-0.380	1.736	0.227	-0.794	-3.897	0.425	-0.259	0.232	1.103
1989	-1.218	2.085	1.902	-0.761	1.306	1.341	-0.815	-6.237	0.141	-1.470	1.611	2.350
1990	-0.655	2.095	2.013	-1.153	1.582	0.568	-1.539	-6.638	0.514	-0.370	0.803	2.938
1991	-0.799	2.429	2.500	-0.955	-0.594	1.687	-2.166	-6.503	0.400	-0.188	2.491	2.254
1992	-0.145	2.031	1.527	-0.922	-1.193	0.581	-2.228	-4.180	0.506	1.651	2.138	0.194
1993	0.085	1.391	0.912	-0.347	-2.338	0.884	-0.971	-2.342	-0.341	1.518	2.789	-1.640
1994	0.471	0.951	-0.030	-0.249	-0.870	-0.121	-0.155	-0.688	-0.350	0.926	1.255	-1.402
1995	0.506	0.538	-0.453	.	.	.	.	.	.	.	.	.

TABLE 4.150. Annual changes in standardized seasonal factors (in %).

$$
\begin{aligned}
& \frac{-0.00955 - 0.00922 - 0.00347 - 0.00249}{8} \\
= \quad & -0.00598.
\end{aligned}
$$

Likewise, for the other months, we would have, in %:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
-0.362	1.525	1.062	**-0.598**	0.080	0.688	-1.156	-4.072	0.199	0.149	1.466	0.513

And 10 times the average of the absolute value of these data leads to the value of M9:

$$
\begin{aligned}
M9 &= 10\,(0.00362 + 0.01525 + 0.01062 + 0.00598 + 0.00080 + 0.00688) + \\
&\quad 10\,(0.01156 + 0.04072 + 0.00199 + 0.00149 + 0.01466 + 0.00513) \\
&= 0.099.
\end{aligned}
$$

For M11, the same procedure would apply, but for data from April 1990 to March 1993, leading to the value 0.151.

The final value of statistic Q is derived as follows:

$$
\begin{aligned}
Q &= \frac{1.08 + 1.199 + 8.71 + 0.233 + 8.569 + 2.45}{100} + \\
&\quad \frac{1.998 + 0.882 + 0.693 + 0.652 + 0.604}{100} \\
&= 0.27.
\end{aligned}
$$

5

Modelling of the Easter Effect

X-11-ARIMA and X-12-ARIMA propose different models for correcting the Easter effect **based on an estimate of the irregular component.** The proposed models and the methods used are sometimes quite different in the two softwares, which is why we have not integrated them directly into the seasonal adjustment example of Chapter 4.

5.1 The Easter Holiday

5.1.1 A Brief History

According to the Gospels, the resurrection of Christ occurred during the *Jewish Passover*, an annual Jewish holiday after the first full moon of Spring during which a lamb is sacrificed[1] to commemorate the exodus from Egypt. The Christians wanted to preserve the symbolic link between this sacrifice and Jesus' crucifixion, so it was decided, at the Council of Nicaea in 325, that the Christian holiday of Easter would be observed on the first Sunday

---

[1] *Then Moses called all the elders of Israel, and said to them, "Select lambs for yourselves according to your families, and kill the passover lamb. Take a bunch of hyssop and dip it in the blood which is in the basin, and touch the lintel and the two doorposts with the blood which is in the basin; and none of you shall go out of the door of his house until the morning. For the LORD will pass through to slay the Egyptians; and when he sees the blood on the lintel and on the two doorposts, the LORD will pass over the door, and will not allow the destroyer to enter your houses to slay you. You shall observe this rite as an ordinance for you and for your sons for ever."* [7, Exod.12].

after the first full moon following the Spring equinox. Unfortunately, the former Julian calendar was based on a slightly longer year than that of today, and gradually the Spring equinox crept closer to the winter months. When, in 1582, Pope Gregory XIII introduced the Gregorian calendar still used today, the Spring equinox fell in early March, and one of the main reasons for changing the calendar was doubtless to shift the Easter holiday back closer to Spring.

5.1.2 Calculation of the Dates of Easter

Determining in advance the dates of Easter has been the subject of works by famous mathematicians, and Gauss himself is the author of scholarly, and unfortunately complex algorithms. Gardner [26] cites a simple algorithm, attributable to Thomas H. O'Beirne [58], valid from 1900 to 2099 inclusive:

- Let Y be the year. Subtract 1900 from Y and let N be the difference.
- Divide N by 19; let A be the remainder of this division.
- Divide $(7A+1)$ by 19 and let B be the quotient of the division whose remainder is ignored.
- Divide $(11A+4-B)$ by 29; let M be the remainder of this division.
- Divide N by 4 and let Q be the quotient of the division whose remainder is ignored.
- Divide $(N+Q+31-M)$ by 7; let W be the remainder.
- The date of Easter is then $25-M-W$. If the result is positive, the month is April, and if negative, it is March (interpreting 0 as March 31, -1 as March 30 and so on up to -9 for March 22).

The dates of Easter, for the years 1985 to 1995, are given in Table 5.1. For example, for 1989, the algorithm gives:

$$
\begin{aligned}
Y &= 1989 \\
N &= 1989 - 1900 = 89 \\
N/19 &= 89/19 = 19 \times 4 + 13 \text{ and so } A = 13 \\
(7A+1)/19 &= 92/19 = 19 \times 4 + 16 \text{ and so } B = 4 \\
(11A+4-B)/29 &= 143/29 = 29 \times 4 + 27 \text{ and so } M = 27 \\
N/4 &= 89/4 = 22 \times 4 + 1 \text{ and so } Q = 22 \\
(N+Q+31-M)/7 &= 115/7 = 16 \times 7 + 3 \text{ and so } W = 3 \\
25 - M - W &= -5
\end{aligned}
$$

Since the result is negative, Easter is in March, and it is on the 26th $(31-5)$.

Year	Date of Easter
1985	April 7
1986	March 30
1987	April 19
1988	April 3
1989	March 26
1990	March 15
1991	March 31
1992	April 19
1993	April 11
1994	April 3
1995	April 16

TABLE 5.1. Date of Easter for the years 1985 to 1995.

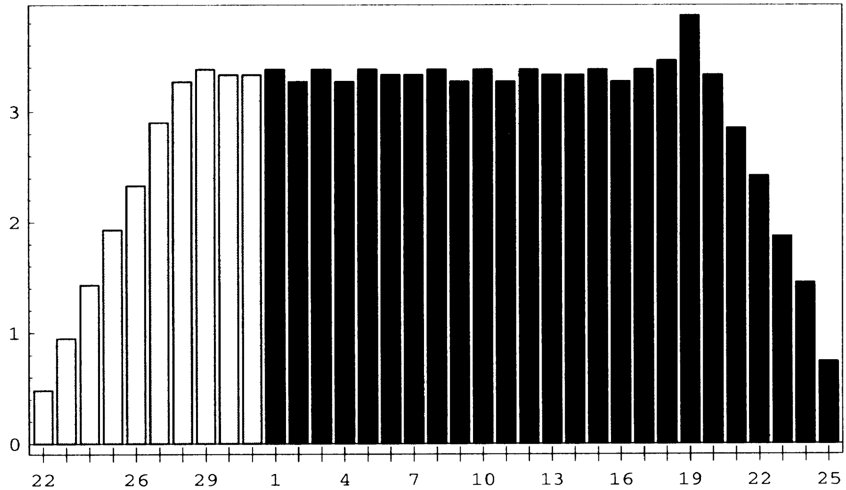


FIGURE 5.1. Distribution of dates of Easter from March 22 to April 25 over a complete cycle of 5,700,000 years.

Other algorithms, some of them more general, can be found in Montes [54] and Tøndering [65]. The sequence of Easter dates repeats itself every 5,700,000 years in the Gregorian calendar and the distribution of the dates of Easter over a complete cycle is illustrated in Figure 5.1. It can be seen that Easter Sunday more often falls in April (76.7%) than in March (23.3%).

5.1.3 *Easter and Seasonal Adjustment*

Why worry about the Easter holiday when it comes to analysing time series and seasonal adjustment? Quite simply because Easter causes a change in the level of activity in many sectors: the statutory holidays of Good Friday and Easter Monday, a change in eating habits (chocolate, lamb, etc.), and so on. Also, Easter Sunday can fall no earlier than March 22 (the last time it did was in 1818, and the next time will be in 2285), and no later than

April 25 (it last did in 1943 and will next in 2038); that is, it falls in either the first or second quarter. The potential effects of this holiday therefore cannot be wholly accounted for by series seasonality.

The models developed to take into account the Easter effect depend on the very nature of the series.

- 1997 is the last year in which Easter Sunday fell in March (on the 30th), and it had not fallen in March since 1991, when it was on March 31. For monthly sales of motor vehicles, for example, the observed variation between March and April 1997 is not comparable to those observed between the same months in previous years. The closing of stores on Good Friday resulted in an unusual drop in sales in March 1997. Inversely, a relatively larger number of marriages was registered in March 1997. For these series, the impact of Easter, positive or negative, is *immediate* in the sense that it is concentrated in the month in which Easter falls.

- For other series, the Easter effect can be felt not only during the holiday itself, but also in the preceding weeks. This *gradual* effect is seen, for example, in sales of chocolates, flowers, and so on. In this case, the observed effect depends not only on the fact that Easter falls in March or April, but also on the date on which it falls in April. The effect on the figures for March will therefore be all the greater the earlier Easter falls in the month of April.

The effects of Easter may be estimated either from the raw data, using regression models with ARIMA errors, or from a preliminary estimate of the irregular component obtained after eliminating the other effects present in the series (trend, seasonality, trading-day effect, etc.). The first approach is offered only by X-12-ARIMA, in its module Reg-ARIMA (Findley et al. [23]), and the second is available in both X-11-ARIMA and X-12-ARIMA software; it is on the models used for this second approach that we will focus for now. We will single out six such models:

- The 3 models proposed by X-11-ARIMA, which estimate the Easter effect using the irregular component from Table D13: the *Immediate Impact* model, the *Corrected Immediate Impact* model and the *Gradual Impact* model.

- The 3 models proposed by X-12-ARIMA: the *Bateman-Mayes model*, which estimates the Easter effect based on Table D13, and two *two gradual impact* models (*"Sceaster"* and *"Easter"*), which estimate the Easter effect using the irregular component estimates from Tables B13 and C13, as appropriate, along with the trading-day effect.

The irregular components from Table B13 (Table 4.35, Section 4.1.12) and Table D13 (Table 4.126, Section 4.3.14) will be used to illustrate the different models.

5.2 The X-11-ARIMA Models

As we have seen, Easter usually falls in April and, in the following models, it is this situation which is considered the norm: if an Easter effect is detected, only the data for the years in which Easter affects March will be corrected. This correction will affect the months of March and April of the years concerned and, as it is desirable not to modify the series' level, the sum of the correction coefficients will be set equal to 0 for an additive model, or to 2 for a multiplicative model.

Finally, we should point out that in working on the estimates of the irregular component of Table D13, it is a *residual* Easter effect that is estimated.

5.2.1 The Immediate Impact Model

The *Immediate Impact* model is the simplest version of the six models described in this chapter.

Model and Estimation of Effects

- Let $I_{i,j}$ be the value of the irregular component from Table D13 corresponding to month j of year i.
- $I_{i,4}$ and $I_{i,3}$, $(i = 1, \ldots, N)$ are thus the values of the irregular component for the months of April and March for the N available years. Their differences[2] are written $Y_i = I_{i,4} - I_{i,3}$.
- Let Z_i be the number of days between Easter Sunday in year i and March 22 (the earliest date for this holiday) and let $X_i = f(Z_i)$ be the function defined by:

$$X_i = f(Z_i) = \begin{cases} 1 & \text{if } Z_i \leq 9 \quad \text{(Easter falls in March)} \\ 0 & \text{if } Z_i > 9 \quad \text{(Easter falls in April)} \end{cases}$$

  While Z_i does not intervene explicitly in the calculations of the immediate effect model, the variable Z_i and the function $X_i = f(Z_i)$ facilitate the comparison of the different models.

The Easter effect may be obtained by explaining the values of the irregular component by the variable X_i, that is, by measuring the impact on the irregular component of the fact that Easter falls in March. We therefore have the following models:

$$\begin{aligned} I_{i,3} &= a_M + b_M X_i + \eta_i \\ I_{i,4} &= a_A + b_A X_i + \xi_i \end{aligned}$$

[2] In this Chapter, the usual notation in regression analysis is used.

where b_M and b_A measure the impact of Easter on the values of the months of March and April, and where η_i and ξ_i are the regression error terms. To preserve the series' level, b_A and b_m must be such that $b_A = -b_M = b/2$ and, subtracting these models, we have:

$$Y_i = I_{i,4} - I_{i,3} = a + bX_i + \epsilon_i.$$

- The parameter a measures the average impact of the Easter holiday on the difference of the irregulars. It therefore expresses a structural difference between the months of March and April (or the first and second quarters). Since, in principle, the data in Table D13 show neither trend nor seasonality, the parameter a must theoretically be null and its estimate close to zero.

- The parameter b is the supplementary effect due to Easter falling in March. It is therefore this quantity that measures the specific impact of the Easter holiday.

Let $\hat{b}$ be the estimator of b. Then, with preservation of the series' level, the correction coefficients are:

	Additive model	Multiplicative model
Month of March	$-\hat{b} \times X_i/2$	$1 - \hat{b} \times X_i/2$
Month of April	$\hat{b} \times X_i/2$	$1 + \hat{b} \times X_i/2$

Thus the data from the years in which Easter falls in March, when $X_i = f(Z_i)$ is not equal to 0, will therefore be corrected.

The estimator $\hat{b}$ of b, obtained by ordinary least squares, can be explicitly calculated[3].

- Let N_M and N_A be the number of years in which Easter falls in March or in April respectively. Of course, we have $N = N_M + N_A$.

- Let YM and YA be the sums of the values Y_i for the years in which Easter falls in March and April, respectively.

Then, it is easily shown that:

$$\hat{b} = \frac{YM}{N_M} - \frac{YA}{N_A} = \bar{Y}_M - \bar{Y}_A$$

and

$$\hat{a} = \frac{YA}{N_A} = \bar{Y}_A.$$

[3]From the usual OLS formula: $\hat{b} = \mathrm{Cov}(X,Y)/\mathrm{Var}(X)$ and $\hat{a} = \bar{Y} - \hat{b}\bar{X}$ with $\mathrm{Var}(X) = \frac{1}{n}\sum_i \sum_j X_{i,j}^2 - \bar{X}^2$ and $\mathrm{Cov}(X,Y) = \frac{1}{n}\sum_i \sum_j X_{i,j}Y_{i,j} - \bar{X}\bar{Y}$.

Year	Date of Easter	Z_i	$X_i = f(Z_i)$	$I_{i,3}$	$I_{i,4}$	Y_i
1986	March 30	8	1	0.99099	1.06850	0.07751
1987	April 19	28	0	1.00837	0.99721	-0.01116
1988	April 3	12	0	1.00020	0.99587	-0.00433
1989	March 26	4	1	0.97009	1.02498	0.05489
1990	April 15	24	0	1.00315	0.99900	-0.00415
1991	March 31	9	1	0.99502	1.00659	0.01156
1992	April 19	28	0	1.00018	1.00294	0.00276
1993	April 11	20	0	1.00691	0.97966	-0.02726
1994	April 3	12	0	0.99280	0.99774	0.00494

TABLE 5.2. Easter effect, *Immediate Impact* model data.

	SS	DF	RMSE	F	PROB>F
Easter Effect	0.0059	1	0.0059	14.2262	0.0070
Error	0.0029	7	0.0004		
Total	0.0089				

TABLE 5.3. Easter effect, F-Test for the *Immediate Impact* model.

Clearly $\hat{a}$ should be close to zero. In fact, as the case in which Easter falls in April corresponds to the "normal" situation, the Easter effect is, for these years, included in the seasonality and therefore does not affect the values of the irregular for the months of March and April in these years. The average of these values is therefore, in theory, equal to 0 or to 1 depending on the decomposition model, and their difference thus averages to zero. Therefore, we also have

$$\hat{b} = \bar{Y}_M - \bar{Y}_A \approx \bar{Y}_M.$$

Finally, the presence of an Easter effect can be tested with an F-test.

Example

Only the data from 1986 to 1994 allow for the calculation of the difference of the irregulars of March and April (shown in Table 5.2). From this table, it is easily inferred that

$$\begin{aligned}
\hat{b} &= \frac{0.07751 + 0.05489 + 0.01156}{3} - \\
&\quad \frac{-0.01116 - 0.00433 - 0.00415 + 0.00276 - 0.02726 + 0.00494}{6} \\
&= 0.04799 + 0.00653 \\
&= 0.05452, \\
\hat{a} &= -0.00653.
\end{aligned}$$

The results of the analysis of variance and of the F-test for this regression are displayed in Table 5.3. The Easter effect is therefore deemed significant at the 1% level.

Year	Mar	Apr
1985	.	.
1986	97.274	102.726
1987	100.000	100.000
1988	100.000	100.000
1989	97.274	102.726
1990	100.000	100.000
1991	97.274	102.726
1992	100.000	100.000
1993	100.000	100.000
1994	100.000	100.000
1995	100.000	.

TABLE 5.4. A11: Easter effect, *Immediate Impact* model of X-11-ARIMA, values for the months of March and April (100 otherwise).

The correction factors are zero for the years in which Easter falls in April and, the decomposition being multiplicative, for the years in which it falls in March they are:

- $1 + 0.05452/2 = 1.02726$ for the data for the month of April;
- $1 - 0.05452/2 = 0.97274$ for the data for the month of March.

This results, in X-11-ARIMA, in Table 5.4.

5.2.2 The Corrected Immediate Impact Model

X-11-ARIMA incorporates a *Corrected Immediate Impact* model, developed at the Australian Bureau of Statistics (Laker [44, 45]), which is only slightly different from the previous model by taking into account cases where the Easter weekend straddles March and April, that is, when Easter Sunday falls on March 31, April 1 or 2. In such cases, it is assumed that both months are affected by the Easter holiday with an effect taken to be half the effect observed when the entire Easter weekend falls in March.

Model

The model is therefore written as $Y_i = a + bX_i + \epsilon_i$ with:

- $I_{i,j}$ $(i = 1, \ldots, N)$ being the value of the irregular component from Table D13 corresponding to month j of year i, and $I_{i,4}$ and $I_{i,3}$ being the values of the irregular component for the months of April and March, for the N available years, and $Y_i = I_{i,4} - I_{i,3}$ their differences.
- Z_i being the number of days between Easter Sunday in year i and March 22 (the earliest date for this holiday), and $X_i = f(Z_i)$ the variable defined by

$$X_i = f(Z_i) = \begin{cases} 1 & \text{if } Z_i \leq 8 & \text{(Easter falls in March)} \\ 0.5 & \text{if } 9 \leq Z_i \leq 11 & \text{(Easter straddles March and April)} \\ 0 & \text{if } Z_i \geq 12 & \text{(Easter falls in April).} \end{cases}$$

The value of $\hat{b}$ may still be explicitly calculated.

- Let N_M, N_{MA} and N_A be the number of years in which Easter falls in March, straddles March and April, and falls in April, respectively. Of course, we have $N = N_M + N_{MA} + N_A$.

- Let YM, YMA and YA be the sums of values Y_i for the years in which Easter falls in March, straddles March and April and falls in April, respectively; and let $YT = YM + YMA + YA$.

We then have:

$$\hat{b} = 2 \times \frac{YM + 0.5 \times YMA - r \times YT}{2 \times N \times r \times (1 - r) - 0.5 \times N_{MA}}$$

with

$$r = \frac{N_M + 0.5 \times N_{MA}}{N}$$

Here again, and as done in the X-11-ARIMA program, the presence of an Easter effect may be tested with an F-test.

Estimation of Effects for an Additive Model

The natural estimates for the corrections are the same as before for an additive decomposition model:

	Additive model
Month of March	$-\hat{b} \times X_i/2$
Month of April	$\hat{b} \times X_i/2$

Thus, for a year in which Easter straddles March and April, the correction will be half that for years in which Easter falls in March.

Estimation of Effects for a Multiplicative Model

In the case of a multiplicative decomposition model, the program imposes an additional correction.

For the years in which Easter falls in April, and for those years only, let us write as AIA the average of the irregular for April ($I_{i,4}$) and as AIM the average of the irregular for March ($I_{i,3}$). Therefore, the corrections are:

	Multiplicative model
Month of March	$1 - \hat{b} \times X_i/(2 \times AIM)$
Month of April	$1 + \hat{b} \times X_i/(2 \times AIA)$

In the theoretical presentation of the model, Laker [44, 45] estimates the Easter effect from an estimate of the seasonal-irregular component. He therefore posits, in a way similar to what we have seen thus far:

$$\begin{aligned} SI_{i,3} &= \overline{SI}_{.3} - bX_i + \eta_i \\ SI_{i,4} &= \overline{SI}_{.4} + bX_i + \xi_i \end{aligned}$$

or, taking differences,

$$(SI_{i,4} - SI_{i,3}) = \left(\overline{SI}_{.4} - \overline{SI}_{.3}\right) + 2bX_i + \epsilon_i$$

where:

- $SI_{i,3}$ and $SI_{i,4}$ designate the values of the seasonal-irregular component for the months of March and April,
- b is the Easter effect,
- $\overline{SI}_{.3}$ is the theoretical average of the seasonal-irregular component for a "normal" month of March, that is, one unaffected by an Easter effect,
- and $\overline{SI}_{.4}$ is the theoretical average of the seasonal-irregular component for a "normal" month of April, that is, one affected by an Easter effect.

Under these conditions, if $2\hat{b}$ is the estimate resulting from the difference model, we will have, in the case of a multiplicative decomposition, and, for example, for the month of March:

$$SI_{i,3} \approx \overline{SI}_{.3} - \hat{b}X_i = \overline{SI}_{.3}\left(1 - \frac{\hat{b}X_i}{\overline{SI}_{.3}}\right)$$

or

$$\overline{SI}_{.3} = \frac{SI_{i,3}}{\left(1 - \frac{\hat{b}X_i}{\overline{SI}_{.3}}\right)}$$

and so the ratio

$$\left(1 - \frac{\hat{b}X_i}{\overline{SI}_{.3}}\right)$$

makes it possible to move from a value of the seasonal-irregular component affected by the Easter effect to a corrected value of this effect. Unfortunately, the quantity $\overline{SI}_{.3}$ is unknown, and Laker proposes estimating it by the average of the values of the seasonal-irregular component of the months of March of those years in which Easter falls in April (therefore, "normal" months of March).

Year	Date of Easter	Z_i	$f(Z_i)$	$I_{i,3}$	$I_{i,4}$	Y_i
1986	March 30	8	1	0.99099	1.06850	0.07751
1987	April 19	28	0	1.00837	0.99721	-0.01116
1988	April 3	12	0	1.00020	0.99587	-0.00433
1989	March 26	4	1	0.97009	1.02498	0.05489
1990	April 15	24	0	1.00315	0.99900	-0.00415
1991	March 31	9	0.5	0.99502	1.00659	0.01156
1992	April 19	28	0	1.00018	1.00294	0.00276
1993	April 11	20	0	1.00691	0.97966	-0.02726
1994	April 3	12	0	0.99280	0.99774	0.00494

TABLE 5.5. Easter effect, *Corrected Immediate Impact* model data.

A correction in the same vein is proposed for the values of the seasonal-irregular component for the months of April.

In our case, as the estimation of the Easter effect is based on estimates of the irregular component, Laker proposes adopting the same correction principle, which produces the formulas in the preceding section.

This correction proves superfluous here in that we know the theoretical average of the irregular component, here equal to 1, and it need not be estimated.

Example

In order to obtain the output of this example, the following code can be submitted to X-11-ARIMA:

```
DATA ipi    12 85 10 ;
   (the data)
;
TITLE ipi;
RANGE 12 85  10  95  3  ;
SA (ipi, 0 ,1) TDR 2 EASTER 1 CHART 1 PRTDEC 3 PRINT 5;
END;
```

The data are the same as in the example of Section 5.2.1 and only the function $f(Z_i)$ changes for the year 1991, when, Easter Sunday falling on March 31, the Easter weekend straddles March and April (see Table 5.5).

We therefore have:

$$\begin{aligned}
r &= \frac{N_M + 0.5 \times N_{MA}}{N} \\
&= \frac{2 + 0.5 \times 1}{9} \\
&= 0.27778 \\
\hat{b} &= 2 \times \frac{YM + 0.5 \times YMA - r \times YT}{2 \times N \times r \times (1-r) - 0.5 \times N_{MA}} \\
&= 2 \times \frac{0.13240 + 0.5 \times 0.01156 - 0.27778 \times 0.104762}{2 \times 9 \times 0.27778 \times (1 - 0.27778) - 0.5 \times 1}
\end{aligned}$$

	SS	DF	RMSE	F	PROB>F
Easter Effect	0.0076	1	0.0076	43.8393	0.0003
Error	0.0012	7	0.0002		
Total	0.0089				

TABLE 5.6. Easter effect, F-Test for the *Corrected Immediate Impact* model.

$$= \quad 0.07012.$$

The results of the analysis of variance and of the F-test for this regression are displayed in Table 5.6. The Easter effect is therefore deemed very significant.

As our decomposition is multiplicative, to calculate the correction factors it is necessary to know the averages of the irregular for the years in which Easter falls in April. We have:

$$\begin{aligned} AIM &= \frac{1.00837 + 1.00020 + 1.00315 + 1.00018 + 1.00691 + 0.99280}{6} \\ &= 1.001935, \end{aligned}$$

$$\begin{aligned} AIA &= \frac{0.99721 + 0.99587 + 0.99900 + 1.00294 + 0.97966 + 0.99774}{6} \\ &= 0.995403. \end{aligned}$$

And the correction factors are:

	Easter in March
Month of March	$1 - \hat{b}/(2 \times AIM) = 0.9650$
Month of April	$1 + \hat{b}/(2 \times AIA) = 1.0352$

	Easter straddling March and April
Month of March	$1 - 0.5 \times \hat{b}/(2 \times AIM) = 0.9825$
Month of April	$1 - 0.5 \times \hat{b}/(2 \times AIA) = 1.0176$

This results, in X-11-ARIMA, in Table 5.7.

5.2.3 The Gradual Impact Model

The Easter holiday can also have an effect on the days leading up to it. Thus, it being customary to give and consume chocolate or offer flowers during this period, the related industries will adapt their production, when the time comes, in order to be able to meet the demand.

Year	Mar	Apr
1985	.	.
1986	96.501	103.522
1987	100.000	100.000
1988	100.000	100.000
1989	96.501	103.522
1990	100.000	100.000
1991	98.250	101.761
1992	100.000	100.000
1993	100.000	100.000
1994	100.000	100.000
1995	100.000	.

TABLE 5.7. A11: Easter effect, *Corrected Immediate Impact* model of X-11-ARIMA, values for the months of March and April (100 otherwise).

Model and Estimation of Effects

With the *Gradual Impact* model, it is assumed that the effect varies linearly during k days (k being able to assume the values 1 to 9) leading up to Easter Sunday.

- Let $I_{i,3}$ and $I_{i,4}$, $(i = 1, \dots, N)$, be the values of the irregular component of the months of March and April from Table D13 for the N available years, and let $Y_i = I_{i,4} - I_{i,3}$ be their differences.
- Let Z_i be the number of days between Easter Sunday of year i and March 22 (earliest date for this holiday), and let $X_i = f(Z_i)$ be the function defined by:

$$X_i = f(Z_i) = \begin{cases} 1 & \text{if } Z_i \leq 9 & \text{(Easter falls in March)} \\ \frac{k+9-Z_i}{k} & \text{if } 9 < Z_i < k+9 & \text{(Easter falls in April,} \\ & & \text{ before April } k) \\ 0 & \text{if } Z_i \geq 9+k & \text{(Easter falls in April,} \\ & & \text{ on or after April } k) \end{cases}$$

The Easter effect is then estimated, by ordinary least squares, based on the model $Y_i = a + bX_i + \epsilon_i$. **With this model, then, the data for the years in which Easter falls between April 1 and April k are not used.**

The value of $\hat{b}$ may still be explicitly calculated, using the results of the *Immediate Impact* model estimated in the years for which $X_i = f(Z_i) = 0$ (or $X_i = f(Z_i) = 1$). Writing:

- N_M and N_{LA} for the numbers of years in which Easter falls in March and "late" April, respectively,
- YM and YLA for the sums of the values Y_i for the years in which Easter falls in March and "late" April, respectively,

we immediately have

$$\hat{b} = \frac{YM}{N_M} - \frac{YLA}{N_{LA}} = \bar{Y}_M - \bar{Y}_{LA},$$

and

$$\hat{a} = \bar{Y}_{LA}.$$

The correction factors are:

	Additive model	Multiplicative model
Month of March	$-\hat{b} \times X_i/2$	$1 - \hat{b} \times X_i/2$
Month of April	$+\hat{b} \times X_i/2$	$1 + \hat{b} \times X_i/2$

Comments

- The user can ask the X-11-ARIMA program to choose automatically the "optimum" value of k from among the possible values 1 to 9, i.e. the value of k producing the smallest mean-square error.
- The user can ask the program to exclude extreme values before estimating the model. Only the values corresponding to the years in which Easter is in April are affected: there are generally too few years in which Easter falls in March to identify correctly the extreme values in these cases.

  1. The standard deviation of the values Y_i is calculated for the years in which Easter falls on or after April k:

  $$\sigma_{LA} = \left(\frac{1}{N_{LA} - 1} \sum_{i \in LA} \left(Y_i - \bar{Y}_{LA}\right)^2 \right)^{1/2} \qquad (5.1)$$

  2. And the values such that $\left|Y_i - \bar{Y}_{LA}\right| > 2\sigma_{LA}$ are excluded.

- This model is very close to the *Sceaster* model used in X-12-ARIMA (see Section 5.3.2) apart from the fact that the estimation is done differently.

Example

By way of example, we will assume that $k = 5$. The data are provided in Table 5.8. The output can be obtained by running the following code:

```
DATA ipi    12 85 10 ;
   (the data)
;
```

Year	Date of Easter	Z_i	$f(Z_i)$	$I_{i,3}$	$I_{i,4}$	Y_i
1986	March 30	8	1	0.99099	1.06850	0.07751
1987	April 19	28	0	1.00837	0.99721	-0.01116
1988	April 3	12	0.4	1.00020	0.99587	-0.00433
1989	March 26	4	1	0.97009	1.02498	0.05489
1990	April 15	24	0	1.00315	0.99900	-0.00415
1991	March 31	9	1	0.99502	1.00659	0.01156
1992	April 19	28	0	1.00018	1.00294	0.00276
1993	April 11	20	0	1.00691	0.97966	-0.02726
1994	April 3	12	0.4	0.99280	0.99774	0.00494

TABLE 5.8. Easter effect, *Gradual Impact* model data ($k = 5$).

```
TITLE ipi;
RANGE 12 85  10  95  3  ;
SA (ipi, 0 ,1) TDR 2 EASTER 4 BUILDUP 5 EASTXM 0
CHART 1 PRTDEC 3 PRINT 5;
END;
```

The years in which Easter falls on April 1, 2, 3 or 4 are therefore given a weight different from 0 and 1. This is the case for years 1988 and 1994, in which Easter was on April 3 ($Z_i = 12$). The value of X_i for these years is:

$$X_i = f(Z_i) = \frac{k + 9 - Z_i}{k} = \frac{5 + 9 - 12}{5} = \frac{2}{5} = 0.4.$$

The regression is performed on the data for which $X_i = f(Z_i) = 1$ (Easter in March) or $X_i = f(Z_i) = 0$ (Easter after April 4), or 7 years. We have:

$$\begin{aligned}\hat{b} &= \frac{0.07751 + 0.05489 + 0.01156}{3} - \\ &\quad \frac{-0.01116 - 0.00415 + 0.00276 - 0.02726}{4} \\ &= 0.05794,\end{aligned}$$

and $\hat{a}$ is equal to the average of the differences for the years in which Easter falls after April 4, and so

$$\begin{aligned}\hat{a} &= \frac{-0.01116 - 0.00415 + 0.00276 - 0.02726}{4} \\ &= -0.009953.\end{aligned}$$

The results of the analysis of variance and the F-test for this regression are displayed in Table 5.9. The Easter effect is therefore deemed significant.
Our decomposition being multiplicative, the correction factors are:

	SS	DF	RMSE	F	PROB>F
Easter Effect	0.0058	1	0.0058	10.4940	0.0230
Error	0.0027	5	0.0005		
Total	0.0085				

TABLE 5.9. Easter effect, F-Test for the *Gradual Impact* model ($k = 5$).

Year	Mar	Apr
1985	.	.
1986	97.103	102.897
1987	100.000	100.000
1988	98.841	101.159
1989	97.103	102.897
1990	100.000	100.000
1991	97.103	102.897
1992	100.000	100.000
1993	100.000	100.000
1994	98.841	101.159
1995	100.000	.

TABLE 5.10. A11: Easter effect, *Gradual Impact* model ($k = 5$) of X-11-ARIMA, values for the months of March and April (100 otherwise).

	Easter in March (86, 89, 91)
Month of March	$1 - \hat{b}/2 = 0.9710$
Month of April	$1 + \hat{b}/2 = 1.0289$

	Easter on April 3 (88, 94)
Month of March	$1 - 0.4 \times \hat{b}/2 = 0.9884$
Month of April	$1 + 0.4 \times \hat{b}/2 = 1.0116$

For a value of k equal to 5, we therefore obtain Table 5.10.

The estimated values of the differences $\hat{Y}_i$ are therefore given by $\hat{Y}_i = \hat{a} + \hat{b} \times X_i$. The prediction errors can be calculated for the years in which Easter falls in April. They are displayed in Table 5.11. The average of the squares of these errors for $k = 5$ equals 0.0001455.

"Optimum" choice of the duration k

Year	Y_i	$X_i = f(Z_i)$	$\hat{b}X_i$	$\hat{Y}_i = \hat{a} + \hat{b}X_i$	$Y_i - \hat{Y}_i$
1987	-0.01116	0	0	-0.00995	-0.001206
1988	-0.00433	0.4	0.02318	0.01322	-0.017553
1990	-0.00415	0	0	-0.00995	0.005800
1992	0.00276	0	0	-0.00995	0.012709
1993	-0.02726	0	0	-0.00995	-0.017303
1994	0.00494	0.4	0.02318	0.01322	-0.008284

TABLE 5.11. Easter effect, *Gradual Impact* model ($k = 5$), prediction errors.

k	1	2	3	4	5	6	7	8	9
Error	0.01132	0.01132	0.01132	**0.00958**	**0.01455**	0.02065	0.02639	0.03144	0.03581

TABLE 5.12. Easter effect, *Gradual Impact* model, quadratic errors for various values of k ($\times 100$).

Y_i	$f(Z_i)$	Y_i^2	$\lvert Y_i - \bar{Y}_{LA}\rvert$
-0.011159	0	0.0001245	0.0012063
-0.004152	0	0.0000172	0.0058004
0.002757	0	0.0000076	0.0127093
-0.027256	0	0.0007429	0.0173034
$\bar{Y}_{LA} = -0.009953$		$\sum Y_i^2 = 0.0008922$	

TABLE 5.13. Easter effect, *Gradual Impact* model, extreme values detection.

If the program were asked to choose the optimum value of k, it would take the value minimizing this quadratic error. In our example, we would get the values displayed in Table 5.12. The minimum value occurs when $k = 4$.

For k equal to 1, 2 or 3, it can be seen that for the years studied (1986 to 1994), the variables $X_i = f(Z_i)$ are the same, which explains the equality of the first 3 quadratic errors.

Detection of extreme values

As we saw earlier, $\hat{a} = \bar{Y}_{LA} = -0.009953$ and the standard deviation of the values $\lvert Y_i - \bar{Y}_{LA}\rvert$, displayed in Table 5.13, for the years in which Easter falls on or after April 5 (i.e. $k = 5$) is calculated, using Equation (5.1):

$$\begin{aligned} \sigma_{LA} &= \left(\frac{1}{4-1}\left[0.0008922 - 4 \times (-0.009953)^2\right]\right)^{1/2} \\ &= 0.01286. \end{aligned}$$

No value of Y_i are more than 2 standard deviations (or 0.0257) away, in absolute values, from the average, and so no point is excluded from the regression.

5.3 The X-12-ARIMA Models

X-12-ARIMA proposes different Easter effect correction models than those of X-11-ARIMA. The *Sceaster* and *Easter* models evaluate the Easter effect based on estimates of the irregular component from Tables B13 and C13. The *Bateman-Mayes* [4] model uses the data from Table D13.

The *Sceaster* model is very close to the *Gradual Impact* model proposed by X-11-ARIMA, discussed earlier (see Section 5.2.3). The *Sceaster* model, like those of X-11-ARIMA, considers that the normal situation is that in which Easter falls in April; so, it corrects only the data for the years in which Easter affects March. This is not the case with the *Easter* model and the *Bateman-Mayes* model which, as will be seen in the following sections, correct the data for the months of March, April, and, in some situations, even February.

5.3.1 The Bateman-Mayes Model

The *Bateman-Mayes* model is used only in the case of a multiplicative decomposition model.

Model and Estimation

The Easter effect is estimated, based on the values of the irregular component from Table D13 for the months of March and April, in several steps:

- Let $I_{i,3}$ and $I_{i,4}$, $(i = 1, \ldots, N)$, be the values of the irregular component from Table D13 for the N available years. The values of the months of April are transformed and it is the variables $2 - I_{i,4}$ and $I_{i,3}$ that will subsequently be used.

- Let k be the number of days between Easter Sunday for a given year and March 22.

- The years are distributed among 4 groups defined in relation to k:
  $G_1 : 0 \leq k \leq 10$ from March 22 to April 1
  $G_2 : 11 \leq k \leq 17$ from April 2 to April 8
  $G_3 : 18 \leq k \leq 24$ from April 9 to April 15
  $G_4 : 25 \leq k \leq 34$ from April 16 to April 25.

Next, trimmed means of the combined values $2 - I_{i,4}$ and $I_{i,3}$ are calculated for each group.

- The trimmed mean m_1 is the average of the values of group G_1 that are less than two standard deviations away from the simple average of group G_1. The trimmed mean m_4 for group G_4 is calculated in the same way.

- The trimmed means m_2 and m_3 are calculated in several steps:

  1. Let $\bar{m}_2$ and $\bar{m}_3$ be the simple averages of groups G_2 and G_3 respectively.
  2. Preliminary correction factors $E_t(k)$ for the Easter effect on the month of March corresponding to observation t are calculated

as follows:

$$E_t(k) = \begin{cases} m_1 + \frac{(\bar{m}_2 - m_1)(k-10)}{4} & \text{if } 11 \leq k \leq 14 \\ \bar{m}_2 + \frac{(\bar{m}_3 - \bar{m}_2)(k-14)}{7} & \text{if } 15 \leq k \leq 20 \\ \bar{m}_3 + \frac{(m_4 - \bar{m}_3)(k-21)}{4} & \text{if } 21 \leq k \leq 24. \end{cases}$$

3. The preliminary correction factor for the month of April (observation $t+1$) is obtained as $E_{t+1}(k) = 2 - E_t(k)$.
4. Calculate the average quadratic errors of the preliminary correction factors for $11 \leq k \leq 17$ and $18 \leq k \leq 24$.
5. The trimmed means m_2 and m_3 are the averages of the values of groups G_2 and G_3 that are less than two standard deviations away from the values of the preliminary factors assigned to them.

- The Easter effect for a month of March is finally calculated thus:

$$E_t(k) = \begin{cases} m_1 & \text{if } 0 \leq k \leq 10 \\ m_1 + \frac{(m_2 - m_1)(k-10)}{4} & \text{if } 11 \leq k \leq 14 \\ m_2 + \frac{(m_3 - m_2)(k-14)}{7} & \text{if } 15 \leq k \leq 20 \\ m_3 + \frac{(m_4 - m_3)(k-21)}{4} & \text{if } 21 \leq k \leq 24 \\ m_4 & \text{if } 25 \leq k \leq 34 \end{cases}$$

For the month of April $t+1$, $E_{t+1}(k)$ is obtained as $2 - E_t(k)$, and for all other months, the correction factor is 1.

The function which associates the Easter effect for a month of March with k is therefore a piecewise linear function.

- These correction factors must then be adjusted to take into account the distribution of Easter insofar as certain dates are more frequent than others (see Figure 5.1). To do this, we calculate the quantity:

$$\bar{E} = \sum_{k=0}^{34} w(k) E_t(k),$$

where $w(k)$ is the proportion of years in which Easter falls on the k^{th} day after March 22. The values of $w(k)$, calculated for the period 1583-1982, used by X-12-ARIMA are given in Table 5.14.

The final estimates of the correction factors for the Easter effect are therefore:

$$\tilde{E}_t(k) = \begin{cases} E_t(k)/\bar{E} & \text{if } t \text{ is a month of March} \\ E_t(k)/(2 - \bar{E}) & \text{if } t \text{ is a month of April} \\ 1 & \text{otherwise.} \end{cases}$$

k	0	1	2	3	4	5	6
w(k)	0.0100	0.0150	0.0050	0.0175	0.0300	0.0325	0.0250
k	7	8	9	10	11	12	13
w(k)	0.0300	0.0300	0.0400	0.0375	0.0350	0.0250	0.0275
k	14	15	16	17	18	19	20
w(k)	0.0425	0.0425	0.0275	0.0300	0.0225	0.0400	0.0425
k	21	22	23	24	25	26	27
w(k)	0.0325	0.0300	0.0350	0.0300	0.0425	0.0375	0.0350
k	28	29	30	31	32	33	34
w(k)	0.0300	0.0250	0.0350	0.0300	0.0100	0.0100	0.0100

TABLE 5.14. Easter effect, proportion of years in which Easter falls on the k^{th} day after March 22; calculated for the period 1583-1982.

Raw data			Transformed data		
Year	March	April	k	March	April
1986	99.099	106.850	8 (G1)	0.99099	0.93150
1987	100.837	99.721	28 (G4)	1.00837	**1.00279**
1988	100.020	99.587	12 (G2)	1.00020	1.00413
1989	97.009	102.498	4 (G1)	0.97009	0.97502
1990	100.315	99.900	24 (G3)	1.00315	1.00100
1991	99.502	100.659	9 (G1)	0.99502	0.99341
1992	100.018	100.294	28 (G4)	1.00018	0.99706
1993	100.691	97.966	20 (G3)	1.00691	1.02034
1994	99.280	99.774	12 (G2)	0.99280	1.00226
1995	100.038	.	25 (G4)	1.00038	.

TABLE 5.15. Easter effect, *Bateman-Mayes* model data.

Example

To obtain the output of this example the following code can be submitted:

```
series{ data=(115.7 109.8 .... 130.2)
        start= 1985.10
        period= 12
        print=none
        decimals=3}
X11 {mode=mult
     print=(all)
     x11easter=yes
     print1stpass=yes }
X11regression { variables=td
                print=(all)}
```

The raw data (from Table D13) and the transformed data are in Table 5.15. Thus, for example, the transformed data point for April 1987 is

$$APR87 = 2 - 0.99721 = 1.00279.$$

From the values of k are inferred the subsets G_1, G_2, G_3 and G_4:

$$\begin{aligned} G_1 &= \{1986, 1989, 1991\}, \\ G_2 &= \{1988, 1994\}, \\ G_3 &= \{1990, 1993\}, \\ G_4 &= \{1987, 1992, 1995\}. \end{aligned}$$

They are also shown in Table 5.15.

Calculation of m_1.
The data for G_1 are:

k	March	April
8	0.99099	0.93150
4	0.97009	0.97502
9	0.99502	0.99341

The average $\bar{m}_1$ and the standard deviation σ_1 of these six values are: $\bar{m}_1 = 0.97601$ and $\sigma_1 = 0.02201$. The matrix of the absolute differences from $\bar{m}_1$ is:

March	April
0.01498	0.04451
0.00591	0.00098
0.01902	0.01741

And only the first April value (0.93150) is more than two standard deviations from the average $\bar{m}_1$ (deviation of 0.04451) and must therefore be eliminated from the final calculation. We therefore have:

$$\begin{aligned} m_1 &= \frac{0.99099 + 0.97009 + 0.97502 + 0.99502 + 0.99341}{5} \\ &= 0.98491. \end{aligned}$$

Calculation of m_4.
The data for G_4 are:

k	March	April
28	1.00837	1.00279
28	1.00018	0.99706
25	1.00038	.

The average $\bar{m}_4$ and the standard deviation σ_4 of these five values are: $\bar{m}_4 = 1.00176$ and $\sigma_4 = 0.003774$. The matrix of the absolute differences from $\bar{m}_4$ is:

March	April
0.00661	0.00103
0.00158	0.00469
0.00136	.

And here, no point being deemed extreme, the value of m_4 is 1.00176.

Calculation of m_2 and m_3.
The data for G_2 and G_3 are:

G_2			G_3		
k	March	2-April	k	March	2-April
12	1.00020	1.00413	24	1.00315	1.00100
12	0.99280	1.00226	20	1.00691	1.02034

The simple averages of these two groups of 4 values are $\bar{m}_2 = 0.99985$ and $\bar{m}_3 = 1.00785$ respectively. The preliminary estimates of the effects for March and April are thus:

G_2			G_3		
k	March	April	k	March	April
12	0.99238	1.00762	24	1.00328	0.99672
12	0.99238	1.00762	20	1.00671	0.99329

Thus, for $k = 12$ (in G_2),

$$\begin{aligned} E_t(12) &= m_1 + \frac{(\bar{m}_2 - m_1)(k-10)}{4} \\ &= 0.98491 + \frac{(0.99985 - 0.98491)(12-10)}{4} \\ &= 0.99238 \\ E_{t+1}(12) &= 2 - 0.99238 \\ &= 1.00762. \end{aligned}$$

And similarly, for G3 we have:

$$\begin{aligned} E_t(20) &= \bar{m}_2 + \frac{(\bar{m}_3 - \bar{m}_2)(k-14)}{7} \\ &= 0.99985 + \frac{(1.00785 - 0.99985)(20-14)}{7} \\ &= 1.00671 \\ E_{t+1}(20) &= 2 - 1.00671 \\ &= 0.99329 \end{aligned}$$

and

$$\begin{aligned} E_t(24) &= \bar{m}_3 + \frac{(m_4 - \bar{m}_3)(k-21)}{4} \\ &= 1.00785 + \frac{(1.00176 - 1.00785)(24-21)}{4} \\ &= 1.00328 \\ E_{t+1}(24) &= 2 - 1.00328 \\ &= 0.99672. \end{aligned}$$

The matrices of the absolute deviations for the preliminary estimates of the effects are therefore:

G_2			G_3		
k	March	April	k	March	April
12	0.00782	0.00349	24	0.00013	0.00428
12	0.00043	0.00537	20	0.00020	0.02705

The mean square errors are as follow. For I_2 $(11 \leq k \leq 17)$:

$$\begin{aligned} EQM_2 &= \left(\frac{1}{4}\left[(0.00782)^2 + (0.00349)^2 + (0.00043)^2 + (0.00537)^2\right]\right)^{1/2} \\ &= 0.00506. \end{aligned}$$

For I_3 $(18 \leq k \leq 24)$:

$$\begin{aligned} EQM_3 &= \left(\frac{1}{4}\left[(0.00013)^2 + (0.00428)^2 + (0.00020)^2 + (0.02705)^2\right]\right)^{1/2} \\ &= 0.01369. \end{aligned}$$

No point is therefore deemed extreme, and the averages for each group are the simple averages: $m_2 = 0.99985$ and $m_3 = 1.00785$.

Calculation of Easter effects.
These effects are calculated as follows:

1. For the data of G_1, the effects of March are equal to m_1 (0.98491) and those of April are equal to $2 - m_1$.

2. For the data of G_4, the effects of March are equal to m_4 (1.00176) and those of April are equal to $2 - m_4$.

3. For the data of G_2 and G_3, the effects are equal to the preliminary effects calculated earlier, as no extreme point has been detected.

The Easter effects can be calculated for any value of k from the averages m_1, m_2, m_3 and m_4; the curve of these effects is shown in Figure 5.2.
The adjustment for the distribution of Easter can then be calculated:

$$\bar{E} = \sum_{k=0}^{34} w(k)E_t(k) = 0.99717,$$

and the effects adjusted:

Year	Mar	Apr
1985	.	.
1986	98.770	101.223
1987	100.460	99.543
1988	99.519	100.478
1989	98.770	101.223
1990	100.613	99.391
1991	98.770	101.223
1992	100.460	99.543
1993	100.957	99.049
1994	99.519	100.478
1995	100.460	.

TABLE 5.16. H1: Easter effect, *Bateman-Mayes* model of X-12-ARIMA, values for the months of March and April (100 otherwise).

	Easter effects		Adjusted Easter effects	
Year	March	April	March	April
1986	98.491	101.509	**98.770**	**101.223**
1987	100.176	99.824	100.460	99.543
1988	99.238	100.762	99.519	100.478
1989	98.491	101.509	98.770	101.223
1990	100.328	99.672	100.613	99.391
1991	98.491	101.509	98.770	101.223
1992	100.176	99.824	100.460	99.543
1993	100.671	99.329	100.957	99.049
1994	99.238	100.762	99.519	100.478
1995	100.176	99.824	100.460	99.543

With, for example:

$$MAR86 = 98.491/0.99717 = 98.770$$

and

$$APR86 = 101.5091/(2 - 0.99717) = 101.223.$$

Which produces the final result displayed in Table 5.16.

5.3.2 The Sceaster Model

The *Sceaster* model is used in the more general context of the estimation of calendar effects proposed by the X-11 method at the end of Parts B and C, and is applied to the estimates of the irregular component from Tables B13 and C13.

Model and Estimation

As with the *Gradual Impact* model proposed in X-11-ARIMA, see Section 5.2.3, the Easter holiday is assumed to have an impact on the w days ($1 \leq w \leq 24$) leading up to Easter Sunday, and the model $I_{i,j} = a + bX_{i,j}(w) + \epsilon_{i,j}$ is posited where:

- $I_{i,j}$ is the value of the irregular (from Table B13 or C13) corresponding to year i and period (month or quarter) j.

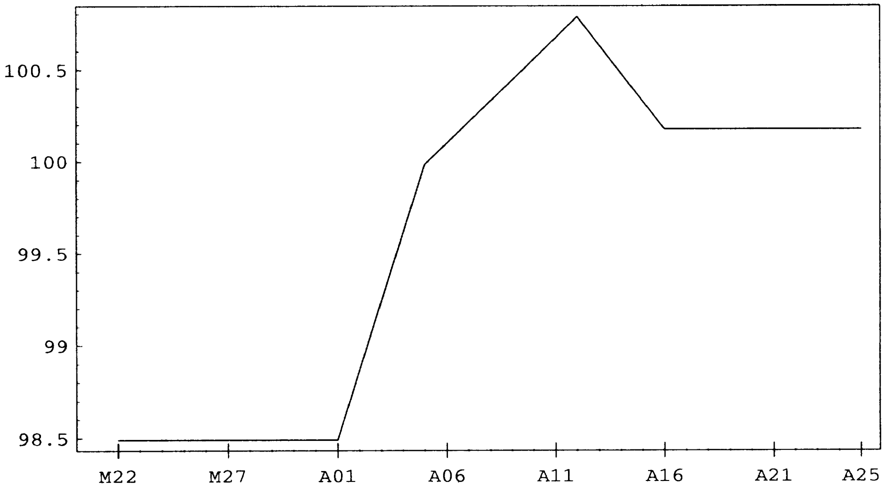


FIGURE 5.2. *Bateman-Mayes* Easter effect according to the date of Easter (March 22 to April 25).

- For a given year i, of the w days before or on Easter, let w_i be the number of days that fall in March (or in the first quarter). Then let:

$$X_{i,j}(w) = \begin{cases} w_i/w & \text{in March or the First quarter } (j = 3 \text{ or } j = 1) \\ -w_i/w & \text{in April or the Second quarter } (j = 4 \text{ or } j = 2) \\ 0 & \text{otherwise.} \end{cases}$$

The constraint imposed on w ($1 \leq w \leq 24$) means that only the values of the regressor for March and April (or of the first and second quarter) are not zero[4]. The values of a and b are estimated by ordinary least squares.

At this stage, several comments may be made.

- First of all, if Easter falls in March, the associated value $X_{i,3}(w)$ is equal to 1. Similarly, if Easter falls after April w, this value $X_{i,3}(w)$ is equal to 0. More generally, if we consider the values of variable X for the months of March, going back to the notations associated with the X-11-ARIMA models we have: $X_{i,3}(w) = f(Z_i)$.

- Given the simple form of the explanatory variable, the estimator $\hat{b}$ can be given a more explicit form.

  The only non-zero values of variable X are those for the months of March and April that are, moreover, opposites. If it is assumed that the sequence of $X_{i,j}$ does not begin in April and does not end

[4]In fact, this is not quite true since, for example, in the year 2008 Easter will fall on March 23, and if $w = 24$, there will then be one day in February.

in March, for every month of March there will be a corresponding month of April and the sum of the two values of variable X of these months will be zero. The average $\bar{X}$ is therefore also null in this case, and in the general case, nearly 0.

Assuming that the series is made up of n observations, we will therefore have:

$$\hat{b} = \mathrm{Cov}(X, I)/\mathrm{Var}(X)$$

with

$$\mathrm{Var}(X) = \frac{1}{n}\sum_i\sum_j X_{i,j}^2 - \bar{X}^2 = \frac{1}{n}\sum_i\sum_j X_{i,j}^2 = \frac{2}{n}\sum_i \left(\frac{w_i}{w}\right)^2$$

and

$$\mathrm{Cov}(X, I) = \frac{1}{n}\sum_i\sum_j X_{i,j}I_{i,j} - \bar{X}\bar{I} = \frac{1}{n}\sum_i \frac{w_i}{w}\left(I_{i,3} - I_{i,4}\right)$$

Hence we see the values of the differences between the irregulars of March and April appear.

- Finally, as $\hat{a} = \bar{I} - \hat{b}\bar{X}$ and the average $\bar{X}$ is close to 0, $\hat{a}$ is very close to $\bar{I}$, the average of the irregular component, (so it is close to its theoretical average of 0 for an additive decomposition, or 1 for a multiplicative decomposition).

The Easter effect is estimated by $\hat{E}_{i,j} = a + \hat{b}X_{i,j}(w)$, ensuring that Easter has no effect beyond the months of March and April.

The model proposed by X-12-ARIMA therefore closely resembles the *Gradual Impact* model of X-11-ARIMA[5], and the *Immediate Impact* model for $w = 1$, but here the estimation of the regression model is performed using all available years, not just those years in which Easter falls in March or after April w. Furthermore, in X-12-ARIMA, the estimation is performed a first time using the data from Table B13, and a second time using the data from Table C13.

We should recall, finally, that at this stage of the operation, in X-12-ARIMA, it is also possible to look for other calendar effects in the irregular component.

Example

If we use this model in X-12-ARIMA, the estimation is performed a first time using the data from Table B13, and then a second time using the data from Table C13.

The output for this example was obtained with the following code:

[5]That is why this model is referred to in X-12-ARIMA as "Sceaster", the SC standing for Statistics Canada.

Year	Easter	5 days before	Number of days in March (w_i)	$X_{i,3}$	$X_{i,4}$
1985	April 7	April 3	0	0	0
1986	March 30	March 26	5	1	-1
1987	April 19	April 15	0	0	0
1988	April 3	March 30	2	0.4	-0.4
1989	March 26	March 22	5	1	-1
1990	April 15	April 11	0	0	0
1991	March 31	March 27	5	1	-1
1992	April 19	April 15	0	0	0
1993	April 11	April 7	0	0	0
1994	April 3	March 30	2	0.4	-0.4
1995	April 16	April 12	0	0	0

TABLE 5.17. Easter effect, *Sceaster* model data. $X_{i,j} = 0$ whenever $j \neq 3$ or 4.

```
series{ data=(100.535 100.471 ..... 100.227)
        start= 1985.10
        period= 12
        print=none
        decimals=3}
X11 {mode=mult
     print=(all) }
X11regression { variables=sceaster[5]
                print=(all)}
```

Here, the values of the dependent variable I are the 114 values of Table B13 (divided by 100). Refer to Table 4.35 for the data.

For this example, we assume that $w = 5$. The values of the variable X (zero everywhere except for the months of March and April) are provided in Table 5.17.

Thus, in 1988, in the 5 days leading up to Easter (including Easter), 2 were in March (March 30 and 31). Hence, the explanatory variable X takes the values $2/5 = 0.4$ in March 1988, $-2/5 = -0.4$ in April 1988, and zero for all the other months in 1988.

We have $\bar{Y} = 0.9999$ and $\bar{X} = 0$, which results in $\hat{a} = \bar{Y} - \hat{b}\bar{X} = 0.9999$. Also,

$$\begin{aligned}
\mathrm{Var}(X) &= \frac{2}{n}\sum_i \left(\frac{w_i}{w}\right)^2 \\
&= \frac{2}{114}\left(1 + 0.4^2 + 1 + 1 + 0.4^2\right) \\
&= 2 \times 3.32/114, \\
\mathrm{Cov}(X, I) &= \frac{1}{n}\sum_i \frac{w_i}{w}\left(I_{i,3} - I_{i,4}\right) \\
&= \frac{1}{114}\frac{1}{100}\left[(95.390 - 107.358) + 0.4 \times (101.498 - 98.202)\right] + \\
&\quad \frac{1}{114}\frac{1}{100}\left[(98.753 - 99.913) + (97.167 - 101.999)\right] +
\end{aligned}$$

	Regression Coefficient	Standard Deviation	T-value	Prob > t
Intercept	0.99989	0.00183	547.89	0.000
Sceaster[5]	-0.02387	0.00756	-3.16	0.001

	SS	DF	RMSE	F-Value	Prob > F
Easter Effect	0.0038	1	0.0038	9.9642	0.0021
Error	0.0425	112	0.0004		
Total	0.0463	113			

TABLE 5.18. Easter effect, F-Test for the *Sceaster* model ($k = 5$).

Year	Mar	Apr
1985	.	.
1986	97.613	102.387
1987	100.000	100.000
1988	**99.045**	100.955
1989	97.613	102.387
1990	100.000	100.000
1991	97.613	102.387
1992	100.000	100.000
1993	100.000	100.000
1994	99.045	100.955
1995	100.000	.

TABLE 5.19. B16II: Preliminary estimates of the Easter effect with the *Sceaster* model of X-12-ARIMA. Values for the months of March and April (100 otherwise).

$$
\begin{aligned}
& \frac{1}{114}\frac{1}{100}\left[0.4 \times (100.309 - 98.327)\right] \\
= \quad & \frac{-0.15849}{114},
\end{aligned}
$$

and so $\hat{b} = -0.15849/6.64 = -0.02387$. The Easter effect is therefore estimated by $\hat{E}_{i,j} = 1 + \hat{b}X_{i,j}(w) = 1 - 0.02387 \times X_{i,j}(w)$, which produces Table 5.19, with, for example, for March 1988:

$$\hat{E}_{1988,3} = 1 - 0.02387 \times 0.4 = 0.99045.$$

The results of the analysis of variance and of the F-test for this regression, that are only partially output by X-12-ARIMA, are displayed in Table 5.18. The Easter effect is therefore deemed significant at the 1% level.

5.3.3 *The Easter Model*

The *Easter* model, like the previous one, falls within the more general context of the estimation of calendar effects proposed by the X-11 method at the end of Parts B and C and is applied to the estimates of the irregular component from Tables B13 and C13.

Model and Estimation

Here again, the Easter holiday is assumed to have an impact on the w $(1 \leq w \leq 25)$ days leading up to Easter Sunday, and the model $I_{i,j} = a + bX_{i,j}(w) + \epsilon_{i,j}$ is posited where:

- $I_{i,j}$ is the value of the irregular (from Table B13 or C13) corresponding to year i and period (month or quarter) j.

- For a given year i, let $w_{i,j}$ be the number of days, of the w days before Easter (**excluding Easter**), that fall in month (or quarter) j. We first define the variable $Z_{i,j}(w) = w_{i,j}/w$. Given the constraint on w, this variable is nil except for the months of February, March and April. However, it has a certain seasonality: the values pertaining to the month of February will, for example, be structurally weaker. The variable $X_{i,j}(w)$ is obtained by subtracting from $Z_{i,j}(w)$ its average $\bar{Z}_{.,j}(w)$ for month j, calculated for the available years. We therefore have: $X_{i,j}(w) = w_{i,j}/w - \bar{Z}_{.,j}(w)$. This preserves the series level by cancelling the Easter effect for the set of months concerned. The explanatory variable therefore has a zero average and no seasonality.

A long-term mean for the Easter regression variable, calculated over the period 1583-1982 (see Section 5.3.1), can be used instead of the average $\bar{Z}_{.,j}(w)$.

The constraint on the value of w $(1 \leq w \leq 25)$ means that corrections can be made only to the months of February, March and April. As previously, the values of a and b are estimated by ordinary least squares, and the Easter effect is estimated by $\hat{E}_{i,j} = a + \hat{b}X_{i,j}(w)$.

Example

If we use this model in X-12-ARIMA, the estimation is performed a first time using the data from Table B13, and a second time using the data from Table C13.

The output for this example was obtained with the following code:

```
series{ data=(100.535 100.471 ..... 100.227)
        start= 1985.10
        period= 12
        print=none
        decimals=3}
X11 {mode=mult
     print=(all) }
X11regression { variables=easter[5]
                eastermeans=no
                print=(all)}
```

Year	Easter	5 days before	# of days in March ($w_{i,3}$)	# of days in April ($w_{i,4}$)	$Z_{i,3}$	$Z_{i,4}$	$X_{i,3}$	$X_{i,4}$
1985	April 7	April 2	0	5	0	1	.	.
1986	March 30	March 25	5	0	1	0	0.6182	-0.6182
1987	April 19	April 14	0	5	0	1	-0.3818	0.3818
1988	April 3	March 29	3	2	0.6	0.4	0.2182	-0.2182
1989	March 26	March 21	5	0	1	0	0.6182	-0.6182
1990	April 15	April 10	0	5	0	1	-0.3818	0.3818
1991	March 31	March 26	5	0	1	0	0.6182	-0.6182
1992	April 19	April 14	0	5	0	1	-0.3818	0.3818
1993	April 11	April 6	0	5	0	1	-0.3818	0.3818
1994	April 3	March 29	3	2	0.6	0.4	0.2182	-0.2182
1995	April 16	April 11	0	0	0	1	-0.3818	.

TABLE 5.20. Easter effect, *Easter* model data. $Z_{i,j}$ and $X_{i,j}$ are zero whenever $j \neq 3$ or 4.

Here, the values of the dependent variable I are therefore the 114 values of Table B13 (divided by 100). Refer to Table 4.35 for the data.

For this example, we assume that $w = 5$. The values of the variable Z (zero everywhere except for the months of March and April) are provided in Table 5.20.

Thus, in 1988, in the 5-day period before Easter, there were 3 days in March (March 29, 30 and 31) and 2 in April (April 1 and 2).

The averages of $Z_{i,3}$ and $Z_{i,4}$ are calculated:

$$\begin{aligned} \bar{Z}_{.,,3} &= (1 + 0.6 + 1 + 1 + 0.6)/11 = 0.38182 \\ \bar{Z}_{.,,4} &= (1 + 1 + 0.4 + 1 + 1 + 1 + 0.4 + 1)/11 = 0.61818. \end{aligned}$$

By correcting the variable Z with these monthly averages, we get the explanatory variable X, displayed in Table 5.20.

In the most general case, the values of the regressor for the months of February, March and April cannot be zero so there is no simple expression for the estimator of b. But, as in the previous case, it will involve only the differences between the values of the irregular for the months of March and April. Using OLS, we get $\hat{a} = 0.99980$ and $\hat{b} = -0.02639$.

The Easter effect is therefore estimated by $\hat{E}_{i,j} = 1 - 0.02639 \times X_{i,j}(w)$, which produces Table 5.22, with, for example, for March 1988:

$$\hat{E}_{1988,3} = 1 - 0.02639 \times 0.2182 = 0.99424.$$

The results of the analysis of variance and of the F-test for this regression, that are only partially output by X-12-ARIMA, are displayed in Table 5.21. The Easter effect is therefore deemed significant at the 1% level.

	Regression Coefficient	Standard Deviation	T-value	Prob > t	
Intercept	0.99980	0.00185	540.55	0.000	
Easter[5]	-0.02639	0.01014	-2.60	0.005	

	SS	DF	RMSE	F-Value	Prob > F
Easter Effect	0.0026	1	0.0026	6.7772	0.0105
Error	0.0437	112	0.0004		
Total	0.0463	113			

TABLE 5.21. Easter effect, F-Test for the *Easter* model ($k = 5$).

Year	Mar	Apr
1985	.	.
1986	98.369	101.631
1987	101.008	98.992
1988	**99.424**	100.576
1989	98.369	101.631
1990	101.008	98.992
1991	98.369	101.631
1992	101.008	98.992
1993	101.008	98.992
1994	99.424	100.576
1995	101.008	.

TABLE 5.22. B16II: Preliminary estimate of the Easter effect with the *Easter* model of X-12-ARIMA. Values for the months of March and April (100 otherwise).

References

[1] Akaike, H. (1980), Seasonal Adjustment by a Bayesian Modeling, *Journal of Time Series Analysis*, *1*, 1-13.

[2] Armatte, M. (1992), Conjonctions, conjoncture et conjecture. Les baromètres économiques (1885-1930), *Histoire et Mesure*, *7*, 99-149.

[3] Bartlett, M. S. (1950), Periodogram Analysis and Continuous Spectra, *Biometrika*. *35*, 1-16.

[4] Bateman, D.V. and Mayes, F. (1970), Holiday Adjustment of Retail Sales, Unpublished memorandum, *US Bureau of the Census*, US Department of Commerce.

[5] Baxter, M. A. (1994), A Guide to Interpreting X-11-ARIMA/88 Diagnostics, Unpublished Memorandum, *Central Statistical Office*, United Kingdom.

[6] Bell, W.R. and Hillmer, S.C. (1984), Issues Involved With the Seasonal Adjustment of Economic Time Series,*Journal of Business and Economic Statistics*, *2*, 291-394.

[7] Bible, Revised Standard Version, Internet: http://www.hti.umich.edu/relig/rsv/.

[8] Bournay, J. and Laroque, G. (1979), Réflexions sur la méthode d'élaboration des comptes trimestriels, *Annales de l'INSEE*, *36*.

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

[10] Burman, J. P. (1980), Seasonal Adjustment by Signal Extraction, *Journal of the Royal Statistical Society, Series A, 143*, 321-337.

[11] Buys-Ballot, C. (1847), *Les changements périodiques de température*, Utrecht: Kemink et Fils.

[12] Cholette, P. A. (1981), A Comparison and Assessment of Various Adjustment Methods of Sub-Annual Series to Yearly Benchmarks, Working paper, Methodology Branch, *Statistics Canada*, Ottawa ON, Canada.

[13] Cholette, P. A. (1981), A Comparison of Various Trend-Cycle Estimators, *Time Series Analysis, O.D. Anderson & M.R. Perryman Editors*, Amsterdam: North-Holland.

[14] Cholette, P. A. and Dagum, E. B. (1994), Benchmarking Time Series with Autocorrelated Survey Errors, *International Statistical Review, 62*, 365-377.

[15] Cleveland, W. S. (1979), Robust Locally Weighted Regression and Smoothing Scatterplots, *Journal of The American Statistical Association, 74*, 829-836.

[16] Cleveland, R. B., Cleveland, W. S., McRae J.E. and Terpenning, I. (1990), STL a Seasonal-Trend Decomposition Procedure Based on Loess, *Journal of Official Statistics. 6*, 3-73.

[17] Cooley, J.W. and Tukey, J.W. (1965), An Algorithm for the Machine Calculation of Complex Fourier Series, *Mathematics of Computation, 19*, 297-301.

[18] Cournot, (1838), *Researches into the Mathematical Principles of the Theory of Wealth*, English translation 1897, New York: Macmillan.

[19] Dagum, E. B. (1975), Seasonal Factor Forecasts from ARIMA Models, *Proceedings of the International Institute of Statistics*, 40th Session, Contributed Papers, 3, Warsaw, 206-219.

[20] Dagum, E. B. (1980), The X-11-ARIMA Seasonal Adjustment Method, *Statistique Canada*, Catalogue 12-564E.

[21] Dagum, E. B. (1988), The X-11-ARIMA/88 Seasonal Adjustment Method, Methodology Branch, *Statistics Canada*, Ottawa ON, Canada.

[22] Doherty, M. (1992), The Surrogate Henderson Filters in X-11, Technical Report, *Department of Statistics*, Wellington, New Zealand.

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

[24] Fisher, A. (1937), A Brief Note on Seasonal Variation, *Journal of Accountancy, 64*, 54-59.

[25] Fourier, J.B. (1822), *The Analytical Theory of Heat*, New York: Dover Publications, published in 1955.

[26] Gardner, M. (1981) Mathematical Games, *Scientific American, February 1981*, 17-20.

[27] Gomez, V. and Maravall, A. (1996), Programs TRAMO and SEATS, *Banco de España*, Documento de Trabajo 9628.

[28] Gouriéroux, C. and Monfort, A. (1997), *Time Series and Dynamic Models*, Cambridge: Cambridge University Press.

[29] Grun-Rehomme, M. and Ladiray, D. (1994), Moyennes mobiles centrées et non centrées : construction et comparaison, *Revue de Statistique Appliquée, XLII*, 33-61.

[30] Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

[31] Henderson, R. (1916), Note on Graduation by Adjusted Average, *Transactions of the Actuarial Society of America, 17*, 43-48.

[32] Henderson, R. (1924), A New Method of Graduation, *Transactions of the Actuarial Society of America, 25*, 29-40.

[33] Herschel, W. (1801), Observations Tending to Investigate the Nature of the Sun in Order to Find the Causes or Symptoms of its Variable Emission of Light and Heat with Remarks on the Use that May be Possibly be Drawn from Solar Observation, *Philosophical Transactions of the Royal Society of London, 91*, 265-318.

[34] Higginson J. (1975), An F-test for the Presence of Moving Seasonality when Using Census Method II-X-11 Variant, Working Paper, Methodology Branch, *Statistics Canada*, Ottawa ON, Canada.

[35] Hillmer, S. C. and Tiao, G. C. (1982), An ARIMA Model Based Approach to Seasonal Adjustment, *Journal of the American Statistical Association, 77*, 63-70.

[36] Hood, C. C. (1998), X-12-Graph: A SAS/GRAPH Program for X-12-ARIMA Output, *US Bureau of the Census*, Washington, D.C.

[37] Hylleberg, S. (1986), *Seasonality in Regression*, Orlando: Academic Press.

[38] Hylleberg, S. (1992), The Historical Perspective, in *Modelling Seasonality*, London: Oxford University Press, 15-25.

[39] Jevons, W.S. (1862), On the Study of Periodic Commercial Fluctuations, *Investigations in currency and finance*, London: Macmillan, 1884.

[40] Kendall, M. (1973), *Time Series*, London: Charles Griffin & Co.

[41] Kitagawa, G. and Gersch, W. (1984), A Smoothness Priors State Space Modelling of Time Series with Trend and Seasonality, *Journal of the American Statistical Association*, *79*, 378-389.

[42] Koopman, S. J., Harvey, A. C., Doornik, J. A. and Shepard, N. G. (2000), *STAMP. Structural Time Series Analyser, Modeller and Predictor*, London: Timberlake Consultants Press.

[43] Koopmans, L. H. (1974), *The Spectral Analysis of Time Series*, New York: Academic Press.

[44] Laker, L.G. (1976a), Mathematical Note on Easter Correction, Working Paper, *Australian Bureau of Statistics.*

[45] Laker, L.G. (1976b), Slightly Less Mathematical Note on the Mathematical Note on Easter Correction, Working Paper, *Australian Bureau of Statistics.*

[46] Laniel, N. (1985), Design Criteria for the 13-term Henderson End-Weights, Working Paper, Methodology Branch, *Statistics Canada*, Ottawa ON, Canada.

[47] Lothian, J. (1978), The Identification and Treatment of Moving Seasonality in the X-11 Seasonal Adjustment Method, Working paper 78-10-004, Methodology Branch, *Statistics Canada*, Ottawa ON, Canada.

[48] Lothian, J. and Morry, M. (1978a), A Test for the Presence of Identifiable Seasonality when Using the X-11 Program, Working Paper, Time Series Research and Analysis Division, *Statistics Canada*, Ottawa ON, Canada.

[49] Lothian, J. and Morry, M. (1978b), A Set of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Method, Working Paper 78-10-005, Methodology Branch, *Statistics Canada*, Ottawa ON, Canada.

[50] Maballée, Colette et Berthe (1906), Peaks and Peaks and Correlograms, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 3*, 139-167.

[51] Macaulay, F.R. (1931), The Smoothing of Time Series, *National Bureau of Economic Research.*

[52] March, L. (1905), Comparaison numérique de courbes statistiques, *Journal de la Société Statistique de Paris*, 255-277.

[53] Menderhausen, H. (1937), Annual Survey of Statistical Technique: Methods of Computing and Eliminating Changing Seasonal Fluctuations, *Econometrica, 5*, 234-262.

[54] Montes, M.J. (1998-02-02), Calculation of the Ecclesiastical Calendar, Internet: http://www.smart.net/~mmontes/ec-cal.html.

[55] Musgrave, J. (1964a), A Set of End Weights to End all End Weights, Working paper, *US Bureau of the Census*, Washington.

[56] Musgrave, J. (1964b), Alternative Sets of Weights for Proposed X-11 Seasonal Factor Curve Moving Averages, Working paper, *US Bureau of the Census*, Washington.

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

[58] O'Beirne, T. (1966), The Regularity of Easter, *Bulletin of the Institute of Mathematics and Its applications, 2*, 46-49.

[59] Persons, W. M. (1919), Indices of Business Conditions, *Review of Economic Statistics, 1*, 5-107.

[60] Poynting, J.H. (1884), A Comparison of the Fluctuations in the Price of Wheat and in the Cotton and Silk Imports into Great Britain, *Journal of the Royal Statistical Society, 47*, 34-64.

[61] Priestley, M.B. (1965), Evolutionary Spectra and Nonstationary Processes, *Journal of the Royal Statistical Society, Series B, 27*, 204-237.

[62] SAS Institute Inc. (1990), *SAS/GRAPH Software Reference, version 6, First Edition, Volume1*, Carry, NC: SAS Institute.

[63] Shiskin, J., Young A. and Musgrave, J. C. (1967), The X-11 Variant of the Census Method II Seasonal Adjustment Program, Washington DC, Technical Paper no 15, *Bureau of the Census*, US Department of Commerce.

[64] Slutsky, E. (1927), The Summation of Random Causes as the Source of Cyclical Processes, *Econometrica, 84*, 105-146 (1937), translation of a Russian paper (Conjoncture Institute, Moscow).

[65] Tøndering, C. (2000), Frequently Asked Questions About Calendars, Internet: http//www.tondering.dk/claus/calendar.html.

[66] Tukey, J. W. (1950), The Sampling Theory of Power Spectrum Estimates, *Proceedings of the Symposium on Applications of Autocorrelation Analysis to Physical Problems*, Washington DC, Department of the Navy, Office of Naval Research.

[67] Young, A.H. (1965), Estimating Trading-Day Variations in Monthly Economic Series, Washington DC, Technical Paper no 12, *Bureau of Census*, US Department of Commerce.

[68] Yule, G.U. (1921), On the Time Correlation Problem, With Especial Reference to the Variate-Difference Correlation Method, *Journal of The Royal Statistical Society, 84*, 497-526.

[69] Yule, G.U. (1927), On a Method of Investigating Periodicities in Disturbed Series with Special Reference to Wolfer's Sunspot Numbers, *Philosophical Transactions of The Royal Society, Series A, 226*, 267-298.

Index

Lecture Notes in Statistics

For information about Volumes 1 to 105, please contact Springer-Verlag

106: Harald Niederreiter, Peter Jau-Shyong Shiue (Editors), Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing. xiv, 372 pp., 1995.

107: Masafumi Akahira, Kei Takeuchi, Non-Regular Statistical Estimation. vii, 183 pp., 1995.

108: Wesley L. Schaible (Editor), Indirect Estimators in U.S. Federal Programs. viii, 195 pp., 1995.

109: Helmut Rieder (Editor), Robust Statistics, Data Analysis, and Computer Intensive Methods. xiv, 427 pp., 1996.

110: D. Bosq, Nonparametric Statistics for Stochastic Processes. xii, 169 pp., 1996.

111: Leon Willenborg, Ton de Waal, Statistical Disclosure Control in Practice. xiv, 152 pp., 1996.

112: Doug Fischer, Hans-J. Lenz (Editors), Learning from Data. xii, 450 pp., 1996.

113: Rainer Schwabe, Optimum Designs for Multi-Factor Models. viii, 124 pp., 1996.

114: C.C. Heyde, Yu. V. Prohorov, R. Pyke, and S. T. Rachev (Editors), Athens Conference on Applied Probability and Time Series Analysis Volume I: Applied Probability In Honor of J.M. Gani. viii, 424 pp., 1996.

115: P.M. Robinson, M. Rosenblatt (Editors), Athens Conference on Applied Probability and Time Series Analysis Volume II: Time Series Analysis In Memory of E.J. Hannan. viii, 448 pp., 1996.

116: Genshiro Kitagawa and Will Gersch, Smoothness Priors Analysis of Time Series. x, 261 pp., 1996.

117: Paul Glasserman, Karl Sigman, David D. Yao (Editors), Stochastic Networks. xii, 298, 1996.

118: Radford M. Neal, Bayesian Learning for Neural Networks. xv, 183, 1996.

119: Masanao Aoki, Arthur M. Havenner, Applications of Computer Aided Time Series Modeling. ix, 329 pp., 1997.

120: Maia Berkane, Latent Variable Modeling and Applications to Causality. vi, 288 pp., 1997.

121: Constantine Gatsonis, James S. Hodges, Robert E. Kass, Robert McCulloch, Peter Rossi, Nozer D. Singpurwalla (Editors), Case Studies in Bayesian Statistics, Volume III. xvi, 487 pp., 1997.

122: Timothy G. Gregoire, David R. Brillinger, Peter J. Diggle, Estelle Russek-Cohen, William G. Warren, Russell D. Wolfinger (Editors), Modeling Longitudinal and Spatially Correlated Data. x, 402 pp., 1997.

123: D. Y. Lin and T. R. Fleming (Editors), Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis. xiii, 308 pp., 1997.

124: Christine H. Müller, Robust Planning and Analysis of Experiments. x, 234 pp., 1997.

125: Valerii V. Fedorov and Peter Hackl, Model-oriented Design of Experiments. viii, 117 pp., 1997.

126: Geert Verbeke and Geert Molenberghs, Linear Mixed Models in Practice: A SAS-Oriented Approach. xiii, 306 pp., 1997.

127: Harald Niederreiter, Peter Hellekalek, Gerhard Larcher, and Peter Zinterhof (Editors), Monte Carlo and Quasi-Monte Carlo Methods 1996, xii, 448 pp., 1997.

128: L. Accardi and C.C. Heyde (Editors), Probability Towards 2000, x, 356 pp., 1998.

129: Wolfgang Härdle, Gerard Kerkyacharian, Dominique Picard, and Alexander Tsybakov, Wavelets, Approximation, and Statistical Applications, xvi, 265 pp., 1998.

130: Bo-Cheng Wei, Exponential Family Nonlinear Models, ix, 240 pp., 1998.

131: Joel L. Horowitz, Semiparametric Methods in Econometrics, ix, 204 pp., 1998.

132: Douglas Nychka, Walter W. Piegorsch, and Lawrence H. Cox (Editors), Case Studies in Environmental Statistics, viii, 200 pp., 1998.

133: Dipak Dey, Peter Müller, and Debajyoti Sinha (Editors), Practical Nonparametric and Semiparametric Bayesian Statistics, xv, 408 pp., 1998.

134: Yu. A. Kutoyants, Statistical Inference For Spatial Poisson Processes, vii, 284 pp., 1998.

135: Christian P. Robert, Discretization and MCMC Convergence Assessment, x, 192 pp., 1998.

136: Gregory C. Reinsel, Raja P. Velu, Multivariate Reduced-Rank Regression, xiii, 272 pp., 1998.

137: V. Seshadri, The Inverse Gaussian Distribution: Statistical Theory and Applications, xi, 360 pp., 1998.

138: Peter Hellekalek, Gerhard Larcher (Editors), Random and Quasi-Random Point Sets, xi, 352 pp., 1998.

139: Roger B. Nelsen, An Introduction to Copulas, xi, 232 pp., 1999.

140: Constantine Gatsonis, Robert E. Kass, Bradley Carlin, Alicia Carriquiry, Andrew Gelman, Isabella Verdinelli, Mike West (Editors), Case Studies in Bayesian Statistics, Volume IV, xvi, 456 pp., 1999.

141: Peter Müller, Brani Vidakovic (Editors), Bayesian Inference in Wavelet Based Models, xi, 394 pp., 1999.

142: György Terdik, Bilinear Stochastic Models and Related Problems of Nonlinear Time Series Analysis: A Frequency Domain Approach, xi, 258 pp., 1999.

143: Russell Barton, Graphical Methods for the Design of Experiments, x, 208 pp., 1999.

144: L. Mark Berliner, Douglas Nychka, and Timothy Hoar (Editors), Case Studies in Statistics and the Atmospheric Sciences, x, 208 pp., 2000.

145: James H. Matis and Thomas R. Kiffe, Stochastic Population Models, viii, 220 pp., 2000.

146: Wim Schoutens, Stochastic Processes and Orthogonal Polynomials, xiv, 163 pp., 2000.

147: Jürgen Franke, Wolfgang Härdle, and Gerhard Stahl, Measuring Risk in Complex Stochastic Systems, xvi, 272 pp., 2000.

148: S.E. Ahmed and Nancy Reid, Empirical Bayes and Likelihood Inference, x, 200 pp., 2000.

149: D. Bosq, Linear Processes in Function Spaces: Theory and Applications, xv, 296 pp., 2000.

150: Tadeusz Caliński and Sanpei Kageyama, Block Designs: A Randomization Approach, Volume I: Analysis, ix, 313 pp., 2000.

151: Håkan Andersson and Tom Britton, Stochastic Epidemic Models and Their Statistical Analysis: ix, 152 pp., 2000.

152: David Ríos Insua and Fabrizio Ruggeri, Robust Bayesian Analysis: xiii, 435 pp., 2000.

153: Parimal Mukhopadhyay, Topics in Survey Sampling, x, 303 pp., 2000.

154: Regina Kaiser and Agustín Maravall, Measuring Business Cycles in Economic Time Series, vi, 190 pp., 2000.

155: Leon Willenborg and Ton de Waal, Elements of Statistical Disclosure Control, xvii, 289 pp., 2000.

156: Gordon Willmot and X. Sheldon Lin, Lundberg Approximations for Compound Distributions with Insurance Applications, xi, 272 pp., 2000.

157: Anne Boomsma, Marijtje A.J. van Duijn, and Tom A.B. Snijders (Editors), Essays on Item Response Theory, xv, 448 pp., 2000.

158; Dominique Ladiray and Benoît Quenneville, Seasonal Adjustment with the X-11 Method, xxii, 220 pp., 2001.