Statistics Canada
Statistique Canada

Methodology Branch

Time Series Research and Analysis Division

Direction de la méthodologie

Division de la recherche et de l'analyse des chroniques

Canada

[redacted]

Working Paper No TSRA-87-014E
Time Series Research and Analysis Division
Methodology Branch

Concepts, Definitions and Principles of

BENCHMARKING AND INTERPOLATION
OF TIME SERIES

by

Pierre A. CHOLETTE

Preliminary version April 1988

COMMENTS WELCOME

Statistics Canada
Methodology Branch
Time Series Research and Analysis Division
Coats Building, 13th floor "J"
OTTAWA, Canada
K1A 0T6

Telephone: (613) 951-0050

TABLE OF CONTENTS

List of figures - Part 1

List of Figures - Part 2

INTRODUCTION

Statistical agencies of developped countries publish figures on a very broad range of socio-economic variables. Most of those numbers take the form of time series, that is of periodic measures of each socio-economic variable considered. For instance, the monthly Unemployment and Price figures and the quarterly Gross National Product figures constitute time series. In many cases, a same socio-economic variable is published with different periodicities. For instance, the Gross National Product is released both quarterly and annually. Furthermore quarterly and montly series are generally made available in their seasonally adjusted forms. Overall, Statistics Canada publishes several hundred thousand time series; the Bank of Canada - which in some respects is a statistical agency - also publishes a large number of time series.

Time series are used by the various socio-economic decisions makers. Governments launch job creation programmes when time series indicate a recession. Central banks start anti-inflation policies, when time series show that prices begin to rise too fast. Car manufacturers slow down production when time series (pertaining to relevant variables) suggest that the market will not absorb the current production rates. Even the person in the street is a consumer of time series. The unemployed may not start looking for a job, if the unemployment statistics are high. The consumer may postpone the purchase of a house or of a car, if interest rates are too high or are rising too fast, or if the employment situation is too uncertain.

Many users of time series hold the natural view that the numbers released by statistical agencies are straight compilations of data originating from various sources of information like surveys, administrative records and censuses. In fact, most of the basic data obtained have to be adjusted, corrected or somehow processed by statisticians in order to arrive at useful, consistent and publishable values: Non-responses or illogical responses to questionnaires are imputed (replaced by reasonable values). Financial year data supplied by firms are adjusted to reflect the conventional year. Weekly data are converted into monthly values. Data supplied by large conglomerates (e.g. an oil company) are splitted into various industrial activities (e.g. exploration, extraction, refining, retailing); or, into various goods and services produced by the company. And all these activities, goods and services are broken down geographically (e.g. by Province). Monthly or quarterly data are often adjusted on the basis of more detailed and reliable yearly information. In many cases, the variables published are not even observable: They are indirectly derived from related information. According to the famous economist Milton Friedman (1962), "Most economic time series are highly manufactured products, constructed out of many bits and pieces that must be shaped and rearranged to yield the final series."

Without being as provocative as Friedman, it can be asserted that **subject-matter expertise** plays a major role in the establishement of most data published by statistical agencies. We define subject-matter as the intimate knowledge, by the time series builder, of the socio-economic processes and variables involved in the phenomena measured by the series considered. Under the same header, one should also include the intimate knowledge of the operational circumstances in which the socio-economic variable are measured. Many of the methods described in this document aim at incorporating in the series that subject-matter expertise in the most rational manner as possible.

Many of the processes and the transformation applied by statistical agencies are carried out by means of mathematical techniques. The present document is concerned with two families of these techniques: interpolation and benchmarking. Benchmarking, examined in Part 1, arises when data about a variable originate from two different sources with different periodicities. For instance the Canadian monthly Retail Trade series originate from a survey; and the corresponding annual values, from a census. The resulting monthly and annual series are generally not perfectly consistent. In particular, the annual totals of the former are not equal to the corresponding values of the latter. Benchmarking is the process of adjusting the sub-annual series to make it consistent with the annual benchmarks.

That classical definition of benchmarking assumes that the benchmarks are fully reliable - as their name implies. Section 1 enquires into the nature and the manifestations of bencmarks and introduces various concepts and notions relevant both to benchmarking and interpolation. It is found that in numerous situations, the reliability of the benchmarks is very questionable. Having no substitute, the word benchmark will continue to prevail. However, a broader definition of benchmarking, also used by Hillmer and Trabelsi (1987), is proposed. Benchmarking is the process of combining the monthly and the annual series, in order to obtain a consistent and more reliable pair of series for the socio-economic variable considered. This new definiton implies that the benchmarks may be binding or non-binding; and, that the classical definition is a particular of the latter.

A very exhaustive examination of benchmarking methods, done by Sanz (1981), recommends a variant of the Denton (1971) method. That method, based on the movement preservation principle, is easy to explain and most appropriate for socio-economic time series. Section 2 generalizes the approach, to accomodate the various benchmarking situations encountered in practice, namely financial year benchmarks, sub-annual benchmarks and unreliable benchmarks. Sections 3 and 4 discuss implementational issues of benchmarking: preliminary benchmarking of the current year, revision of benchmarked values and the computational aspects of benchmarking. Section 5 addresses the problem of benchmarking systems of series subject to aggregation constraints, geographical and industrial aggregation for instance. Section addresses the evaluation of benchmarked series.

Interpolation, examined in Part 2 of the document, arises when the desired quarterly (say) values are simply missing for the variable considered. Estimates are then derived from relevant external quarterly information and from values available annually (say). The components of Consumer Expenditures, for example, are not observed quarterly. The quarterly values are then derived by means of year-to-year growth rates, determined by the series builder on the basis of various sources of information. Sections 2 (of Part 2) proposes a method of growth rate interpolation; and Section 3, a method to achieve similar results by means of ARIMA interpolation. Section 1 presents methods for the "calendarization" of socio-economic data, for instance the conversion of fiscal year data into conventional (i.e. calendar) year values and the conversion (of bundles) of weekly data into monthly values. These problems can be approached as simplest cases of interpolation and benchmarking.

Some methods of benchmarking and interpolation described in this paper are not actually applied in practice. In our opinion, much of the potential for cheaper, improved, better integrated and more aboundant statistics, offerred by benchmarking and interpolation techniques, remains un-tapped.

PART 1: BENCHMARKING

Before proceeding with the content of benchmarking, its context needs to be clarified. Many of the concepts and definitions presented are also relevant for interpolation.

1. CONCEPTS AND DEFINITIONS

This Section enquires into the nature and properties of the benchmarks, of the unbenchmarked series and of other notions relevant to benchmarking and interpolation. Gaps are observed between seemingly straightforward concepts and their practical manifestations. Benchmarking and interpolation methods have to be specifically designed to accomodate the factual situations.

1.1 The Original "Sub-Annual" Series

As mentionned earlier, many socio-economic indicators are made available with different periodicities. Typically, many socio-economic variables are published monthly and annually; or quarterly and annually. The need for benchmarking arises when the sources of information for the more frequent series (e.g. monthly) and for the less frequent series are different. In the absence of benchmarking, the two series, which describe the same variable, could contradict each other. The more frequent (e.g. monthly) series will labelled the "sub-annual series", or the "original series", with respect to the less frequent series. The less frequent series will be labelled the "benchmarks" series. In this paper, "sub-annual" series may thus actually refer to daily or weekly data, if considered against monthly data; or even to annual data, if considered against quinquennial data. By this convention, we hope to achieve a more concrete discussion. In most applications considered indeed, the sub-annual series are either monthly or quarterly. Resuming the example of the Introduction, the original sub-annual series is the monthly values of Retail Trade, obtained from a survey; and the annual benchmarks, the annual values of Retail Trade, originating from a census. The characteristics of sub-annual series are now examined.

"Sub-annual" series are usually less reliable than their benchmarks. This unreliability refers to the fact that the "true" series ξ_t is observed with an error ϵ_t:

$$(1.1) \qquad x_t = \xi_t + \epsilon_t$$

where x_t is the available observation. The larger the error is (in absolute terms), the more unreliable is the series x_t. In this document - at least -, we are not interested in the nature of the error ϵ_t (in whether it is a sampling error, an observation error, an estimation error, etc.); but, in the mere fact that it exists. However, for the purpose of benchmarking, it is useful to decompose that error into **bias** and variance. When the error is substantial but behaves very predictably (from a purely chronological point of view) from period t to period t+1, series x_t is considered biased. For instance, if from period to period a series repeatedly under-estimates the true situation by 10% (negative error), it is ruled biased. When the error is substantial but unpredictable from one period to the next, the series is considered to have a large variance, that is to be **erratic** or more technically speaking non-efficient. Thus unreliability may take the forms of variance, of bias, or both. The **Mean Square Error** statistic embodies both the variance and the bias:

$$\text{M.S.E.}_t = \text{var}(x_t) + \text{bias}_t^{\ 2} \tag{1.2}$$

One purpose of benchmarking is to improve the reliability of the sub-annual series, that is to reduce its M.S.E. Depending on the presence and on the nature of unreliability, different variants of benchmarking will be in order.

Some sub-annual surveys use **sample rotation**, that is the respondents in the sample remain therein for a few periods (e.g. six months). It is well known that, in the absence of corrective action, sample rotation causes the estimates to remain above or under the true target value for several time periods. From the point of view of benchmarking, this effect will also be considered as bias.

Some of the reasons for unreliability are now outlined. In principle, the Canadian Statistics Law makes it compulsory for individuals and businesses to answer survey questionnaires sent to them by Statistics Canada. However, both filling in and processing questionnaires is laborious. In practice, statistical agencies try to avoid burdening the respondents with questionnaires. One way of achieving this is to send to the respondent **less detailled questionnaires sub-annually** and more detailled ones annually; to select **smaller sample sizes sub-annually** (fewer respondents) and larger samples sizes annually. The latter practice, especially, results in more erratic sub-annual series.

Statistical agencies also alleviate the response burden by resorting to **administrative records**. Administrative statistics come as the by-product of the activity of some organization. School boards, hospitals, churches, courts, regulatory agencies, for instance, generate data about school enrollments, incidence of deseases, births, marriages, incidence of crimes. Such statistics fulfil the needs of the organizations producing them; but do not necessarily meet the standards and requirements of the statistical agency. Furthermore, there may be differences in procedures and quality across regions and organizations. For these and other reasons, administrative statistics must usually be adjusted by the statistician in order to make them usable. The resulting sub-annual series may not be fully reliable. More details about the characteristics administrative data can be found in Brackstone (1987).

Another source of unreliability of the sub-annual series is the following. Many respondent companies send their data in **bundles of four or five weeks**. For example, a department store sends its data in bundles of 4, 4 and 5 weeks; 4, 4 and 5 weeks; and so on. This pattern is typical of the Canadian Retail Trade series. For Wholesale Trade, a common practice is to send figures every 4 weeks, that is in 13 bundles per year. In both cases, central statisticians adjust such data to convert them into monthly values. The problem is especially serious when bundles end and start in the middle of a month. The quality of the adjustment made impacts on the reliability of the resulting monthly series - on its variance especially.

Section 1.4 of Part 2 proposes methods to address that kind of problem.

In the context of benchmarking - and interpolation especially -, the sub-annual series may be just an **indicator series**, with a scale of magnitude and with units different from those of the desired series. For instance, the sub-annual series may be an index series of some kind, expressed in percentages; whereas the desired series is to be in billions of dollars. Such original sub-annual series are obviously biased estimates of the desired series. To the extent the indicators are gross, they are also erratic.

There are broad general rules governing the reliability of sub-annual series. Geographically or industrially aggregate series tend to be more reliable than than the corresponding geographical or industrial component series. Series pertaining to more developed (regions of) countries tend to be more reliable than those of less developped countries. Annual series tend to be more reliable than the corresponding monthly or quarterly series. However there are many exceptions to these rules.

In many cases, the reliability of the original sub-annual series may be assessed by graphical examination. Indeed, all the components of time series, except for the irregular and the trading-day components, are smooth with respect to time: By definition, seasonality tends to repeat almost exactly from year to year; and the trend-cycle values are locally monotonic or change direction in a gradual manner. Trading-day fluctuations appear erratic against time. However they are neglibible in stock series; and in quarterly and annual flow series, because all quarters and years have almost exactly the same number of trading-days. Consider a variable known to have no trading-day fluctuations and to behave in an essentially smooth manner (e.g. population). If the series measuring that variable behaves erratically, then it can be ruled as erratic. The reverse is not necessarily true however, because some variables are known to have un-predictable behaviours, like construction series. In other words, such a series cannot be ruled erratic because it behaves erratically.

The purpose of this section is not to exhaust all sources of unreliablilty of sub-annual series, but to give an insight on how it may arise.

1.2 The "Annual" Benchmarks

The original sub-annual series is one of the inputs of benchmarking. Another input consists of the annual benchmarks. A benchmark is the relatively less frequent measurement of a socio-economic variable considered; and the sub-annual series, the more frequent measurement. That definition encompasses quinquennial or decennial census values of population with respect to the annual values. In most applications considered however, the benchmarks will refer to annual values, and the term "annual benchmarks" will prevail.

Benchmarks usually originate from relatively more reliable sources of information, like censuses (e.g. the Canadian annual Census of Manufactures) of the target population. They are therefore considered as less biased and to have lower variances than the corresponding sub-annual series. Before qualifying that statement, it is appropriate to introduce the concepts of flow series, of stock series and of index series.

Conceptually, **flow series** are such that monthly values are the sum of the daily values in the month; quarterly values, the sum of the monthly values; and so forth. For instance the amount of gasoline sold in Canada in 1986 is the sum of the gasoline sold in each month of 1986. Thus all trade series, all export and import series, all income series are flow series. Implicit in flow series is the notion of velocity. The faster socio-economic agents purchase goods and services, for instance, the higher National Expenditures for the period considered; conversely, the more they delay their purchases, the lower National Expenditures.

Stock series, on the other hand, reflect the level of a variable at one particular date. Population series, employment series, inventory series (e.g. oil reserves) are all stock series. The annual values of stock series often correspond to the value of the last sub-annual period of the year. Thus the annual values of inventories correspond to the December 31 value. In some cases, the yearly values of stock series is ruled to be the annual averages of the corresponding sub-annual series. The annual values of unemployment in 1986, for instance, is the average of the monthly 1986 values. For the purposes of benchmarking, only the first type of stock (e.g. inventories) will be labelled as stock series. The other type of stock will fall in the category of index series. Note that annual stock series, as just defined, are essentially seasonal: they are the monthly or quarterly selected as the yearly value.

Stricly speaking, **index** series are those for which the annual values correspond to the average of the sub-annual values and which are expressed as percentage of a base-year. Thus the Consumer Price index (1971=100%) and the Index of Industrial Production (1971=100%) are index series in the strict sense. (Conceptually, the former is also a stock and the latter a flow series.) For the purpose of benchmarking, the second part of the definiton is dropped. Unemployment series would therefore be labelled as index series, if their annual values are defined as the average of the sub-annual figures. As defined herein, the concepts of flow, stock and index series are then mutually exclusive and collectively exhaustive.

The relationship between an annual and the corresponding sub-annual series (which is to be retored by benchmarking), is then straightforward: For flow series, the annual benchmarks correspond to the annual sums of the sub-annual series; for stock series (as defined above), to the same sub-annual value from year to year; and for index series, to the annual averages of sub-annual values. Complications occur in practice however.

In many cases, the annual data actually forwarded to statistical agencies do not refer to the conventional year, i.e. the year ranging from January to December; but, to the financial year of the individual respondents, e.g. July to June. (The expressions calendar year and fiscal year are often encountered.) This situation is typical of all the Canadian surveys of businesses. Consider a respondent with a financial year extending from April to March for instance. The "annual" data supplied by such a respondent tend to over-estimate the true conventional year value, in case of rising activity (positively sloped trend-cycle component); and to under-estimate, in case of decline. Such annual values are then biased, and the bias varies with the phase of the busines cycle. When considered as benchmarks pertaining to conventional years, they can be qualified as erratic, because their bias is unpredictable from a purely chronological point of view.

However, all variants of benchmarking and interpolation described in this document specify the benchmarks as pertaining to the time periods they actually cover. This solves the financial year problem, if all the respondents (contributing to the benchmarks) have the same financial year. That condition does arise in practice - in the institutional sector especially, e.g. local governements, school boards, hospitals. If needed, conventional year values can be obtained as a by-product of benchmarking, by taking the annual sums (or relevant operation) of the series benchmarked to the homogeneous financial year values.

Unfortunately, in most cases the respondents do not share a common financial year. One practice under such circumnstances is to rule any "annual" data covering any financial year which end between April 86 and March 87 (say) as pertaining to year 1986. Under that rule, any respondent with one of those twelve possible **financial years**, May 1985 to April 1986, June 1985 to May 1986, ..., May 1986 to March 1987, is classified in 1986; and the annual value for 1986 is simply the sum of the data of the respondents classified in 1986. As observed by Cholette (1987a), such a scheme creates very serious problems. For flow series, the estimates are biased, and the bias depends on the phase of the business cycle. For the purpose of benchmarking such estimates are then erratic.

But there is more. The biases also depend on the distribution of the respondents over the 12 possible financial years for the variable considered (Ibidem). Indeed, the relevant distribution is not the number of respondents in each financial year but the number of corresponding units of the variable considered (e.g. dollars of sales, volume, persons, etc.). Since that distribution changes with the variable considered - even for a given set of respondent -, the bias varies with the variable considered. Thus relations between variables, which prevailed at the respondent level, are destroyed by the aggregation, because the variables are subject to different biases. This predicament is unacceptable to decisions makers, who base their decisions on sets of socio-economic variables and not on variables considered in isolation; and, unacceptable to statistical agencies, who have the mandate to produce integrated systems of series. For stock series the situation is worst, because the seasonality inherent to stock series (see above) is not preserved.

Unless the respondents have a common financial year, it is necessary to convert financial year data into conventional year values before their aggregation (over respondents) into annual benchmarks. Section 1 of Part 2 proposes methods to do that. However, any such method involves assumptions about the underlying sub-annual values (known or unknown) of the respondent and estimation errors. The resulting estimated conventional year values are certainly less reliable than if there had been not need to carry out such a conversion.

In Statistics Canada, some annual benchmarks originate from the annual input-output model of the economy. Such models are exhaustive and integrated accounting frameworks, which trace the sales (production) and the purchases of each of the goods and services in the real (as opposed to financial) economy. For instance, an input-output model traces how much of the paper sold (produced) by the paper industry was purchased by industry A; how much, by industry B; etc.; how much, by the consumer; and how much was exported. More generally, the production of any industry must be purchased by the other industries, by the consumer or by foreign buyers. It is indeed materially impossible that part of the production of a good or service was not purchased by anyone (nor added to inventories) and disappeared.

Input-output models thus provide a way to compare data on the production of paper provided the paper industry, to the data originating from other sources, on the industrial purchases of paper (say), on the consumer purchases of paper, on the exports of paper. In other words, an input-output framework can be used to cross-validate data originating from different sources like censuses, surveys and administrative records. Combined with subject-matter expertise, the model also reveals the probable location of inconsistencies. The statistician may then correct the data. Furthermore, since the model is exhaustive of all the economy, the values arrived at for any variable must at least be consistent with the values of every other variable in the model. These are the grounds to believe that the values arrived at after such an integration exercise are better annual benchmarks than the annual values originally available from the individual censuses and surveys.

One prime example of annual benchmarks supplied by the input-output model is that of the National Account series. The National Accounts system is literally a sub-set of the input-output framework. The latter covers all "intermediate" goods and services, which enter the production of other goods and services; and all "final" goods and services, on which the Accounts focus. Both frameworks also cover the incomes (sales) corresponding to the expenditures (purchases) on goods and services. Since the Accounts series, namely the income and expenditures series, are quarterly, it is natural to use the annual series from the input-model as benchmarks.

A few more comments about annual benchmarks are in order. In many situations, benchmarks are available every second year or, sometimes, in an irregular fashion. For a year considered, the annual benchmark usually becomes available (if at all) several months and sometimes more than a year after the year is over. In some cases, two or three benchmarks at the time become available. These operational circumstances have implementational implications for any benchmarking method, which are examined in Sections 3 and 4.

Finally, some sub-annual series have no annual benchmarks. For a variable considered in isolation, there would be simply no need to benchmark. The problem arises when the series is part of a system of series, most of which are subject to benchmarking. In order to preserve consistency of the system, some kind of adjusment has to be performed on the variable with missing benchmarks. Section 5 outlines how this may be achieved.

This sub-section revealed some of the problems encountered with annual benchmarks. It should now be clear that annual benchmarks may not be annual - in more than one sense. They may not be reliable, as their name leads to believe. (However having no better word to propose and given its already wide acceptance, the word benchmark - and the derived word benchmarking - will continue to prevail in this document.) The logical attitude resulting from this kind of conclusion is the following. The annual benchmarks should not always be considered as binding values to be complied with by the sub-annual series. In many cases, they should merely be considered as extra observations available for the socio-economic variable considered, besides the sub-annual series, and should be treated as non-binding. Consequently, all the benchmarking variants presented in this document allow for both **binding** and **non-binding** benchmarks.

1.3 The "Annual" Discrepancies

A sub-annual series and independently obtained annual benchmarks for a same socio-economic variable give rise to **annual discrepancies**. These measure the degree of inconsistency between the original sub-annual series and the annual benchmarks. For flow series, the annual discrepancies are the differences between the annual benchmarks and the annual sums of the sub-annual series; for stock series (as defined above), the difference between the annual benchmarks and the one applicable sub-annual values; and for index series the difference between the annual benchmarks and the annual averages of the sub-annual series. A more useful concept is that of **annual proportional discrepancies**, which are the annual descrepancies expressed as ratios instead of differences.

Both the proportional and the additive annual discrepancies contain information about the benchmarking situation. Table 1.1 displays a simplified classification of possible benchmarking situations, depending on the reliability of the benchmarks and of the sub-annual series. The table also shows the corresponding behaviours of the discrepancies and the appropriate benchmarking variant to use in each situation. The table is simplified in that it assumes polar cases of un-biasedness and erraticity (efficiency); and, in that it seeks simple rules for selecting the appropriate benchmarking variants.

In the first column, the annual benchmarks are reliable: they are both un-biased and non-erratic (efficient). That column pertains to the "classical" benchmarking situations. Indeed, except for Hillmer and Trabelsi (1987), all benchmarking methods in the litterature, from Bassie (1939) to Litterman (1983), force the sub-annual series to comply with the benchmarks. In other words, they consider the benchmarks as binding and therefore - at least implicitly - as reliable. Reliable annual benchmarks may be specified as binding, and benchmarking may be carried out regardless

TABLE 1.1: Behaviour of the annual discrepancies and appropriate benchmarking variants depending on the reliability of the sub-annual series and of the annual benchmarks

		1	2	3	4	
		Non-Biased Annual Benchmarks		Biased Annual Benchmarks		
		Non-Erratic (reliable)	Erratic (unreliable)	Non-Erratic (unreliable)	Erratic (unreliable)	
Non Biased Sub-Annual Series	1 Non-Erratic (reliable)	constant discrepancies	erratic discrepancies	constant discrepancies	erratic discrepancies	behaviour of discrepancies
		binding benchmarking	non-binding benchmarking	special* benchmarking	special* benchmarking	appropriate action
	2 Erratic (unreliable)	erratic discrepancies	erratic discrepancies	erratic discrepancies	erratic discrepancies	behaviour of discrepancies
		binding benchmarking	non-binding benchmarking	special* benchmarking	special* benchmarking	appropriate action
Biased Sub-Annual Series	3 Non-Erratic (unreliable)	constant discrepancies	erratic discrepancies	constant discrepancies	erratic discrepancies	behaviour of discrepancies
		binding benchmarking	non-binding benchmarking	special* benchmarking	special* benchmarking	appropriate action
	4 Erratic (unreliable)	erratic discrepancies	erratic discrepancies	erratic discrepancies	erratic discrepancies	behaviour of discrepancies
		binding benchmarking	non-binding benchmarking	special* benchmarking	special* benchmarking	appropriate action

* special benchmarking variants explained in Section 2.3

of the reliability of the original sub-annual series. Indeed, if the sub-annual is also reliable (row 1), then the discrepancies are likely to be very small (i.e. trivially constant), benchmarking is a mere formality, and whether the benchmarks are specified as binding or not binding will have little effect. In all the other cases, where the sub-annual series is not reliable (row 2, 3 and 4) benchmarking is advisable, because it improves the reliability of the sub-annual series. The reliability of the sub-annual series is not a pre-requisite for benchmarking. (Section 2.1.2 on the assumptions of benchmarking will examine this in more detail.) Note that graphical examination of the original sub-annual series provides information about its reliability (see Section 1.1).

Section 1.2 described many situations with potentially unreliable benchmarks. These correspond to columns 2 to 4 of Table 1.1. In the second column, the annual benchmarks are not reliable, because they are erratic (although unbiased). If the sub-annual series is reliable (row 1), then no benchmarking should be attempted, because benchmarking could only deteriorate an already good sub-annual series. The benchmarks could be discarded. The best annual values of the socio-economic variable considered are simply the annual sums (or the appropriate arithmetic operation) of the sub-annual series. In practice however, the sub-annual series is not likely to be absolutely unbiased and could probably be improved. It is then possible to specify the benchmarks as non-binding. (Note that no benchmarking is equivalent to specify non-binding benchmarks with no weight attributed to them.) If the sub-annual series is also not reliable (row 2, 3 and 4), benchmarking should take place with non-binding benchmarks (See Section 2). The resulting benchmarked series is a combination - a compromise in a sense - between the sub-annual and the annual series. That combination of two unreliable series is generally more reliable than each of the individual series. (That principle is well established in statistics.)

The discussion to this point can be summarized as follows: When the benchmarks are reliable (colum 1) they may be specified as binding; when they are not reliable but un-biased, as non-binding. This philosophy would be applicable to any benchmarking method. In the third and fourth columns of Table 1.1, the benchmarks are biased. The series builder is then not likely to contemplate benchmarking. For the sake of thoroughness however, the opportunities (and the lack thereof) provided by benchmarking under such conditions are examined. This will be done in Section 2.3, when more background on benchmarking principles is available.

The behaviour of annual discrepancies are thus usefull to assess the benchmarking situation, before benchmarking is attempted. Constant discrepancies can occur in 4 out the 16 cases tabulated. However, since each behaviour of the discrepancies can point to several situations, which call for different variants of benchmarking, subject matter expertise is crucial to tell the situations apart. Section 6 will expand on the assessment of benchmarking situations and benchmarking results.

1.4 The Sub-Annual Benchmarks

Some situations involve another input to benchmarking: the sub-annual benchmarks. These are occasional sub-annual values available besides the corresponding values of the original sub-annual series. Sub-annual benchmarks are usually more reliable than the original series. (The term sub-annual benchmark is selected by analogy to the annual benchmarks.) The process of benchmarking then combines the annual and the sub-annual benchmarks with the original sub-annual series. In case of sub-annual benchmarks specified as binding, the resulting benchmarked series runs through the sub-annual benchmark values.

Both in interpolation and benchmarking, sub-annual benchmarks can be a vehicle for subject-matter expertise. Indeed the series builder may use them to impose certain seasonal patterns or certain vicinities to the desired series for certain periods of time. For instance, sub-annual benchmarks can be used to "freeze" benchmarked series prior to a certain date, by forcing the desired series to start with a value considered as historical. Later sections will examine those opportunities in more details.

The annual benchmarks of stocks series are - technically speaking - equivalent to sub-annual benchmarks. Indeed both pertain to individual sub-periods (e.g. months, quarters). Sub-annual benchmarks can also be considered as particular cases of annual benchmarks. The latter may pertain to several sub-periods; and, the former, to individual sub-periods. For the time being however, we find it useful to keep them as separate concepts.

1.5 The Sub-Annual Discrepancies

As in the case of annual benchmarks, sub-annual benchmarks give rise to **sub-annual discrepancies**. These are the difference between the sub-annual benchmark and the corresponding original sub-annual observation. The **sub-annual proportional discrepancies** are the ratio (instead of the difference) of the same.

In much the same manner as the annual discrepancies, the sub-annual ones can be used to assess the benchmarking situation. Furthermore, the sub-annual and annual discrepancies should align. If for instance all annual and sub-annual discrepancies hover in the neighbourhood of 1.10 except for one sub-annual discrepancy equal to 0.80, the series builder should question the corresponding sub-annual benchmark or the original sub-annual value.

1.6 Constraints Between Series

Benchmarking often has to be performed on systems of series bound together by additivity constraints or more generally by linear constraints. Additivity constraints typically occur when an aggregate variable is the sum of geographical or industrial component variables. For example, the benchmarked Canadian Retail Trade Sale series must be equal to the sum of the corresponding provincial sale series for each period of time.

Constraints also arise with financial variables, and with economic variables seen through the framework of some financial accounting system like the System of National Accounts. Accounting constraints are not essentially different from additivity constraints, except they may involve substractions and fewer variables. For instance Profits are identically equal to Receipts minus Expenditures - for each period of time. This identity which prevails in the original sub-annual series (say) must still hold in the benchmarked series.

Section 5 will show how additivity and accounting constraints between series can be satisfied.

1.7 Illustration of the Benchmarking Problem

The various notions covered in this section are summarized in Figure 1.1. The solid curve x_t stands for the original sub-annual series, that is for the un-benchmarked series. In the case considered, the series is a flow and ranges from the fourth quarter of 1982 to the third quarter of 1987. Each annual benchmarks y_i is represented by its average value for the year. This practice, maintained throughout this document, conveniently displays the annual benchmarks of flow series on the same scale as the sub-annual series and also indicates their reference periods. In the figure, the annual benchmarks refer to conventional years. However, that does not need to be the case for benchmarking to work. No benchmark is available for 1984 and 1987, which is imcomplete (i.e. "current"). The benchmarking situation illustrated also provides for two sub-annual benchamrks z^d_t. These pertain to the first and the sixteenth term of the series, that is to the fourth quarter of 1982 and third of 1986. They specify that the benchmarked series desired should run close to those two points.

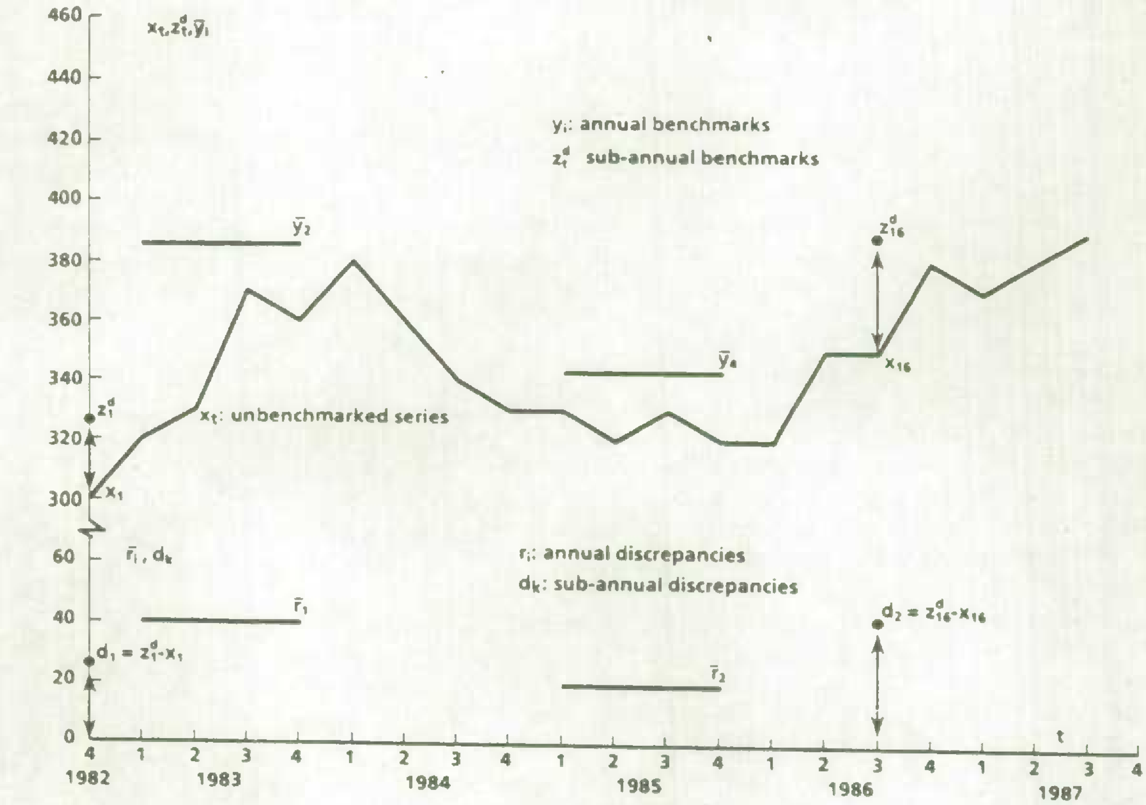


Figure 1.1: Illustration of the benchmarking problem when both annual and sub-annual benchmarks are available

The lower part of the figure displays the annual and the sub-annual discrepancies corresponding to the annual and sub-annual benchmarks of the upper part. Like for the annual benchmarks, each annual discrepancy (of flow series) is best represented by the annual average discrepancies, that is by its average over the reference period. Both the annual average discrepancies and the sub-annual discrepancies are then represented on a common scale. As explained in Section 1.3, the examination of the discrepancies provides a first assessment of the benchmarking situation and suggests which variants of benchmarking to use. This brings the discussion to the principles of benchmarking.

2. PRINCIPLES AND ASSUMPTIONS OF BENCHMARKING

With that body of concepts and definitions now available, a more precise definition of benchmarking may be attempted. Benchmarking is the process of optimally combining the original sub-annual series with the annual benchmarks and with the sub-annual benchmarks, in order to obtain a more reliable sub-annual series and a more reliable annual series. These series with reduced Mean Square error are respectively the **benchmarked series** and the **derived annual series**.

This definition of benchmarking is somewhat unconventional. To our knowledge, the attitude of statistical agencies has been to adjust the sub-annual series to comply to the benchmark values; that is to consider the latter as binding - hence the word benchmarking. This attitude assumes that the benchmarks are not just more reliable than the sub-annual series; but fully reliable. As discovered in Section 1, the relative reliability of the annual benchmarks is, in many situations, very questionnable. In our opinion however, the annual series does not have to be more reliable than the sub-annual series in order for "benchmarking", as defined herein, to be useful. All the variants of benchmarking presented in this paper allow for both binding and non-binding benchmarks.

There are two approaches to benchmarking time series: the numerical approach and the statistical approach. The statistical approach specifies a statistical model followed by the desired series. Hillmer and Trabelsi (1987) and Guerrero (1987) - these are the only two references for the statistical approach to date - specify that the true time series behaves according to an ARIMA model (Box and Jenkins, 1970). If it were possible to observe the true series, there would be no inconsistency between the annual and the sub-annual measurements of the series. In other words, the inconsistencies between the measurements available at different frequencies originate from the fact that the series is observed with some error. The authors accordingly set out to estimate the underlying ARIMA model on the basis of both the available sub-annual and annual observations. The fitted values obtained are the desired consistent values. This still experimental approach is interesting. Note that it does not consider the benchmarks as necessarily binding. However, it would substantially reduce the irregular fluctuations and - at this development stage at least - eleminate the trading-day fluctutations. Furthermore, it does not lend itself to massive application in a statistical agency: It requires too much expertise in time series modelling and forecasting. It would possibly be appropriate for a few key socio-economic indicators.

The wide-spread numerical benchmarking methods (Lisman and Sandee, 1964; Boot et al. 1967; Denton, 1971; Ginsburg, 1973; Laniel, 1986), on the other hand, specify no statistical model followed by the series. However, some of them at least could be argued to specify a descriptive model for the desired series. For instance, the Denton approach to benchmarking adopted in this document is based on the **principle of movement preservation**. That descriptive principle is easy to explain: The benchmarked series preserves as much as possible of the consecutive month-to-month movement of the original sub-annual series (including the movement from one year to the next, e.g. from December of one year to January of the next year). Numerical methods also lend themselves to large scale application. In the only comprehensive review of benchmarking methods to our knowledge, Sanz (1981) recommends a slightly modified variant of the Denton method (Cholette, 1978) for use in statistical agencies.

The initial Denton method (1971), did not preserve the movement of the original series to the maximum possible extent. In some situations, the benchmarked series could display severe movement distortions at the start of series. For the long post-war series, this did not matter much, since interests usually focuses on the latest values. Now, the current and predictable trend is towards short series. As a result, the kind of distortions has become unacceptable, and the problem was corrected by Cholette (1978, 1984). The computational reasons behind the initial specification have also become obsolete. This document generalizes the movement preservation benchmarking approach to encompass the various benchmarking situations encountered in practice and described in Section 1 (unreliable benchmarks, financial year benchmarks, etc.); and, makes explicit the assumptions implicit to the approach.

The principle of movement preservation may be expressed in at least two ways: 1) preserving the simple period-to-period change and 2) preserving the period-to-period percentage change. These two forms of preservation yield two variants of benchmarking: additive benchmarking and proportional benchmarking. Both the additive and proportional variants are applicable to situations involving un-biased benchmarks. When the benchmarks are (un-biased and) non-erratic, they should be specified as binding; when they are (un-biased and) erratic, as non-binding. Those events correspond to the first and second columns of Table 1.1 respectively. Section 2.3 presents other variants of benchmarking applicable to situations with biased benchmarks.

2.1 Additive Benchmarking under un-biased benchmarks

Additive benchmarking aims at producing a benchmarked series which displays the movements of the original series. In other words, the two series are to be as parallel as possible. This parallelism, illustrated in Figure 2.1, is best conveyed by the corrections. The corrections or the adjustment factors are simply the modifications made to the original series x_t to arrive at the benchmarked series z_t. The corrections are then measured by the distance between the two series. Obviously, the benchmarked and the original series will be parallel to the extent the distance between them (the corrections) is constant. Under the movement preservation principle, benchmarking then consists of finding smooth corrections, which are as constant as possible and run through the annual and sub-annual discrepancies. With binding benchmarks, case considered in Figure 2.1, the correction curve runs exactly through the sub-annual discrepancies and crosses each annual average discrepancy in a very specific manner: The surface covered by the corrections and by the annual average discrepancies are exactly the same over the reference period of the annual benchmarks.

2.1.1 Specification of Additive Benchmarking Under Un-Biased Benchmarks

The maximum parallelism between the benchmarked and the original sub-annual series is easily achieved by means of mathematical tools. The desired benchmarked series z_t minimizes the following objective function:

$$f(z) = g^x{}_t \sum_{t=2}^{T} \{(z_t - x_t) - (z_{t-1} - x_{t-1})\}^2 + \sum_{m=1}^{M} g^y{}_m \{(\sum_{t=\tau_m}^{\rho_m} z_t) - y_m\}^2 \qquad (2.1)$$

$$+ \sum_{k=1}^{K} g^z{}_k \{z_t - z^d{}_k\}^2 .$$

Parameters M and K respectively stand for the numbers of annual and sub-annual benchmarks y_m and $z^d{}_k$. The symbols τ_m and ρ_m specify the reference periods of the annual benchmarks. Depending on the values of τ_m and ρ_m, the benchmarks may be specified to be available every year, every second year or irregularly; or to cover conventional year or financial years. For flow series ρ_m is greater than τ_m. For stock series ρ_m is equal to τ_m. For index series, y_m actually stands for the annual benchmarks multiplied by the number of months per year (presumably equal to $\rho_m - \tau_m + 1$). The notation of the annual benchmarks then implies no restriction as to their pattern of availability and as to their reference periods. (The notation could also encompass the sub-annual benchmarks. However we choose to keep the latter distinct and conspicuous for the time being.)

The first term of the objective function embodies the movement preservation principle. This **parallelism criterion** specifies that the corrections $(z_t - x_t)$ change as little as possible from one time period to the next - including between years. In other words, the corrections are as constant as possible. The parallelism criterion specifies the movement of the benchmarked series but not the level. The criterion could allow a benchmarked series with a level very different from that of the original series and still be totally satisfied (i.e. be equal to its minimun of zero).

The level of the benchmarked series is determined by the second and third terms of the objective function, the **benchmarks satisfaction criteria**. The second term specifies that the sums of the desired benchmarked values z_t over the reference periods $[\tau_m, ..., \rho_m]$ are as close as possible to the available benchmarks y_m. And the third term states that the benchmarked values sought are as close as possible to the available sub-annual benchmarks $z^d{}_k$.

Parameters $g_t{}^x$, $g^y{}_m$ and $g^z{}_k$ are known relative weights attributed to the parallelism criterion and to the benchmark satisfaction criteria. High weights $g^y{}_m$ and $g^z{}_k$ relative to $g_t{}^x$ specify binding benchmarks in practice. Binding benchmarks situations are the classical ones covered by column 1 of Table 1.1. Low weights, on the other hand, specify non-binding benchmarks, covered by column 2 of Table 1.1. Specifying non-binding benchmarks could be described as Bayesian benchmarking, in that the statistician set the weight of the benchmarks on the basis of prior knowledge. (A starting point for the weights in the Bayesian framework would be, assuming no sub-annual benchmarks, $g^x = (\sigma^2_x - \rho\sigma_x\sigma_y)/(\sigma^2_x+\sigma^2_y-2\rho\sigma_x\sigma_y)$ and $g^y = 1-g^x$. For large scale applications, determining and using those variances may be just as difficult as determining g^x and g^y directly.)

The implementation recommended in Section 4 actually circumvents the problem of determining those weights. Indeed, the numerical approximation proposed therein makes it un-necessary and un-desirable to formally minimize objective function (2.1) (or (2.3)) to carry out benchmarking. The objective functions can be considered as a tool to formally represent and think about the benchmarking process. For the time being, the weights can therefore be assumed to have been pre-selected to reflect binding benchmarks or non-binding benchmarks when appropriate.

Appendix B shows that formally minimizing objective function (2.1) yields the following solution:

$$z_t = \sum_{t=1}^{T} p^x{}_{t,n}\, x_n + \sum_{m=1}^{M} p^y{}_{t,m}\, y_m + \sum_{k=1}^{K} p^z{}_{t,k}\, z^d{}_k. \tag{2.2}$$

Each benchmarked values maximizing the criteria specified in the objective function is a weighted average of the original sub-annual values and of the annual and sub-annual benchmarks.

2.1.2 Assumptions of Additive Benchmarking Under Un-Biased Benchmarks

Minimizing objective function (2.1) involves certain implicit assumptions on the part of the statisticians. The literature is rather vague about them. An attempt is made here to make the assumptions as explicit as possible:

1) The movement in the sub-annual series is worth preserving. Reliability is not required.

2) The annual benchmarks are unbiased. They may be erratic.

3) The corrections to be made to the original sub-annual series do not depend on the sub-annual series, but on the discrepancies.

The first assumption is rather weak. The original sub-annual series may be both biased and erratic. Indeed, the movements of a very erratic original series will persist in the benchmarked series - with longer and therefore smoother movements introduced by benchmarking. Contrary to statistical benchmarking approaches, numerical benchmarking does not require the presence of a signal in the sub-annual series. If needed, this allows benchmarking to be carried out at very low levels of aggregation, where the sub-annual series are typically unreliable and contain little signal. Subsequent aggregation is that which is likely to increase the signal to noise ratio. Whether to benchmark un-reliable series is a policy decision. Once taken, the decision to benchmark implies that the sub-annual movement is better than nothing and is worth preserving. (More precisely, the sub-annual movement is better than that implicit in the benchmarks alone, see Section 1 of Part 2.)

The second assumption, embodied in the second and third terms of the objective function, states that the benchmarks are unbiased. The possible bias in the original sub-annual series is corrected on the basis of the benchmarks. However, the assumption does not imply reliable benchmarks. They may in fact be erratic. High weights $g^y{}_m$ or $g^z{}_k$ for the benchmark satisfaction criteria specify un-biased non-erratic benchmarks (first column of Table 1.1); and low weights, un-biased erratic benchmarks (second column). The benchmarks are binding in the first case and non-binding in the second case.

Figure 2.2 compares the corrections obtained for annual benchmarks first considered as binding and then as non-binding. The binding corrections incorporate the erratic character of the annual average discrepancies, in the form of smooth oscillatory movement. A reliable (say) sub-annual series would be distorted accordingly. This illustrates the effect of mis-specifying benchmarks as binding. The non-binding corrections, on the other hand, do not incorporate the movement implied by the discrepancies; but, only their (moving) average level, in the form of a nearly constant curve. The resulting benchmarked series therefore keeps the movement of the original series with little distortion, but changes its level by the amounts determined by the correction curve. In order of ascending causality, the level change is thus determined by the corrections, by the discrepancies and by the benchmarks.

Figure 2.2 also calls for a few rather digressive but irresistible comments. Under the non-binding corrections, the third annual benchmarks is almost satisfied, despite the fact it is not considered any more binding than the other ones. The annual benchmarks refer to financial years ranging from the second quarter to the first of the following year. The single sub-annual benchmark depicted is binding (in both cases) and forces the corrections and therefore the benchmarked series to start from a pre-specified value. That feature will exploited for implementational purposes (Section 4).

One can sumarize the two first assumptions as follows. The sub-annual series provide the sub-annual period-to-period movement of the benchmarked series; and the (appropriately weighted) benchmarks, the correct level.

According to the third assumption, the corrections $(z_t - x_t)$ do not depend on the (individual) values of the sub-annual series; but only on the annual and sub-annual discrepancies - and their weights. This carries the following implication. If the source of sub-annual bias (in x_t) is undercoverage in a survey, for instance, the behaviour of the respondents not covered (which are in (z_t-x_t)) is governed by the discrepancies. In other words, benchmarking may be used as an implicit and aggregate method of imputation. The implicitly imputed values are determined by the discrepancies. More specifically, the behaviour of the inputed values is specified to be as smooth and constant as possible. Put differently, benchmarking assumes that the aggregate under-coverage moves very gradually from one period to the next; and therefore, displays no sub-annual fluctuations, in particular no seasonality. (With proportional benchmarking by contrast, the behaviour of the individuals not covered is governed both by the discrepancies and by each value of the sub-annual series, that is by the behaviour of the respondents covered.)

Similarly if the source of discrepancy is a slightly inadequate industrial classification of the sub-annual series (say), benchmarking can be used to correct for that. Again this involves certain assumptions which must be clear to the statistician: The aggregate effect of the sub-annually mis-classified units is smooth and gradual from one period to the next. Furthermore, in the case of classification, it is assumed that the mis-classification does not affect the sub-annual movement (e.g. the seasonal pattern) of the original series, which is being preserved by benchmarking. That assumption is rather strong, since different industries (activities) are likely to display different seasonalities, and since economic agents usually try to have seasonally complementary activities (e.g. logging in the winter and farming in the summer).

One can reason in an analogous manner with every possible sources of annual discrepancies. The practitionner should in fact clarify the implicit assumtions of benchmarking depending on the nature of the discrepancies.

None of the three assumptions required the discrepancies to be small. They may actually be very large compared to the original sub-annual values. If they are constant, the benchmarked series will be exactly parallel to the original series. In other words, under constant discrepancies, however big, the original series is a perfect indicator of the sub-annual movement; and the benchmarks, a perfect indicator of the level. (This also implies that the longer movement implicit in the benchmarks is totally consistent with the corresponding movement of the original series.) This property of benchmarking is most relevant, when the original series has an order of magnitude other than that of the benchmarks.

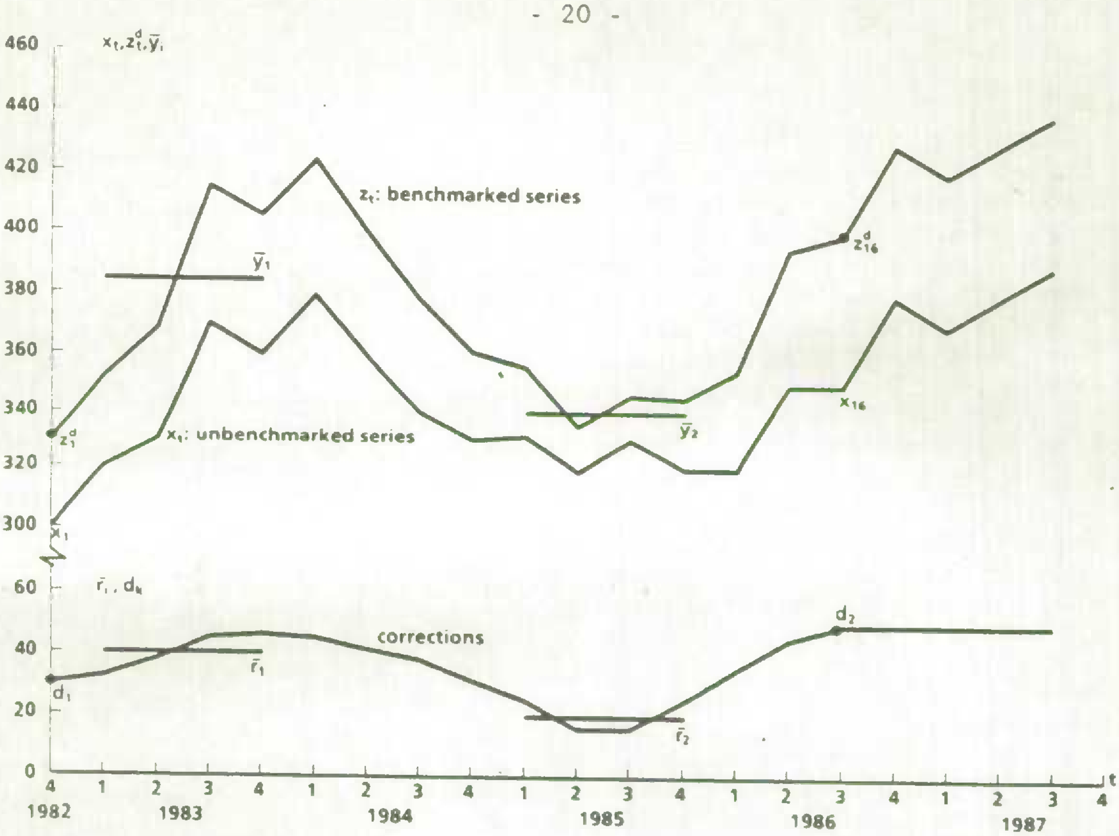


Figure 2.1: Additive benchmarking according to the principle of movement preservation in the presence of annual and sub-annual benchmarks

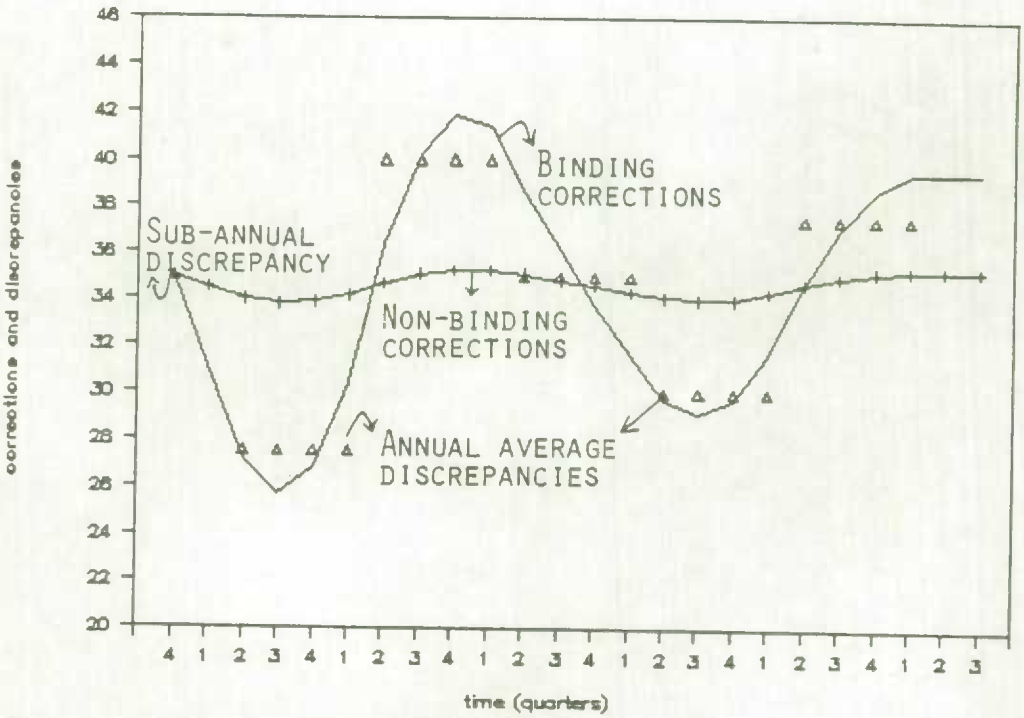


Figure 2.2: Additive corrections under binding benchmarks and under non-binding benchmarks

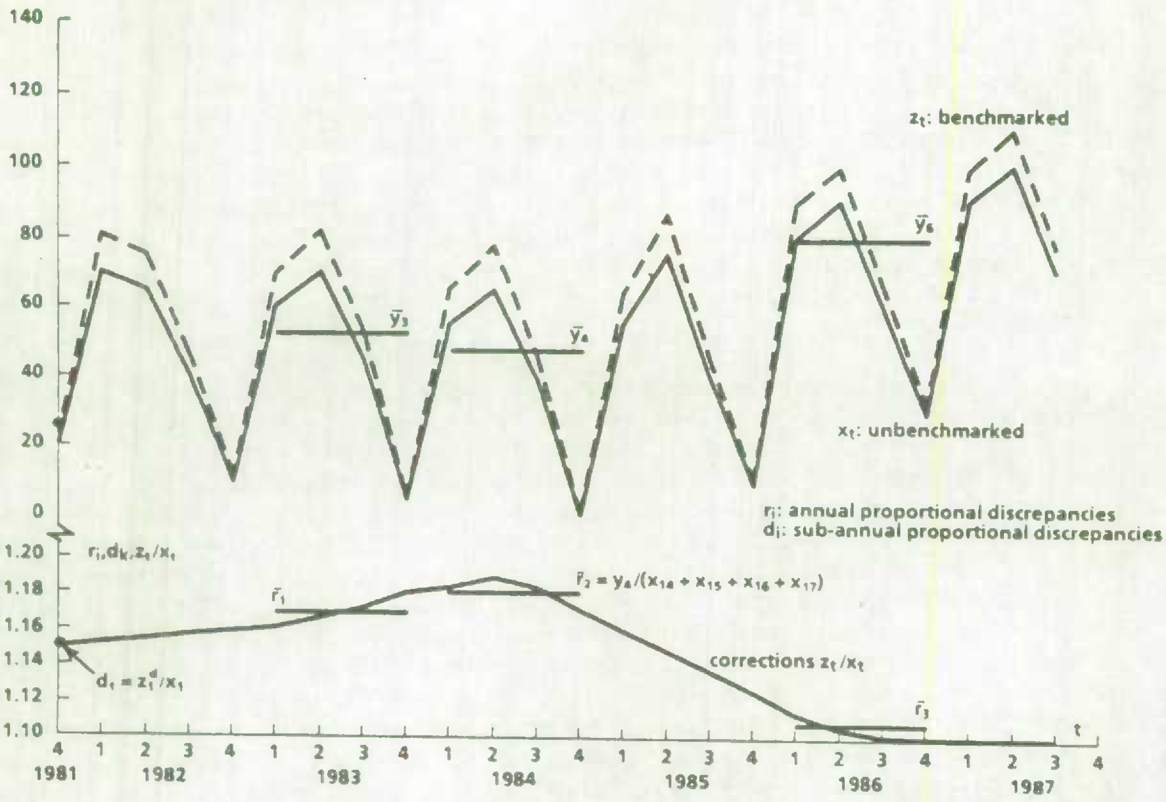


Figure 2.3: Proportional benchmarking according to the movement preservation principle in the presence of annual and sub-annual benchmarks

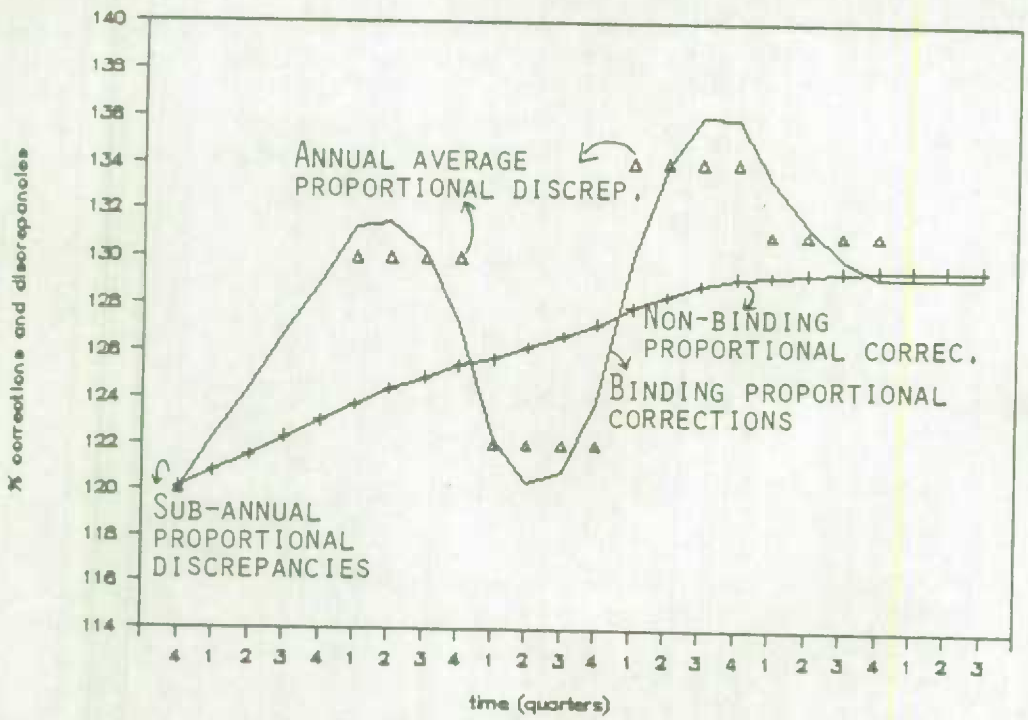


Figure 2.4: Proportional benchmarking corrections under binding benchmarks and under non-binding benchmarks

2.2 Proportional Benchmarking Under Un-Biased Benchmarks

As already mentionned, the principle of movement preservation may also be expressed in terms of percentage changes from one period to the next. It will be shown that two series, namely the benchmarked and the original series, have the same percentage movement, to the extent they are proportional to each other. Consequently proportional benchmarking can be considered as an approximation of percentage change preservation. However, as will soon become apparent, proportional benchmarking is very defendable per se (even if it were not an approximation of percentage change preservation).

Figure 2.3 illustrates a case where proportional benchmarking would be appropriate. When an original series x_t is very seasonal, it can be argued that the seasonal trough values can not reasonably account for the annual discrepancies to the same extent as the seasonal peak values. Indeed, the activity of the variable is almost nill in all fourth quarters and should presumably remain small after benchmarking. Proportional benchmarking fulfills such a requirement. **Proportional corrections** are ratios to be multiplied by the original sub-annual series. Their effect is obviously larger for larger original observations and smaller for smaller ones. As intended, the corresponding benchmarked series z_t is very close to the original sub-annual values in trough quarters. Like additive corrections, proportional corrections z_t/x_t are as smooth and constant as possible. In the case of binding benchmarks displayed in the figure, the corrections run exactly through the sub-annual proportional discrepancies and cross the annual proportional discrepancies.

2.2.1 Specification of Proportional Benchmarking Under Un-Biased Benchmarks

Like in the additive variant, the mathematics of proportional benchmarking are specified on the corrections. However, the corrections are expressed as the ratio of the desired benchmarked series z_t to the corresponding original values x_t. The desired benchmarked series minimizes the following objective function:

$$f(z) = g^{x}{}_{t} \sum_{t=2}^{T} \{(z_t/x_t) - (z_{t-1}/x_{t-1})\}^2 + \sum_{m=1}^{M} g^{y}{}_{m} \{((\sum_{t=\tau_m}^{\rho_m} z_t)/y_m) - 1\}^2 + \sum_{k=1}^{K} g^{z}{}_{k} \{(z_t/z^{d}{}_{k}) - 1\}^2. \quad (2.3)$$

where all symbols retain the same meaning as for additive benchmarking (equation (2.1)).

The first term of the objective function spells the **proportional movement preservation criterion.** That criterion specifies that the proportional corrections (z_t/x_t) change as little as possible from one time period to the next - including between years. In other words, the corrections are as constant as possible. The proportionality criterion specifies the movement of the benchmarked series but not the level. The criterion could allow a benchmarked series with a level very different from that of the original series and still be totally satisfied (i.e. be equal to its minimum of zero). This properties is very useful when the original series has an order of magnitude or a scale different from the desired benchmarked series. In other words, benchmarking - the proportional variant especially - may be used as a tool to transform percentages, for instance, into billions of dollars.

The level of the benchmarked series is determined by the second and third terms of the objective function, the **benchmark satisfaction criteria**. These criteria are expressed as ratios of the desired series to the benchmarks. Compared to the additive variant, this greatly simplifies the choice of weights $g^x{}_t$, $g^y{}_m$ and $g^z{}_k$. Indeed, all the individual terms (inside the summations) in the objective function have comparable magnitudes, because they lie immediately above zero. (For more details about the weights, see Section 2.1.)

The first criterion is equivalent to a percentual movement preservation criterion. Indeed two series z_t and x_t which are proportional to each other

$$z_t / x_t = z_{t-1} / x_{t-1}$$

have identical period-to-period growth rates:

$$\Rightarrow \quad z_t / z_{t-1} = x_t / x_{t-1}.$$

More generally, to the extent two series are proportional to each other, they tend to have the same growth rates:

$$z_t / x_t = z_{t-1} / x_{t-1} + a_t \quad \Rightarrow$$

$$z_t / z_{t-1} = x_t / x_{t-1} + a_t \, x_t / z_{t-1}.$$

$$= x_t / x_{t-1} \quad \text{as } a_t \text{ tends towards zero}$$

where a_t represent deviations from proportionality. To the extent proportionality is realized, the proportionally benchmarked series z_t displays the same growth rates as the original sub-annual series x_t.

Appendix C gives the solution to the formal minimization of objective (2.3). Section 4 provides a numerical approximation of the solution which is much more economical.

2.2.2 Assumptions of Proportional Benchmarking Under Un-Biased Benchmarks

The assumptions of proportional benchmarking, as specified by objective function (2.3), are the following:

1) The percentual movement in the sub-annual series is worth preserving. Reliability is not required.

2) The annual benchmarks are unbiased. They may be erratic.

3) The corrections to be made to the original sub-annual series depend on the annual and the sub-annual discrepancies and on the sub-annual series.

The two first assumptions are identical to those in the additive benchmarking: The sub-annual series provide the sub-annual period-to-period movement of the benchmarked series; and the (appropriately weighted) benchmarks, the correct level.

Like in the additive variant, the benchmarks may be specified as binding or non-binding, depending on the weights g^y_m and g^z_k chosen by the series builder. Figure 2.4 compares the corrections obtained for annual benchmarks first considered as binding and then as non-binding. The binding corrections incorporate the erratic character of the annual average discrepancies, in the form of smooth oscillatory movement. A reliable (say) sub-annual series would be distorted accordingly. This illustrates the effect of mis-specifying benchmarks as binding. The non-binding corrections, on the other hand, do not incorporate the erratic movement implied by the discrepancies; but, - in the example chosen - only their trend. The resulting benchmarked series therefore keeps the movement of the original series with little distortion, but changes its level (and trend) by the amount determined by the correction curve (hence by the discrepancies and by the benchmarks). In the case illustrated, the annual benchmarks refer to conventional years. The single sub-annual benchmark depicted is binding (in both cases) and forces the corrections and therefore the benchmarked series to start from a pre-specified value.

According to the third assumption, the corrections depend both on the original sub-annual series and on the annual and sub-annual discrepancies. Indeed, since the corrections multiply the original series, their effect depend on the values of that series. If the source of sub-annual bias (in x_t) is under-coverage, for instance, the behaviour of the respondents not covered (which are in z_t-x_t) is consequently governed by the behaviour of the individuals covered (in x_t) and by the discrepancies. In other words, benchmarking may be used as an implicit and aggregate method of impution. With proportional benchmarking, the implicit imputed values depend on the available non-imputed values and on the discrepancies. Apart from that distinction, the assumptions are the same as for additive benchmarking. If unclear about the assumption of benchmarking, the reader is urged to refer to Section 2.1.2.

Another virtue of proportional benchmarking is to avoid negative benchmarked values, when both the benchmarks and the sub-annual series are positive. With the additive variant on the other hand, negative benchmarked values would be very likely to occur in a situation like that depicted by Figure 2.3. However the proportional variant produces negative values if some of the sub-annual or of the benchmark values are negative. The proportional variant is unusable when the sub-annual series contains zero values. In such a case zeroes may be replaced by values different from zero but infinitesimally close to zero.

Like in the additive case, proportional benchmarking does not require small discrepancies. This allows one to build a series in millions of persons (say) from an original series expressed in percentages. The proportional variant is very appropriate in such situations.

2.3 Benchmarking Under Biased Benchmarks

Both variants of benchmarking considered until now assumed that the benchmarks were unbiased. As a result, they provided the level of the benchmarked series, whereas the sub-annual series supplied the sub-annual movement. This section enquires into the possible usefulness of biased benchmarks. Such benchmarks correspond to the situations referred to by columns 3 and 4 of Table 1.1. Before specifically examining those situations, more general objective functions, based on the principle of proportional movement preservation, are presented.

Biased benchmarks may contain information about the non-seasonal part of the sub-annual movement - especially when they are non-erratic. With biased benchmarks (erratic or not), the following objective function would be appropriate:

$$
\begin{aligned}
f(z) = {} & g^{x}{}_{t} \sum_{t=2}^{T} [(z_t / x_t) - (z_{t-1} / x_{t-1})]^2 \\
& + \sum_{m=2}^{M} g^{y}{}_{m} [((\sum_{t=\tau_m}^{\rho_m} z_t) / y_m) - ((\sum_{t=\tau_{m-1}}^{\rho_{m-1}} z_t) / y_{m-1})]^2 \\
& + \sum_{k=2}^{K} g^{z}{}_{k} [(z_{tk} / z^{d}{}_{k}) - (z_{tk-1}) / z^{d}{}_{k-1})]^2 \qquad \text{(read subscripts as } t_k \text{ and } t_{k-1}) \\
& + \sum_{i=1}^{I} \{[(\sum_{t=(i-1)J+1}^{iJ} z_t) / (\sum_{t=(i-1)J+1}^{iJ} x_t)] - 1\}^2
\end{aligned}
\tag{2.5}
$$

where J is the number of months per year, I is the number of complete years in series.

The first term is the proportional movement preservation criterion. (For the additive criterion, all divide signs in the objective function should be replaced by minuses.) The second and third terms specify that the movement observed in the annual and sub-annual benchmarks should be reflected in the desired benchmarked series z_t. The weights $g^{y}{}_{m}$ and $g^{z}{}_{k}$ are the weights attributed to the movement in the benchmarks y_m and $z^{d}{}_{k}$. High values for those weights specify that the movement of the benchmarks is binding; this would be appropriate for non-erratic benchmarks (column 3). Low values specify non-binding movement of the benchmarks; this would be appropriate for erratic benchmarks (column 4). The fourth term specifies that the benchmarked series z_t should adopt the annual levels of the original sub-annual series x_t. That term assumes that the original sub-annual series is unbiased (row 1 and 2). (Such a term specifying the level is required. Its absence could result into a benchmarked series with an order of magnitude different from those of either the benchmarks and of the original series.)

It is possible to specify a yet more general objective function for situations where the levels of both the benchmarks and the sub-annual series are wrong. This new objective function could thus apply to all cases referred to in columns 3 and 4 of Table 1.1:

$$
\begin{aligned}
f(z) \; = \; & g^{x}{}_{t} \sum_{t=2}^{T} [(z_t / x_t) - (z_{t-1} / x_{t-1})]^2 \\
& + \sum_{m=2}^{M} g^{y}{}_{m} \, [((\sum_{t=\tau_m}^{\rho_m} z_t) / y_m) - ((\sum_{t=\tau_{m-1}}^{\rho_{m-1}} z_t) / y_{m-1})]^2 \\
(2.6) \qquad & + \sum_{k=2}^{K} g^{z}{}_{k} \, [(z_{tk} / z^{d}{}_{k}) - (z_{tk-1}) / z^{d}{}_{k-1})]^2 \\
& \qquad\qquad\qquad\qquad \text{(read subscripts as } t_k \text{ and } t_{k-1}) \\
& + \sum_{i=1}^{I} [(\sum_{t=(i-1)J+1}^{iJ} z_t) - (\alpha \sum_{t=(i-1)J+1}^{iJ} x_t)]^2 \\
& + \sum_{m=1}^{M} g^{y}{}_{m} \, [(\sum_{t=\tau_m}^{\rho_m} z_t) - \beta \, y_m]^2
\end{aligned}
$$

where α and β respectively stand for the known proportional bias of the sub-annual series and of the annual benchmarks (e.g. 1.10 and 0.70). Those two parameters could also be (linearly) estimated in the minimization process.

The eight situations of columns 3 and 4 of Table 1.1 may now be addressed in the framework of objective function (2.6). If the sub-annual series is absolutely reliable (row 1, column 3 and 4), no benchmarking is advisable. The annual values may be obtained from the original sub-annual series. Even if the benchmarks are non-erratic (column 3), there is no movement in the annual benchmarks which is not present in the sub-annual series, otherwise the discrepancies would not be constant; and benchmarking would not improve the sub-annual series. In practice however, the sub-annual series x_t is not absolutely un-biased. The appropriate variant of benchmarking would then be given by objective function (2.6), where the level of the benchmarked series is specified to be that of the original ($\alpha = 1$ $\beta = 0$) and where the movement of the benchmarks is specified as binding in case of non-erratic benchmarks (colunm 3) and non-binding in case of erratic benchmarks (column 4). The same variant is also appropriate in case of erratic sub-annual series (row 2), except the movement of the benchmarks should be given more weight than when the sub-annual is non-erratic. In both cases (rows 1 and 2), benchmarking preserves the annual level of the un-biased sub-annual series and imposes some of the movement in the benchmarks, as this could improve the movement in the sub-annual series.

In the two situations where the sub-annual series is biased but absolutely non-erratic and the benchmarks are biased (row 3, columns 3 and 4), no benchmaking is advisable. One could then simply adjust the level of the sub-annual series on the basis of subject matter expertise (for instance raise the series by 10%). Indeed, if the benchmarks are non-erratic (column 3), all their movements are also contained in the sub-annual series (otherwise the discrepancies would not be constant); if the annual series is erratic (column 4), then any benchmarking would make the sub-annual series more erratic then it is. In practice however, the sub-annual series is not likely to be absolutely non-erratic. The appropriate variant of benchmarking would then be given by objective function (2.6), where the level of the benchmarked series z_t is specified to be higher (or lower) than that of the original series by a factor α (>0), where the movement in the benchmarks is specified as binding in case of non-erratic benchmarks (column 3) and as non-binding in case of erratic benchmarks (column 4). In case of non erratic benchmarks (column 3), the level of the benchmarked series may also be specified to be higher (or lower) than that of the benchmarks by a factor β (both α and $\beta > 0$). The same variant is also appropriate in case of erratic sub-annual series (row 4), except the movement of the benchmarks should be given more weight than when it is non-eratic. In both cases (rows 3 and 4), benchmarking exogenously changes the annual level of the sub-annual series (by means of pre-selected factors α and β) and imposes some of the movement in the benchmarks, as this could improve the movement in the sub-annual series.

Simpler and probably more practical solutions to the problem of biased benchmarks are the following. Use (publish) the original sub-annual series as it is (and forego benchmarking), when the sub-annual series is unbiased (rows 1 and 2 and columns 3 and 4 of Table 2.1); and use (publish) only the movement of the sub-annual series, when it is biased (rows 3 and 4). This results into Table 2.1 which is simpler than Table 1.1. Indeed annual benchmarks cannot be good indicators of sub-annual trend-cycle movements: The movements derived from annual discrepancies cannot provide precise dates (in terms of month or quarter) of turning-points in the business cycle. Furthermore, benchmarking by means of (2.6) does not lend itself to massive application in statistical agencies, because it would require too much expertise. This sub-section at least put the more practical objective functions (2.1) and (2.3), as well as their underlying assumptions into perspective.

Section 2 raised the opportunity of benchmarking when the benchmarks are un-biased and clarified the assumptions and the implications of benchmarking by means of the movement preservation principle. The next section discusses the issue of preliminary benchmarking.

TABLE 2.1: Behaviour of the annual discrepancies and appropriate benchmarking variants depending on the reliability of the sub-annual series and of the annual benchmarks

		1	2	3	4	
		Non-Biased Annual Benchmarks		Biased Annual Benchmarks		
		Non-Erratic (reliable)	Erratic (unreliable)	Non-Erratic (unreliable)	Erratic (unreliable)	
1 Non Biased Sub-Annual Series	Non-Erratic (reliable)	constant discrepancies	erratic discrepancies	constant discrepancies	erratic discrepancies	behaviour of discrepancies
		binding benchmarking	non-binding benchmarking	no benchmarking	no benchmarking	appropriate action
2	Erratic (unreliable)	erratic discrepancies	erratic discrepancies	erratic discrepancies	erratic discrepancies	behaviour of discrepancies
		binding benchmarking	non-binding benchmarking	no benchmarking	no benchmarking	appropriate action
3 Biased Sub-Annual Series	Non-Erratic (unreliable)	constant discrepancies	erratic discrepancies	constant discrepancies	erratic discrepancies	behaviour of discrepancies
		binding benchmarking	non-binding benchmarking	use only change in x_t	use only change in x_t	appropriate action
4	Erratic (unreliable)	erratic discrepancies	erratic discrepancies	erratic discrepancies	erratic discrepancies	behaviour of discrepancies
		binding benchmarking	non-binding benchmarking	use only change in x_t	use only change in x_t	appropriate action

3. PRELIMINARY BENCHMARKING

As mentionned earlier, the annual benchmark of a year typically becomes available well after the year is over. Recent or current sub-annual data are then released to the public before they are benchmarked. This operational circunstance raises the issue of preliminary benchmarking. Figure 3.1 (a) presents a case of preliminary benchmarking under binding benchmarks. Comparing the current unbenchmarked values of 1987 with past benchmarked of 1986, that is the absence of preliminary benchmarking, produces a movement discontinuity equal to CB. The user of such data is lead to believe that the socio-economic variable is going down. This may impair decision making. When the 1987 benchmark does become available, benchmarking will eliminate the decline CB (as shown later). Comparing the preliminary benchmarked values with the past benchmarked values, on the other hand, yields no discontinuity (movement CD).

There are at least two approaches to preliminary benchmarking. One consists of forecasting the sub-annual series and the annual benchmark for the current incomplete year and of benchmarking as if those forecasts were genuine data. This is the avenue taken by Bassie (1939), Lisman and Sandee (1964); Laniel (1986) proposed the use of simple ARIMA models (Box and Jenkins, 1970).

Another avenue consists of forecasting the corrections for the current year. The approach to benchmarking proposed in Section 2 provides the required forecasted corrections as a by-product. Under the movement preservation criterion used, the optimal corrections merely repeat the last correction calculated for the last year for which there was a benchmark. In other words, such preliminary benchmarking factors, displayed in Figure 3.1 (a), are also the ones which result by actually minimizing objective function (2.1) (or (2.3)). Indeed the criterion minimized does not distinguish between years which have benchmarks from those which do not; nor, between years which are complete from those which are not. The benchmarked series is kept as parallel (or proportional) as possible to the original for all years, and this applies to the current year. For the current year without benchmark, complete parallelism (or proportionality) is achievable, hence the constant corrections.

Repeating the last correction is also equivalent to cumulatively applying the growth observed in the current unbenchmarked observations to the last benchmarked value; in Figure 3.1 (a), to applying additive growth AB to point C, which yields point D. This also holds in the proportional model. Indeed, applying the percentage growth x_{T+p} / x_{T+p-1} in the current unbenchmarked series yields the following preliminary benchmarked values for periods T+1, T+2,...:

$$z_{T+1} = (x_{T+1} / x_T) * z_T = x_{T+1} * (z_T / x_T)$$

$$z_{T+2} = (x_{T+2} / x_{T+1}) * z_{T+1}$$

$$= (x_{T+2} / x_{T+1}) * x_{T+1} * (z_T / x_T)$$

$$= x_{T+2} * (z_T / x_T)$$

and so on; or more generally

$$z_{T+p} = (x_{T+p} / x_{T+p-1}) * z_{T+p-1}$$

$$\text{(3.1)} \qquad \Rightarrow \quad z_{T+p} = x_{T+p} * (z_T / x_T), \quad p=1,2,3,...$$

where z_T / x_T is the last correction made for the last year with a benchmark. As shown by (3.1), that same correction is applied to all current observations. (The use of growth rates for preliminary benchmarking is not to be confused with series interpolation by means of growth rates. See Section 2 of Part 2.)

Repeating the last correction for preliminary benchmarking is equivalent to forecasting the next annual average discrepancy, at level E in Figure 3.1 (a) (in the additive case); and therefore, implicitly equivalent to forecasting both the annual benchmarks and the sub-annual series. The forecasted discrepancy is slightly higher than the last one, if the previous one was lower than the last one (case displayed). Converserly, the forecasted discrepancy is slightly lower than the last one if the previous one was higher. The forecasts also level off, which provides protection against "turning points" in the discrepancies.

Figure 3.1 (b) illustrates the same situation as Figure 3.1 (a), after the 1987 benchmark is available. Three scenarios corresponding to three 1987 benchmark values are considered. In the first scenario, the 1987 benchmark $y^1{}_3$ exactly confirms preliminary benchmarking. This fortunate scenario entails no revision to the preliminary benchmarked series $z^1{}_t$. In the second scenario, the 1987 benchmark $y^2{}_3$ is much higher than anticipated by preliminary benchmarking. The series is then revised from $z^1{}_t$ to $z^2{}_t$. The resulting revised change C'D' between 1986 and 1987 is still comparable to the preliminary change CD. In that case at least, change CD was a better predictor of C'D' than CB was (without preliminary benchmarking).

The third scenario is the worst: The 1987 benchmark $y^3{}_3$ is much lower than anticipated and yields a negative annual discrepancy (when they used to be positive). The series is accordingly revised from $z^1{}_t$ to $z^3{}_t$. It is now debatable whether CD was a better predictor of C"D" than CB was. However one point is not debatable: nor the preliminary benchmarked series $z^1{}_t$ nor the benchmarked series $z^3{}_t$ display discontinuity CB, which occurs in - and is uniquely due to - the absence of preliminary benchmarking. Furthermore the preliminary benchmarked series, represented by $z^1{}_t$, lies in the middle of the extreme scenarios and is therefore the most likely.

The normative considerations discussed under the assumptions of benchmarking actually rule out the possibility of extreme scenarios. In an eventual erratic discrepancy situation, non-binding benchmarks would be appropriate. The benchmarked series would be much more comparable with the preliminary benchmarked series than in the case of Figure 3.1.

The preliminary benchmarking scheme just described is no argument againts the forecasting approach first mentionned. In some critical situations, forecasting the original series and the next annual benchmark may produce superior results (lower revisions). The approach should be entertained for key socio-economic indicators, provided the availability of the forecasting expertise. For mass production however, the preliminary correction factors derived from the movement preservation criterion are satisfactory - as illustrated by Figure 3.1 - and trivially easy to calculate. Their foreknowledge also means that the benchmarking exercise can be conducted only when a new benchmark becomes available (e.g. once a year), instead of every month or quarter. In the meantime, the original sub-annual series can be preliminarily benchmarked by adding (or multiplying) the predicted correction factors to the current sub-annual values. This situation is analogous to seasonal adjustment with the X-11-ARIMA method, which provides seasonal factor forecasts to be used for preliminary seasonal adjustment of current observations.

Preliminary benchmarking may actually be regarded as one of the implementational issues of benchmarking, examined in the next section.

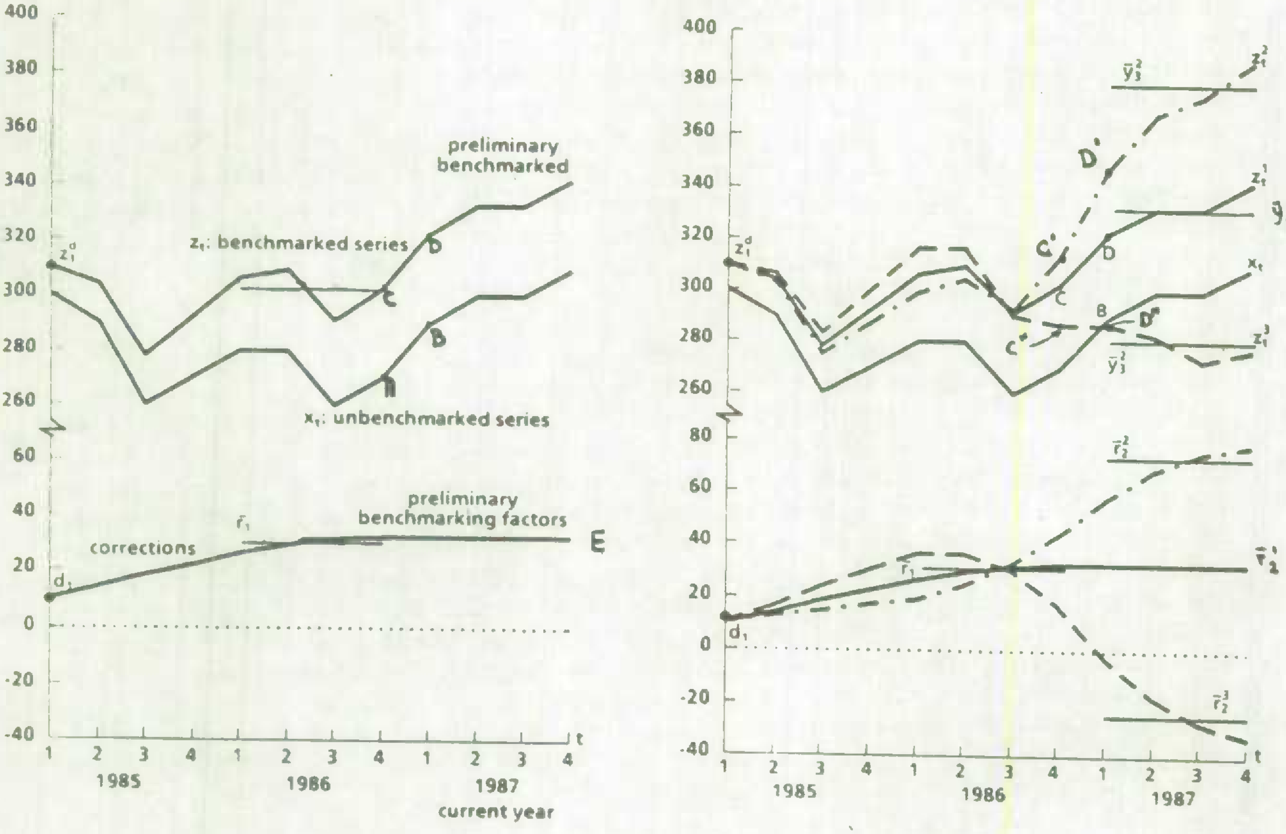


Figure 3.1: (a) Preliminary benchmarking according to the modified Denton method; (b) Real benchmarking under a realistic scenario and two extreme scenario

4. IMPLEMENTATION OF BENCHMARKING

The implementation of benchmarking is now considered. A natural way to implement any benchmarking method is to recalculate the whole benchmarked series from the sub-annual original series and from the benchmarks, each time a new sub-annual observation is available and a corresponding benchmarked value is desired: e.g. every month, every quarter. For a series of 120 observations, benchmarking with the Denton approach would then involve the inversion of matrices of 120 rows by 120 columns, which is a substantial computational problem. The preliminary benchmarking strategy explained in the previous section allows one to carry out benchmarking only when a new benchmark becomes available, typically once a year. (In the meantime predicted preliminary benchmarking factors are used.) However, benchmarking under the movement preservation principle can be further alleviated 1) by taking advantage of the properties of the method and 2) by resorting to numerical approximations to the formal minimization process described in Section 2.

4.1 Simplification based on the properties of the method

With most benchmarking methods, the introduction of new years of data (i.e. of new original and benchmark data) has no impact on the estimates pertaining to the distant past. Put differently, the incorporation of the new observations causes negligible change in the past benchmarked values lying far away from the end of the series. Of course this does not hold if past sub-annual values or past benchmarks are modified when the new year is incorporated. But since, in practice, statistical agencies leave data untouched after a certain number of years (e.g. after 3 years), the past benchmarked values do eventually stabilize. In order to facilitate the analysis, however, it is first assumed that the only change occurring from year to year is the incorporation of the extra year of sub-annual values x_t and of benchmark(s) y_m or $z^d{}_t$.

Under such a pattern, the benchmarked values of a year i, derived by the modified Denton method, stabilize in year i+3. This is illustrated in Figure 4.1. The proportional corrections z_t/x_t pertaining to 1982 are the same in 1985 (dotted curve) as in 1986 (solid curve). Indeed, the 1982 values of the two curves are undistinguishable. The two corresponding benchmarked values are therefore practically identical. In other words, the benchmarked values of 1982 stabilized in 1985 (i.e. when the 1984 benchmarak is incorporated). The two correction curves of 1983 are almost the same in 1985 (dotted curve) as in 1986 (solid curve) and will coincide in 1987. The benchmarked values of 1983 consequently stabilize in 1986. This built-in stabilization of estimates is faster under erratic discrepancies; and slower, under monotonic discrepancies. With that stabilization, it is useless to recalculate distant past benchmarked values year after year. It is sufficient to calculate only the last few years, e.g. 5 or 3 years.

The modified Denton method allows for further alleviations of the benchmarking operation. As illustrated in Figure 2.1 and 2.3, binding sub-annual benchmarks can specify the starting values of the benchmarked series. In particular, the benchmarked series may be specified to start from values considered as historical or final by the statistician. In Figure 4.1, the corrections curve calculated in 1986 ("z_t/x_t in 1986") is forced to start from one sub-annual discrepancy in the fourth quarter of 1981. That discrepancy corresponds to an historical value of the benchmarked series. (The values of 1981 and before are historical.) Under such a scheme, proposed by Laniel (1986) and Baldwin (19??), one needs (re)calculate only the last four complete years of the series, from 1982 to 1985 in the figure. (The 1986 corrections are a trivial repetition of the last 1985 correction and do not actually require calculation.) In 1987, when the 1986 data are incorporated, the sub-annual benchmark is moved to the fourth quarter of 1982, and only years 1983 to 1986 is (re)calculated.

Now assume that on incorporating one benchmark, the benchmarks and the sub-annual series of the two previous years are potentially modified. The benchmarked series then takes two extra years to stabilize. Consequently the annual benchmarking exercise need involve the six last years of the series. This generally translate into important reduction in the scale of the computations. Yet massive reductions are still achievable through algebraic and numerical approximations of the formal minimization which underlies benchmarking.

Before discussing approximation, the following point deserves to be made. On incorporating a benchmark pertaining to one year, at least the preceeding year must be revised. Figure 4.2 depicts the previously benchmarked series accompanied by the new benchmarked series. Failling to carry out the revision of 1986 (shaded area) amounts to publishing discontinuity CB instead of the correct movement AB. (Note that CD was more correct than CB.) Point E in the figure represents a sub-annual benchmark specified so that 1985 is not modified. The warning, which is illustrated for preliminary benchmarking, i.e. for the end of the series, also holds when modifying an already available past benchmark value: the year preceeding that value must be revised, if no discontinuity is to be introduced in the series.

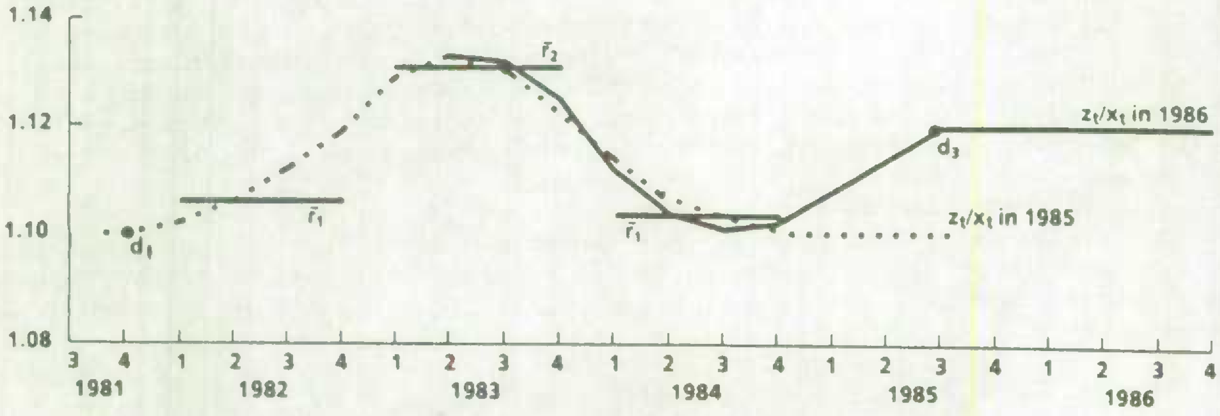


Figure 4.1: Built-in Stability of the corrections and therefore of the benchmarked series after a few years

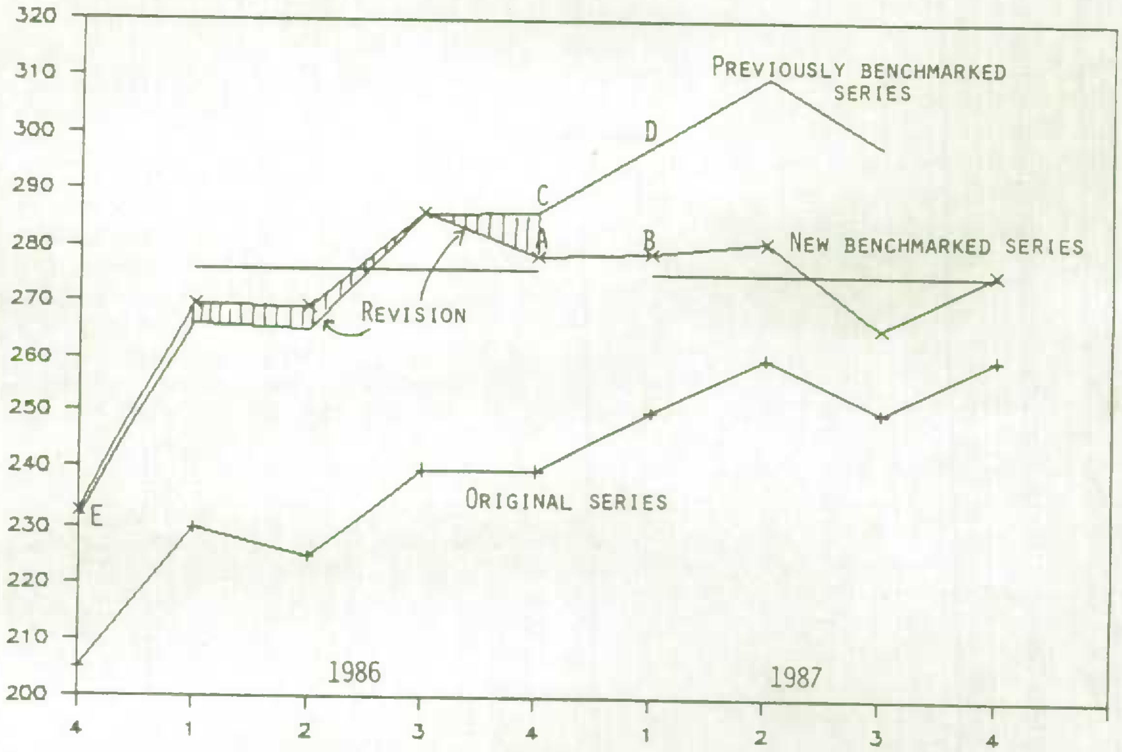


Figure 4.2: Revisions in the past benchmarked values required by the incorporation of a new benchmark or by the modification of existing benchmarks

4.2 Numerical approximations to the formal minimization

As explained in Section 2, benchmarking according to the modified Denton method is based on the principle of movement preservation of the original series; and, in order to achieve that goal, the corrections are kept as smooth and constant as possible. This suggest benchmarking may be attempted from the discrepancies. The approach proposed here is to iteratively fit smooth and flat corrections curves to the discrepancies and adjust those curves with simple arythmetic tools, until the discrepancies are totally allocated, i.e. until the benchmarking constraints are satisfied. This is analogous to iterative ratio-to-moving average fitting in the X-11-ARIMA series component estimation method (Dagum, 1980).

4.2.1 Approximation for stock series

In the case of stock series (as defined in Section 1.2) with binding benchmarks, an exact approximation is readily available. The following algebraic formulae are those of the corrections c_t obtained by formally minimizing (2.1) or (2.3):

(4.1) $$c_t = d_k + (t-t_k) * [(d_{k+1} - d_k)/(t_{k+1} - t_k)], \qquad t_k \leq t \leq t_{k+1},\ k=1,..,K-1$$

(4.2) $$c_t = d_1, \qquad t \leq t_1$$

(4.3) $$c_t = d_K, \qquad t \geq t_K$$

where d_k and t_k stand for the sub-annual additive or proportional discrepancies and their reference periods respectively. (For stock series, the annual benchmarks are technically sub-annual benchmarks, hence notation d_k instead of r_m.) Variable K represents the number of discrepancies, i.e. of benchmarks. The corrections obtained from (4.1) are displayed in Figure 4.3. They are linear interpolation between the discrepancies on each side of a time period considered. The ordinate and the slope of each line are d_k and $[(d_{k+1} - d_k) / (t_{k+1} - t_k)]$. For the periods preceeding the first and following the last discrepancies, the corrections simply repeat the first and last discrepancies. This scheme would also work for flow series which have sub-annual but no annual benchmarks. However, such flow series are not likely to occur in practice.

4.2.2 Approximation for Flow Series

For flow series with binding benchmarks, the approximation is not exact and requires iterations. However, it is based on the same approach as for stock series. The steps required are simple but intricate. They are documented in Appendix D.

Figure 4.4 constrast the approximated and the exact corrections obtained by formal minimization of the objective function. Even in the unfavourable benchmarking situation depicted, the approximated are very close to the exact corrections. It could be argued that the preliminary benchmarking factors resulting from the approximation (for periods 16, 17, ...) are preferable. Indeed the implicitly anticipated discrepancy (see Section 3) is more "conservative". For a more favourable benchmarking situation, the approximated and exact corrections would be even closer.

4.2.3 Approximation for Second Difference

In some applications of benchmarking, but especially of interpolation (see Part 2, Section 1), the objective function minimizes second differences in the corrections. The resulting corrections behave smoothly and as linearly as possible. The approximation described in Sections 4.2.1 and 4.2.2 remains valid with one very little change. Whenever values of discrepancies or corrections are repeated at the start or at the end of series, the corrections are linear extrapolations of the corrections calculated for the central part of the series. For instance equations (4.2) and (4.3) repsectively become

$$c_t = c_{t1} + (t-t_1)*[c_{t1+1} - c_{t1}] \qquad t \leq t_1$$

(read subscript t1 and tK as t_1 and t_k)

$$c_t = c_{tK} + (t-t_K)*[c_{tK+1} - c_{tK}] \qquad t \geq t_K$$

For flow series, that substitution is also made in equations (3") and (4") of Appendix D (mutatis mutandis).

Coding to perform benchmarking according to the proposed approximation with programme SAS/IML may be obtained from the author. The reduction in the amount of calculations and of required computer memory afforded by the approximation is massive. As shown in Appendix B and C, formally minimizing objective function (2.1) or (2.3) over five years of monthly data involves multiplicating and inverting matrices of dimensions 60 by 60. With the numerical approximation, the benchmarking operation can efficiently be carried out on micro-computers. The operation may also be so cheap as to make the implementational simplifications proposed in Section 4.1 unnecessary. In other words, it may be logistically easier and more feasible to recalculate distant past values of series year after year (i.e. to run benchmarking over the whole series), like this is done for seasonal adjustment with the X-11-ARIMA method.

In the case of non-binding benchmarks, the approximation simply consists of fitting a smooth curve through the discrepancies with no concern to allocate the discrepancies. The curve should only capture the local level of the discrepancies. One approach would consist of modifying the benchmarks until they yield monotonic discrepancies, of considering the modified benchmarks as binding and of applying the approximation just presented. (We have nothing more specific to propose at this stage. The difficulty was with binding benchmarks, not with non-binding.) Figure 2.2 and 2.4 provide examples of such curves. This strategy applies both to stock and to flow series. The numerical approximation in non-binding benchmarks situation also circumvents the problem of chosing the weights g^x, ${g^y}_m$ and ${g^z}_k$ in the objective function. (Maybe, the choice of those weights is a non-issue.)

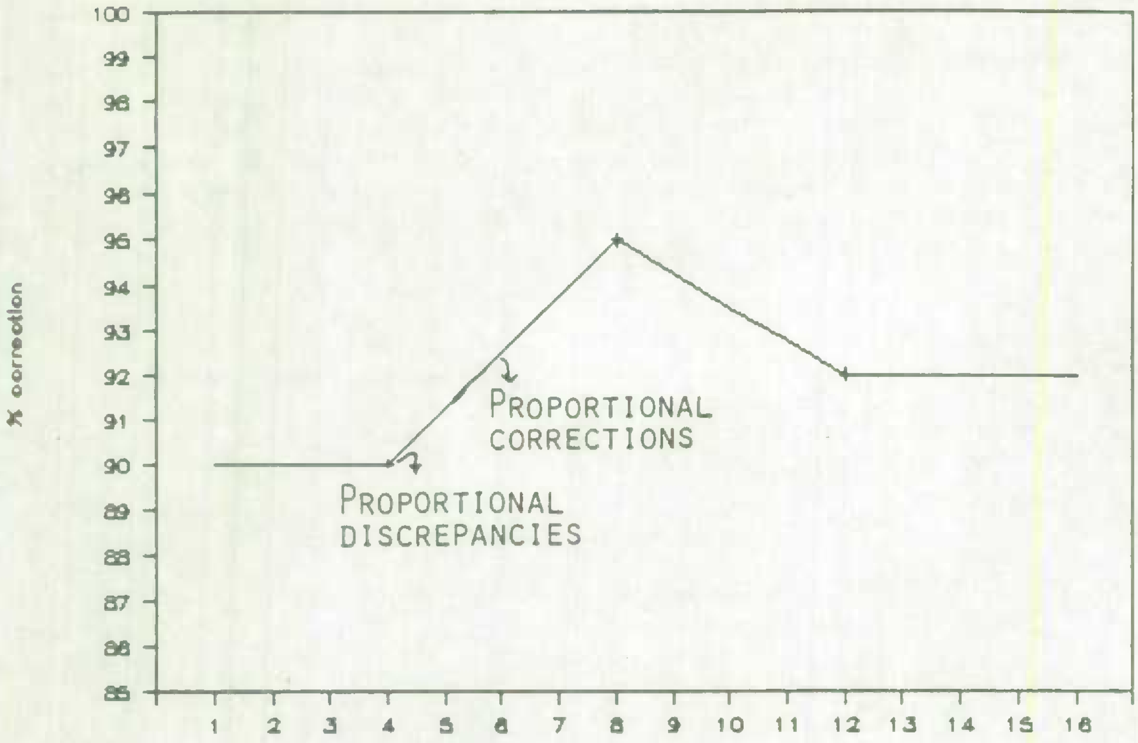


Figure 4.3: Proportional corrections for stock series. The approximation is exact for stock series.

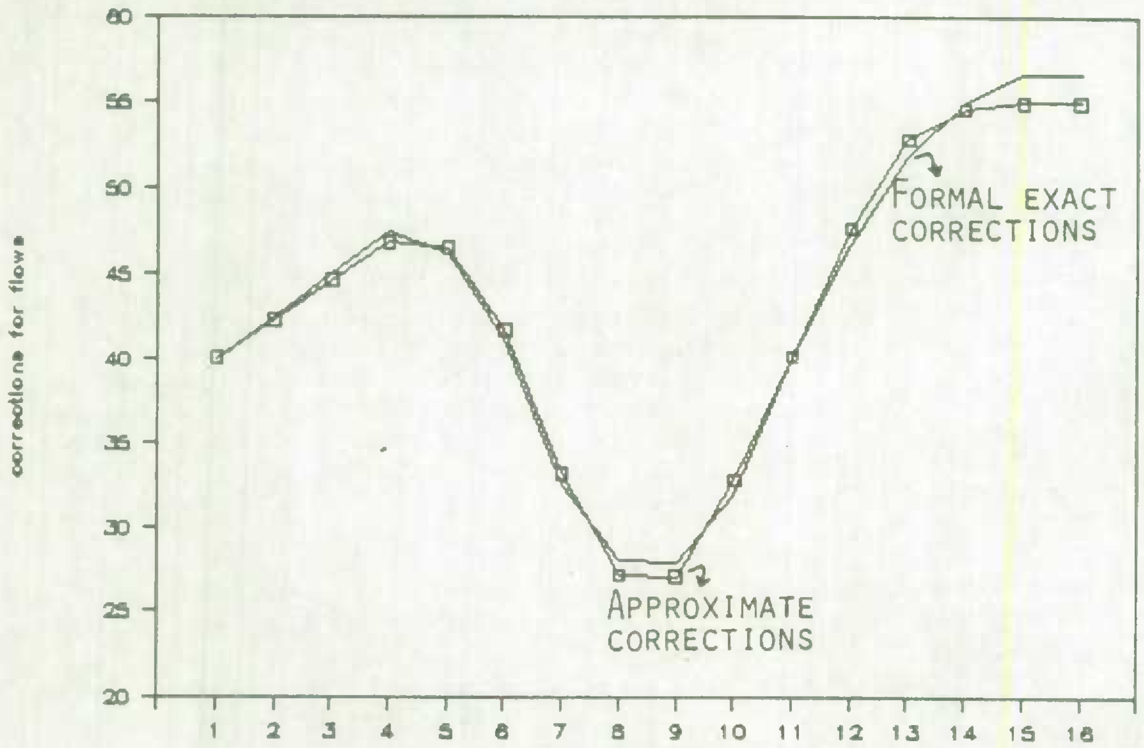


Figure 4.4: Approximated and exact corrections for flow series

5. BENCHMARKING SYSTEMS OF SERIES

As explained in Section 1.6, situations occur where identities must be preserved between series for each period of time. First, one-way classification cases are examined, where the series are components to one aggregate, for instance wholesale trade by industry with respect to the total of all industries; second, two-way classification cases, for instance wholesale trade by region and by industry with respect to the totals of all industries and of all regions and with respect to the grand total of all industries and all regions.

5.1 Benchmarking in situations of one-way classification

Consider the case of N industrial original sub-annual series x_{1t}, $x_{2t}, \ldots,$ x_{Nt}, $t=1,\ldots,T$, in which the Nth series is the total of the other N-1 series for each period of time. The N benchmarked series must then satisfy T binding constraints:

$$(5.1.1) \qquad z_{1,t} + z_{2,t} + \cdots + z_{N-1,t} = z_{N,t}, \quad t=1,\ldots,T$$

Each series must (say) satisfy its own annual and sub-annual benchmarks, $y_{n,m}$ and $z^d_{n,tk}$:

$$(5.1.2) \qquad \sum_{\tau_{n,m}}^{\rho_{n,m}} z_{n,t} = y_{n,m}, \quad n=1,\ldots,N;\ m=1,\ldots,M_n$$

$$(5.1.3) \qquad z_{n,tk} = z^d_{n,k}, \quad n=1,\ldots,N;\ k=1,\ldots,K_n, \quad \text{(read tk as } t_k\text{)}$$

where M_n and K_n respectively stand for the number of annual and sub-annual benchmarks for series $z_{n,t}$. The notation allows each of the N series to have annual benchmarks with different reference periods, $[\tau_{n,m},\ldots,\rho_{n,m}]$. The notation makes it possible to denote the sub-annual benchmarks by means of $y_{n,m}$ with the appropriate reference periods $\tau_{n,m}$ $(=\rho_{n,m})$. In other words $y_{n,m}$ may be used to denote benchmarks whether annual or sub-annual. For the sake of simplicity and alleviation, the sub-annual benchmarks $z^d_{n,k}$ are henceforth dropped.

One approach with a system of series is to benchmark each component separately to its benchmarks and to define the aggregate benchmarked series as the sum of the benchmarked components. The analogous practice in seasonal adjustment is called indirect adjustment. On comparison to its original sub-annual values $x_{N,t}$, the resulting indirectly benchmarked aggregate series $z_{N,t}$ may be unacceptable. For instance, it may display movements which were not present in the original series and which are ruled as highly improbable by the series expert.

5.1.1 Additive Simultaneous Specification

A direct approach to systems of series, adopted by Taillon (1987) and pursued in this section, is to benchmark the N series simultaneously. The principle of movement preservation is specified on both the components and the aggregate. This leads to the following global objective function which also incorporates the constraints between series:

$$f(z) = \sum_{n=1}^{N} g^x{}_n \sum_{t=2}^{T} ((z_{n,t} - x_{n,t}) - (z_{n,t-1} - x_{n,t-1}))^2$$

$$(5.1.4) \qquad + \sum_{n=1}^{N} \sum_{m=1}^{M_n} g^y{}_{n,m} \left(\left(\sum_{t=\tau_{n,m}}^{\rho_{n,m}} z_{n,t} \right) - y_{n,m} \right)^2$$

$$+ \; G \sum_{t=1}^{T} \left(\left(\sum_{n=1}^{N-1} z_{n,t} \right) - z_{N,t} \right)^2$$

The first term specifies the movement preservation principle for each of the N series; the second term, the annual (and the sub-annual) benchmark constraints of each series. The third term specifies the aggregation constraint (5.1.1) which must prevail between series for each period of time. The weights $g^y{}_{n,m}$ specify the relative importance of the various benchmarks. High weights imply binding benchmarks; and low weights, non-binding benchmarks. The weights $g^x{}_n$ standardize the components in order to make them equally important in the objective function. (In the absence of standardization, the relatively small series, i.e. with small absolute values, could be corrected by relatively large amounts, potentially changing the sign of the series. This would indeed be likely to happen when some of the other large-valued series display large and erratic discrepancies. The movement preservation criterion of the small series would have little importance in the objective function and could therefore be violated with little penalty.)

5.1.2 Proportional Simultaneaous Specification

In the additive model (5.1.4), the standardization and the choice of weights is very intricate, because of the possibly different orders of magnitude of the various series. The proportional variant is much more appropriate for simultaneous benchmarking, because all terms are expressed in percentages (more precisely in ratios):

$$f(z) = \sum_{n=1}^{N} g^x{}_n \sum_{t=2}^{T} ((z_{n,t} / x_{n,t}) - (z_{n,t-1} / x_{n,t-1}))^2$$

$$(5.1.5) \qquad + \sum_{n=1}^{N} \sum_{m=1}^{M_n} g^y{}_{n,m} \left(\left(\left(\sum_{t=\tau_{n,m}}^{\rho_{n,m}} z_{n,t} \right) / y_{n,m} \right) - 1.0 \right)^2$$

$$+ \; G \sum_{t=1}^{T} \left(\left(\sum_{n=1}^{N-1} z_{n,t} \right) - z_{N,t} \right)^2$$

In some situations with potentially legitimate negative values for some components series, e.g. profits, such components may be specified additively and standardized through $g^x{}_n$. Assuming no negative components and weights $g^x{}_n$ equal to 1, the movement preservation is equally important for all components. Higher weights may be chosen for the more reliable components. For instance a weight $g^x{}_N$ equal to N-1 specifies that movement preservation is as important for the aggregate $z_{N,t}$ as for all the other component series collectively. Similarly, by means of $g^{y}{}_{n}$, it is possible to weight the more reliable benchmarks, e.g. those of the aggregate, more than the unreliable ones. (Note that if the benchmards already satisfy the aggregation constraints, they will tend to reinforce each other across series. In other words, benchmarks specified as binding for one series will tend to make the benchmarks of the other series binding.)

One statistical advantage of simultaneous benchmarking is that it operates as some kind of accounting framework to cross-validate the benchmarked series. For each period t, the sales (say) must be exhaustively allocated to the industries (or regions, etc.). Furthermore, through the constraints between series, the higher reliability of some components improves that of other components. Another advantage, of course, is to produce components which add up to the aggregate.

(Note that series $z_{N,t}$ may be replaced by the sum of the N-1 other series in both (5.1.4) and (5.1.5). Proceeding with Langrangian multipliers would require such a substitution in order to avoid singularities. The last term of the objective function would then disappear but the remaining terms would be substantially complicated. For the sake of clarity and simplicity, such aggregate series are explicitely kept in the equations; and this will also be the case in the next sub-sections.)

5.1.3 Computational Strategies

Simultaneously benchmarking several series by means of (5.1.5) or (5.1.4) is mathematically possible and straightforward. Computationally however, it is problematic. (So far we have successfully processed 25 quarterly series on 3-year intervals with the standard main computer version of package SAS/IML.) Simultaneously processing 30 monthly series on 3-year (moving) intervals requires the manipulation and the inversion of matrices with dimensions 1080 by 1080, which is probihitive (on an operational basis at least) for the computers available today.

In order to overcome that problem, one strategy is to process the N monthly series jointly but in two steps. The first step consists of benchmarking on a quarterly basis. The monthly original series are collapsed into quarterly series and benchmarked simultaneously. Processing 30 quarterly series on 3-year intervals thus requires the manipulation and inversion of matrices with dimension 360 by 360, instead of 1080 by 1080, which reduces the order of magnitude of the computations. The second step consists of simultaneously benchmarking the original monthly series to the quarterly values obtained in the first step and now considered as "annual" benchmarks. This step involves matrices of 270 by 270.

The strategy could also consist of three steps, the first specifying the problem on semi-annual series (involving matrices of 180 by 180); the second step, on quarterly basis (180 by 180); and the third, on a monthly basis (270 by 270). The latter strategy requires the reference periods of all the annual benchmarks to coincide with the conventional year or with the end of semesters; and the first strategy, with the conventional year or the end of quarters. One problem with that multi-step approach is the possible presence of monthly (sub-annual) benchmarks.

Another strategy to overcome the magnitude of the problem is to collapse the N series $x_{n,t}$ into a much lower number N' of sub-aggregate series $x'_{n,t}$, to carry out benchmarking at that level of sub-aggregation and then to benchmark at the level of the individual series within each sub-aggregate. For instance, processing 30 original monthly industrial series $x_{n,t}$ is done in the following two steps. First the 30 series are collapsed into 6 sub-aggregate series $x'_{n,t}$. These are the sum of 4 to 6 individual series, except for the 6th series which coincides with the overall aggregate ($x'_{N',t}=x_{N,t}$). (The corresponding collapse of the benchmarks is done and this implies homogenoeous reference periods.) Simultaneous benchmarking is carried out on the 6 sub-aggregate series by means of (5.1.5). Over 3-year intervals, the matrix involved have dimensions 216 by 216 (instead of 1080). The second step is to simultaneously benchmark the 4 to 6 individual series within each sub-aggregate under the constraint that all the individual series sum to the sub-aggregate values $z'_{n',t}$ established in the first step. This can be accomplihed by replacing the last term of objective function (5.1.5) by

$$(5.1.6) \qquad + \; G \sum_{t=1}^{T} \left(\left(\sum_{n=1}^{N} z_{n,t} \right) / \; z'_{n',t} \right) - 1.0)^2$$

which embodies the required constraint. This second step involves matrices of dimensions 144 by 144 or 180 by 180 or 216 by 216 depending on the number of series in the sub-aggregate considered being 4, 5 or 6.

One statistical advantage of this second strategy is to take advantage of the higher reliability of the input series at higher levels of aggregation. Furthermore, the movements derived at the more reliable higher levels of aggregation are thus imposed to lower levels. In the same example, the movement in the aggregate benchmarked series and the 5 sub-aggregate series are determined jointly: Those estimated for the 5 sub-aggregates influence those of the aggregate and vice versa. Once the movements in the sub-aggregates are thus determined, the second step imposes those movements on the series within each sub-aggregate. Indeed, the sub-aggregate values - and the corresponding movements - are specified as constraints (in (5.1.6)).

In this second strategy, the collapsing is performed over the components, instead of over time in the first strategy. In that sense, both strategies are similar. They may actually be combined, i.e. the series may be collapsed both over time and components.

5.2 Benchmarking in situations of two-way classification

Now consider a system of series pertaining to R regions and N industries (say). As illustrated in Table 5.1, the original values $x_{r,n,t}$ measure the wholesale trade (say) for region r and for industry n at time t. (There are in fact T tables like 5.1. under consideration.) Assume that $x_{r,N,t}$ are the regional totals over the industries (last column of table); $x_{R,n,t}$, the industrial totals over the regions (last row); and $x_{N,R,t}$, the grand total over both industries and regions, e.g. the "national" total (last column and row). The last industry "N" and the last region "R" are thus the totals for the regions and for the industries respectively.

Table 5.1: Illustration of a two-way classification by region r and by industry n, for a given moment of time t

$\underline{t}$	Industry n				Regional totals over industries $x_{r,N,t}$
	$x_{1,1,5}$	$x_{1,2,t}$	. . .		$x_{1,N,t}$
R	$x_{2,1,t}$	$x_{2,2,t}$	. . .		$x_{2,N,t}$
e					
g	.	.			.
i	.	.			.
o	.	.			.
n				$x_{r,n,t}$	
r					
industrial totals over regions $x_{R,n,t}$	$x_{R,1,t}$	$x_{R,2,t}$	. . .		$x_{R,N,t}$ "National" total

5.2.1 Proportional Specification

The same approach and strategies described in the previous sub-section may be used. The aggregation constraints are now be defined for both the industrial and regional break-downs

$$\sum_{r=1}^{R-1} z_{r,n,t} = z_{R,n,t}, \quad n=1,\ldots,N-1; \; t=1,\ldots,T \tag{5.2.1}$$

$$\sum_{n=1}^{N-1} z_{r,n,t} = z_{r,N,t}, \quad r=1,\ldots,R-1; \; t=1,\ldots,T \tag{5.2.2}$$

The sum of regional totals must also be equal to the sum of the industrial totals, for each period of time:

$$\sum_{r=1}^{R-1} z_{r,N,t} = \sum_{n=1}^{N-1} z_{R,n,t}, \quad (=z_{R,N,t}) \; t=1,\ldots,T \tag{5.2.3}$$

(That constraint may be made implicit in (5.2.2) by letting index r therein reach R instead of R-1; or in (5.2.1), by letting index n reach N.)

The principle of movement preservation is specified on the R regional series and on the N industrial series. The proportional objective function is then:

$$f(z) = \sum_{r=1}^{R} \sum_{n=1}^{N} g^{x}_{r,n} \sum_{t=2}^{T} \{(z_{r,n,t} / x_{r,n,t}) - (z_{r,n,t-1} / x_{r,n,t-1})\}^2$$

$$+ \sum_{r=1}^{R} \sum_{n=1}^{N} \sum_{m=1}^{M_{r,n}} g^{y}_{r,n,m} \{((\sum_{t=\tau_{r,n,m}}^{\rho_{r,n,m}} z_{r,n,t}) / y_{r,n,m}) - 1.0\}^2$$

(5.2.4)

$$+ G \sum_{t=1}^{T} \sum_{n=1}^{N-1} [(\sum_{r=1}^{R-1} z_{r,n,t}) - z_{R,n,t}]^2 + G \sum_{t=1}^{T} \sum_{r=1}^{R-1} [(\sum_{n=1}^{N-1} z_{r,n,t}) - z_{r,N,t}]^2,$$

$$+ G \sum_{t=1}^{T} [(\sum_{r=1}^{R} z_{r,N,t}) - (\sum_{n=1}^{N} z_{R,n,t})]^2$$

Mathematically, it is then possible to generalize the approach of Section 5.1 to any level of classification (and to apply the same collapsing strategies), by incorporating the appropriate constraints and criteria into the objective function.

5.2.2 Computational Strategies

Another approach may be more appropriate in situations where statistical agencies publish only the regional totals $z_{r,N,t}$ and the industrial totals $z_{R,n,t}$, that is the marginal totals of Table 5.1. The other series inside Table 5.1 are available to the statistical agency but not released to the public - as time series at least -, because confidential or not sufficiently reliable, etc. The total number of series to consider may therefore be reduced to R+N-1 instead of R*N. For instance with 6 regions (R=6) and 11 industries (N=11), including the totals, the number of series to benchmark is 16 instead 66. Keeping the same notation, the objective function (5.2.4) becomes:

$$f(z) = \sum_{r=1}^{R} g^{x}_{r,N} \sum_{t=2}^{T} \{(z_{r,N,t} / x_{r,N,t}) - (z_{r,N,t-1} / x_{r,N,t-1})\}^2$$

$$\sum_{n=1}^{N} g^{x}_{R,n} \sum_{t=2}^{T} \{(z_{R,n,t} / x_{R,n,t}) - (z_{R,n,t-1} / x_{R,n,t-1})\}^2$$

$$+ \sum_{r=1}^{R} \sum_{m=1}^{M_{r,N}} g^{y}_{r,N,m} \{((\sum_{t=\tau_{r,N,m}}^{\rho_{r,N,m}} z_{r,N,t}) / y_{r,N,m}) - 1.0\}^2$$

$$
(5.2.5) \qquad \begin{aligned}
&+ \sum_{n=1}^{N} \sum_{m=1}^{M_{R,n}} g^{y}_{R,n,m} \{((\sum_{t=\tau_{R,n,m}}^{\rho_{R,n,m}} z_{R,n,t}) / y_{R,n,m}) - 1.0\}^2 \\
&+ \sum_{t=1}^{T} [(\sum_{r=1}^{R} z_{r,N,t}) - (\sum_{n=1}^{N} z_{R,n,t})]^2
\end{aligned}
$$

The first two terms specify the movement preservation principle on the regional and on the industrial series respectively; the third and fourth terms, the annual benchmark constraints; and the fifth term, the overall aggregation constraint.

One statistical advantage of this approach is to take advantage of the higher reliability of the input series at higher levels of aggregation, namely at the level of the totals.

Like in Section 5.1, the benchmarking problem (5.2.5) may be first specified on a quarterly basis; and then on a monthly basis using the benchmarked quarterly values as "annual" benchmarks. Similarly, regions and industries may be collapsed into sub-aggregates to be later dis-aggregated.

Minimizing objective function (5.2.5) yields benchmarked values $z_{R,n,t}$ and $z_{r,N,t}$ for the totals, that is for the margins (last row and last column) of Table 5.1, for all moments of time t (t=1,...,T). If necessary, the other series, $x_{r,n,t}$ (for r<R and n<N) inside Table 5.1, may be adjusted to sum to the row and to the column totals. This may be achieved by Iterative Proportional Fitting or "raking" (Bishop, Fienberg and Holland, 1975; Brackstone and Rao, 1979). In the context of Table 5.1 raking consists of adjusting each column to the its total, by multiplying it by the proportion of the corresponding desired and actual totals; of adjusting each row to its total in the same manner; of re-adjusting the columns; of re-adjusting the rows; and so on. This method has been proven to converge exactly.

In the context of benchmarking, there are problems with the 2-dimensional raking just described. First, it is not likely to preserve the period-to-period movement of the original series $x_{r,n,t}$ (for r<R and n<N), because each time period, i.e. each of the T tables like Table 5.1, is processed separately. If such discontinuities are to be avoided, objective function (5.2.5) may be reverted to, probably with the collapsing strategies.

Second and for the same reason, the 2-dimensional raking just described will not satisfy the benhmark $y_{r,n,m}$ of the components series $x_{r,n,t}$ (r<R, N<N). However this second problem may be corrected by 3-dimensional raking, processing one year (of tables like 5.1). The two first dimensions consist of the regional and the industrial totals $z_{R,n,t}$ and $z_{r,N,t}$; and the third dimension, of the benchmarks $y_{r,n,m}$ for a given year. (This implies the availability of benchmarks every year for every series.) The problem remaining is the possibility of movement discontinuities between years,

since each year of values $x_{n,r,t}$ is processed separetaly. However, this method may represent a trade-off between quality and feasibility which may provide sufficient quality in many situations.

Practical experience with simultaneous benchmarking may show that very similar results can be achieved with some combination of individual benchmarking with raking. If such were to be the case, the methods proposed in this section should not be applied. However, simultaneous benchmarking does provide a standard, i.e. a norm, against which alternative and simpler approaches may be assessed.

5.3 Other Problems with System of Series

In many situations, the source of discrepancy between the components and an aggregate - even after benchmarking - is rounding and not the lack or the faillure of benchmarking. For instance adding 33, 23 and 63 should give 119. However, the computer may yield 120, because internally (in its memory) the numbers were actually 33.3263, 23.3919 and 63.4019. If there are 100 components, that kind of discrepancy may distribute between -50 and 50. This phenomenon can also explain the residual discrepancies left after benchmarking, between the benchmarks and the annual sums of the corresponding benchmarked series.

One approach to solve that problem is to distribute the rounding discrepancies to the values with the largest fractions. The distribution should take place over the components. However this would disturb the benchmark constraints. The most pratical attitude is probably to tolerate the benchmark constraint disturbances. The yearly financial reports of private corporations usually footnote rounding discrepancies.

Simultaneous benchmarking may also be complicated by the fact that some components have no benchmarks. Those components may still be specified to maximize movement preservation. If a component is the only one without benchmarks, those benchmarks are implicit in the those of the other series. If several series are without benchmarks, the danger is that their level may be changed drastically. In order to avoid this, it is possible to specify that such series have corrections close to those of another component series. The following term is then added to the objective function:

$$h_u \sum_{t=1}^{T} (z_{u,t} / x_{u,t} - z_{s,t} / x_{s,t})^2 \qquad (5.3.1)$$

where the u-th component is the one adopting the corrections of the s-th component. Depending on the value of h_u, that term can be made binding or non-binding. The term can be used as a substitute for the movement preservation criterion for the u-the component (by setting $g^x{}_u=0$) or in conjonction with it. It is also possible to specify that the uth components adopt the corrections of several other components, by incorporating several terms like (5.3.1) for the uth component.

5.4 "Small Area" Data

Consider a situation with R=6 regions and N=10 industries. The sixth region and the tenth industry are respectively the industrial and the regional totals and correspond to the margins of Table 5.1. The 15 marginal total series are reliable. However, the 45 component series by region and industry inside Table 5.1 display erratic movements, more specifically, grossly unacceptable growth rates. The conventional approach is to correct each of the 45 components separately using subject matter expertise. However, this destroys the consistency of the system, i.e. the rows and columns of Table 5.1 do not sum to the margins; and those sums may now also display unacceptable movements. Another rouond of such tedious and laborious adjustment is done over until an acceptable solution is reached.

This type of situation may be addressed by a variant of simultaneaous benchmarking. That variant integrates the subject matter expertise about each individual component series and combines it with the regional and industrial additivity constraints. Assume that the integration exercise takes place on 3 years of annual data, I=3, J=1 and T=3. (The resulting series may be temporally dis-aggregated in another step.) Consider the following objective function:

$$f(z) = \sum_{r=1}^{R} \sum_{n=1}^{N} g^x{}_{r,n} \sum_{t=2}^{T} ((z_{r,n,t} / x_{r,n,t}) - 1)^2$$

$$\sum_{r=1}^{R-1} \sum_{n=1}^{N-1} g^x{}_{r,n} \sum_{t=2}^{T} ((z_{r,n,t} - r_{r,n,t} \;\; z_{r,n,t}) / \sigma_{r,n,t})^2$$

$$+ \sum_{r=1}^{R-1} \sum_{n=1}^{N-1} \sum_{k=1}^{K_{r,n}} g^y{}_{r,n,m} ((z_{r,n,tk}) / y_{r,n,k}) - 1.0)^2$$

(5.4.1) (read subscript tk as t_k)

$$+ G \sum_{t=1}^{T} \sum_{n=1}^{N-1} [(\sum_{r=1}^{R-1} z_{r,n,t}) - z_{R,n,t}]^2 + G \sum_{t=1}^{T} \sum_{r=1}^{R-1} [(\sum_{n=1}^{N-1} z_{r,n,t}) - z_{r,N,t}]^2,$$

$$+ G \sum_{t=1}^{T} [(\sum_{r=1}^{R} z_{r,N,t}) - (\sum_{n=1}^{N} z_{R,n,t})]^2$$

The first term of the objective function specifies that the desired series $z_{r,n,t}$ are to be somewhat close to the data $x_{r,n,t}$ available for the regions and the industries. In other words, the available data should be taken into account in deriving the desired values. The importance of the data is determined by weights $g^x{}_{t,n,t}$. (For the totals the weights could be binding.)

The second term, the growth rate criterion, specifies that the components series (inside Table 5.1) behave according to growth rates $r_{r,n,t}$ from year to year. Theses growth rates embody the subject matter expertise. The components were ruled unacceptable on the basis of their actual growth rates (in $x_{r,n,t}$). The subject matter accordingly specifies more acceptable growth rates. Growth rates equal to 1 (which may be good starting values) specify that the desired series $z_{r,n,t}$ behave smoothly from on period to the next. (More information about growth rates is available in Section 2 of Part 2.) The standard deviations of $x_{r,n,t}$, $\sigma_{r,n,t}$, standardizes the components so that the (growth rate criteria of) small components receive the same importance as large components in the objective function.

The subject matter expertise may also be embodied in benchmarks $y_{r,n,t}$ chosen for individual series and individual periods ($K_{r,n} \leq T$). Such benchmarks are appropriate to specify the starting point (e.g. "historical") of all the component series.

The last three terms specify the aggregation constraints over regions and industries.

With six regions and ten industries on 3-year annual interval, the solution requires the manipulation of matrices of dimensions 180 by 180. If needed, sub-annual values may be obtained by dis-aggregating the annual values specified as annual benchmarks, as described in the previous sub-sections.

6. THE ASSESSMENT OF BENCHMARKING

As explained in Section 1, the benchmarking situation may be assessed by the behaviour of the discrepancies, before benchmarking is attempted. The situation is favourable when the discrepancies are constant or evolve monotonically. Some of the measures of monotonicity developped by Raveh (1986) might be used to quantify that statement.

The same approach could be used to assess the result of the benchmarking process. The monotonicity statistics would then apply to the corrections, that is to the proportional or additive differences between the benchmarked series and the original sub-annual series. Erratic corrections may suggest unreliable benchmarked series.

More specific statistics are in order however. Since the benchmarking methods proposed are based on the principle of movement preservation, a natural statistic is the **average absolute proportional movement deviation** (i.e. non-preservation):

(6.1)
$$\sum_{t=2}^{T} |z_t/x_t - z_{t-1}/x_{t-1}|/(T-1)$$

and the corresponding additive variant for additive benchmarking. As pointed out by Laniel (1986), since proportional benchmarking may be viewed as an approximation to growth rate preservation, the **average absolute growth rate deviation** (i.e. non-preservation is also an appropriate statistic:

(6.2)
$$\sum_{t=2}^{T} |z_t/z_{t-1} - x_t/x_{t-1}|/(T-1)$$

Another relevant statistic is the **residual percentage discrepancy** with respect to the benchmarks:

(6.3)
$$[(\sum_{t=\tau_m}^{\rho_m} z_t) / y_m \ -1.0]*100, \ m=1,...,M$$

This statistic checks the fulfilment of the benchmarking constraints, whether annual or sub-annual. In situations with non-binding benchmarks however, the residual percentage discrepancies are also useful to further assess the benchmarking situation and the compatibility of the benchmarks with the original un-benchmarked series.

In the case of systems of series, statistics (6.1) to (6.3) are calculated for each component series to the aggregate and for the aggregate series. Furthermore, it is certainly appropriate to check the fulfillment of the aggregation constraints for each period of time. This may be achieved with the residual percentage aggregation discrepancy:

$$(6.4) \qquad \{[(\sum_{n=1}^{N-1} z_{n,t}) / z_{N,t}] - 1\}*100 \quad , \; t=1,\ldots,T$$

where n indicates the component and N the aggregate considered. (In case of two-way classification (6.4) is calculated for both classifications.)

Further research remains to be done on the assessment of benchmarking.

PART 2: INTERPOLATION

Benchmarking situations, examined in the first part of this document, are characterized by the availability of more frequent "sub-annual" and less frequent "annual" data about the socio-economic variable considered. Benchmarking consists of deriving a sub-annual series on the basis of the sub-annual data and of the annual data, referred to as benchmarks. The resulting sub-annual series consequently are based on some factual information at the sub-annual level, namely the sub-annual data.

Much of the data published by statistical agencies, however, are - to varying degrees - interpolations. The Canadian quarterly Capital Expenditure series and to some extent the annual Population estimates are interpolations. Interpolation situations are characterized by a lack of direct sub-annual data about the socio-economic variable considered. Interpolation consists of calculating sub-annual values from "annual" benchmarks. This process must therefore rely on assumptions about the sub-annual variable considered. Compared to series obtained through benchmarking, interpolated sub-annual series may be based on little factual information at the sub-annual level.

As will become obvious in this document, the frontier between interpolation and benchmarking is in fact arbitrary. Some cases of interpolation can indeed be viewed as special cases of benchmarking. Those cases will be considered first. At the other end of spectrum, interpolation may - at least in theory - involve no benchmarks at all. The values are then a pure estimations (from related series). This part of the document will progress from one end of the spectrum to the other.

1.TEMPORAL DIS-AGGREGATION AND CALENDARIZATION

This section considers cases of interpolation which are special - and sometimes trivial - applications of benchmarking: Benchmark data are temporally disaggregated into quarterly, monthly and even daily values. The resulting monthly estimates (say) can be re-aggregated into quarterly values. This dis-aggregation and re-aggregation process can be applied to correct the reference periods of data available on an unsuitable time basis; that is, to "calendarize". For example, it is possible to convert financial year data into conventional year estimates; to convert data pertaining to bundles of weekly data into monthly estimates; etc.

All variants considered may be implemented by means of the numerical approximation described in Section 4.2 of Part 1. (The formal presentation of the objective functions remains very useful to understand the various specifications.)

1.1 Interpolating Non-Seasonal Values Between Annual Benchmarks

It is technically possible to create a sub-annual series from annual benchmarks in the absence of any original sub-annual series or of any sub-annual related sub-annual variable. This problem can be casted as a trivial additive benchmarking problem in which the original series x_t is equal to zero. Objective function (2.1) of Part 1 then reduces to:

$$f(z) = g^x \sum_{t=2}^{T} (z_t - z_{t-1})^2 + \sum_{m=1}^{M} g^{y}{}_m \left(\left(\sum_{t=\tau_m}^{\rho_m} z_t \right) - y_m \right)^2 \tag{1.1}$$

$$\sum_{k=1}^{K} g^{z}{}_k (z_t - z^{d}{}_k)^2 .$$

The second and third terms are the benchmark satisfaction criteria. (See Section 2.1.1 of Part 1 for more details.)

The first term of (1.1) specifies that the required interpolated values are as gradual as possible. In situations where the benchmarks predictably behave monotonically (e.g. linearly, exponentially), second differences, instead of first differences, may be more appropriate. The objective function then reads:

$$f(z) = g^x \sum_{t=3}^{T} (z_t - 2\, z_{t-1} + z_{t-2})^2 + \sum_{m=1}^{M} g^{y}{}_m \left(\left(\sum_{t=\tau_m}^{\rho_m} z_t \right) - y_m \right)^2 \tag{1.1'}$$

$$\sum_{k=1}^{K} g^{z}{}_k (z_t - z^{d}{}_k)^2 .$$

The resulting interpolated values behaves as linearly as possible. Whenever in doubt about the predictable monotonic character of the series, first differences are to be used. The series obtained with second differences differs from that obtained with first differences mainly in the first and last years, where the latter series levels off as depicted in Figure 1.1. For the other years there is very little difference, as the behaviour of the series is mainly governed by the benchmark constraints.

Figure 1.1 displays a sub-annual series derived with first differences; that is, by means of objective function (1.1). Such interpolated values cannot have any seasonality. They correspond to the trend-cycle component of the series. However, the dating of the resulting turning points should not be taken seriously. The fact that a turning point occurs at time period 10 (2nd quarter of year 3) in the figure connot be interpreted as significant. The true turning point may in fact lie anywhere between periods 8 and 13. In order to interpolate values which display seasonality and possibly precise trend-cycle movements, one needs a (non-trivial) original sub-annual series or indicator x_t.

The yearly values displayed pertain to financial years ranging from the second to the first quarter of the following year. The interpolated quarterly values may then be re-aggregated into conventional year values ranging from the first to the fourth quarters. This is how interpolation, more precisely temporal dis-aggregation, may be used to calendarize financial year values.

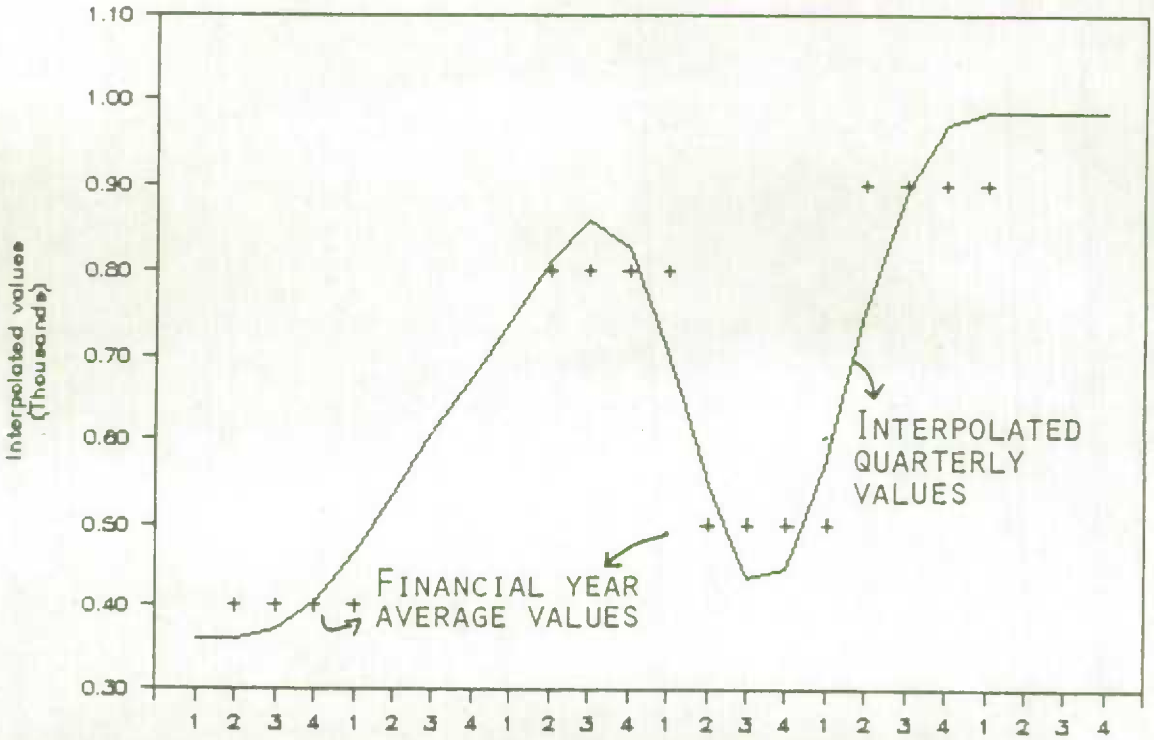


Figure 1.1: Non-seasonal interpolation by means of the modified Boot, Feibes and Lisman method

For regularly spaced and binding annual benchmarks (high values of g^{y}_{m}) and in the absence of sub-annual benchmarks (g^{z}_{k}=0), the problem specified by objective functions (1.1) or (1.1′) is that addressed by Boot, Feibes and Lisman (1967). Both objective functions are indeed generalizations of the method proposed by the authors. Incidently, additive benchmarking (Part 1 equation (2.1.1)) is equivalent to interpolate corrections between the annual and sub-annual discrepancies according to (1.1); and, to adding the corrections to the original sub-annual series x_t. This is one link between benchmarking and interpolation.

The interpolation of non-seasonal values can alternatively be carried out with the proportional variant of benchmarking, if the original series x_t is set equal to any non-zero constant. Indeed minimizing

$$f(z) = g^{x} \sum_{t=2}^{T} \{(z_t/x_t) - (z_{t-1}/x_{t-1})\}^2 + \sum_{m=1}^{M} g^{y}_{m} \{(\sum_{t=\tau_m}^{\rho_m} z_t) / y_m - 1\}^2 \tag{1.2}$$

$$\sum_{k=1}^{K} g^{z}_{k} \{(z_t / z^{d}_{k}) - 1\}^2 .$$

with x_t constant, yields identical results as minimizing (1.1). Similarly specifying second difference in (1.2) is equivalent to minimizing (1.1′). In other words, in the absence of sub-annual series, the additive variant may be seen as a particular case of the proprotional variant. Further developments in this section will consequently be based on the proportional variant.

Applied to interpolation, the proportional variant of benchmarking lends itself to a variety of applications: interpolation of sub-annual values with seasonal and possibly trading-day variations, interpolation of daily values from weekly data.

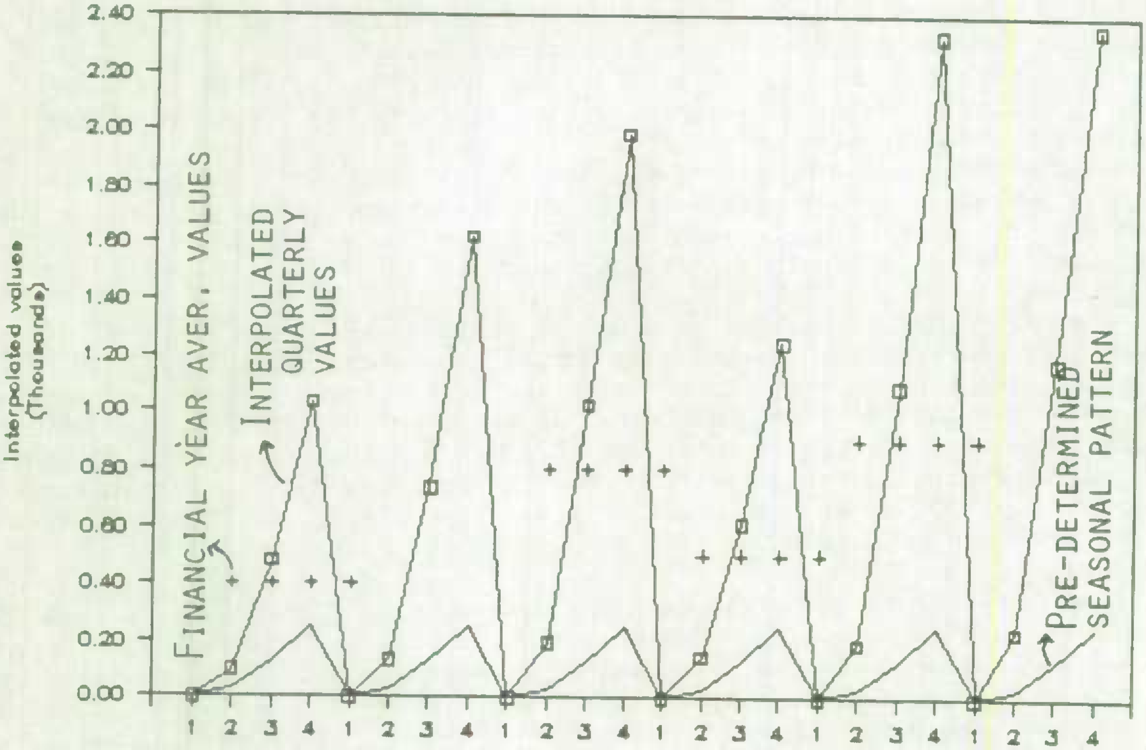


Figure 1.2: Seasonal interpolation from a seasonal pattern equal to 0.01%, 25%, 125% and 250%

1.2 Interpolating seasonal values

In order to interpolate seasonal values, the proportional variant of benchmarking is appropriate. Objective function (1.2) is then minimized, where x_t is a seasonal indicator s_t. The latter may be any seasonal sequence of numbers , e.g. for a quarterly series s_t = [10, 25, 20, 30; 10, 25, 20, 30; 10, 25,...]. In fact a formal seasonal pattern is intuitively easier to select. A value equal to 150% means the month or quarter considered is 50% higher than an average month; a value equal to 80%, 20 percent lower than average, and so forth. A value equal to 0.0001% specifies the absence of activity (in practice) for the month considered. (Zero values of x_t are not acceptable in the proportional model.) It is also possible to select a seasonal pattern which evolves from year to year. The annual average of the seasonal pattern does not have to be 100%, since the level is adjusted in the benchmarking process. This implies the seasonal pattern may be chosen to reflect the percentage of the yearly activity carried out in each month. This alternative way of building a seasonal pattern may be the easiest.

The resulting interpolation problem amounts to benchmarking a pre-selected seasonal pattern x_t to the annual (and possibly sub-annual) benchmarks y_m. The benchmarks should presumably be binding (high values of $g^{y}{}_{m}$). In cases where the benchmarks predictably behave monotonically (e.g. linearly or exponentially), second differences may be preferable in objective function (1.2); in doubt, first differences are preferable.

That seasonal interpolation process is illustrated in Figure 1.2. The quarterly series is interpolated with first differences from a constant seasonal pattern and from the same benchmarks as Figure 1.1. The constant seasonal pattern chosen, 0.01%, 25%, 100% and 250%, would be adequate for some agricultural series. It specifies no activity in the first quarter and most of the activity in the fourth and the third quarters. The annual benchmarks cover from the second to the first quarter of the following year. The quarterly values resulting from the interpolation may be used as such or re-aggregated in conventional year, i.e. "calendar", values (ranging from the first to the fourth quarters)

Similarly, monthly values obtained in that manner may be used as such or re-aggregated into quarterly or annual estimates (or both). This kind of re-aggregation is appropriate to correct the reference period of annual data, pertaining to financial years. Benchmarks covering from February to January of the following year for instance are dis-aggregated into monthly values. The monthly estimates are then re-aggregated into conventional quarters (i.e. January to March, April to June, etc.) and/or into conventional years (extending from January to December).

The same approach may be used for correcting the reference periods of financial quarter data. The presence of seasonality in such benchmark values would make the three first and last monthly estimates subject to heavy revisons. The situation may be improved by artificially extending the series by one financial quarter at each end of the series. The following trivial ARIMA model would be adequate for most situations (provided the reference periods of the benchmarks are regular):

$$y_{I+1} = y_I + y_{I-4} - y_{I-5}$$

$$y_0 = y_I + y_5 - y_6$$

For more details about this type of reference period correction, one may refer to Cholette, (1987a) and Cholette and Baldwin (1988a, 1988b, 1988c) .

1.3 Interpolating Values with Trading-Day Variations

Monthly flow series are likely to contain seasonal but also trading-day variations. The same approach to dis-aggregation may be followed. In order to interpolate monthly values with trading-day variations, objective function (1.2) is minimized with x_t representing a trading-day pattern TD_t:

$$x_t = TD_t = \sum_{k=1}^{7} n_{tk} D_k / 30.4375, \quad t=1,\ldots,T \tag{1.3}$$

where D_1, D_2, ..., D_7 are the daily weights constituting the weekly trading pattern. Number 30.4375 is the average number of day in a month (365.25/12); and $n_{t,k}$, the number of each day k in month t. The weekly pattern is chosen in much the same way as the seasonal pattern. A weight D_1=10% means that Monday is only 10% as important as an average day; a weight D_5=200%, that Friday is twice (100% more) as important as an average day. A daily weight equal to zero, e.g. D_7=0% specify no activity for that day.

The weekly pattern may be chosen to evolve through time. Objective function (1.2) is then minimized with x_t equal the following trading-day pattern:

$$(1.4) \quad x_t = TD_t = \sum_{k=N_t+1}^{N_t+n_{t.}} D_{kt} / 30.4375, \qquad N_t = \sum_{\theta=1}^{t-1} n_{\theta .}, \quad t=1,\dots,T$$

where $n_{t.}$ is the total number of days in month t and N_t is the cumulative number of days elapsed before month t. The number of input daily weights D_k is thus (at least) equal to the number of days covered by the T months, as opposed to seven in (1.3).

In general, a monthly series contains both seasonality and trading-day variations. The appropriate indicator x_t is then the product of TD_t (of (1.3) or (1.4)) and of a pre-selected seasonal pattern s_t. However the length of the months which is seasonal in nature would be taken into account twice, once by the chosen seasonal pattern and once (although implicitely) by the trading-day component defined in (1.3) or (1.4). In order to avoid that specification error, the demominator of (1.3) or (1.4) is replaced by the total number of days in the month $n_{t.}$. These equations become

$$(1.3') \qquad x_t = s_t * TD_t = s_t * \sum_{k=1}^{7} n_{tk} D_k / n_{t.}, \qquad t=1,\dots,T$$

for constant weekly pattern; and

$$(1.4') \quad x_t = s_t * TD_t = s_t * \sum_{k=N_t+1}^{N_t+n_{t.}} D_{kt} / n_{t.}, \qquad N_t = \sum_{\theta=1}^{t-1} n_{\theta .}, \quad t=1,\dots,T$$

for evolving weekly pattern.

The indicator series x_t may also contain a cyclical indicator c_t, in which case x_t is equal to the product of c_t, TD_t and s_t.

Figure 1.3 illustrates monthly trading-day interpolations obtained from a constant monthly trading-day pattern x_t (calculated by means of (1.3)) and from financial year data covering from February of one year to January of the following year. The interpolated values may be used as such or recombined into conventional years or conventional quarters.

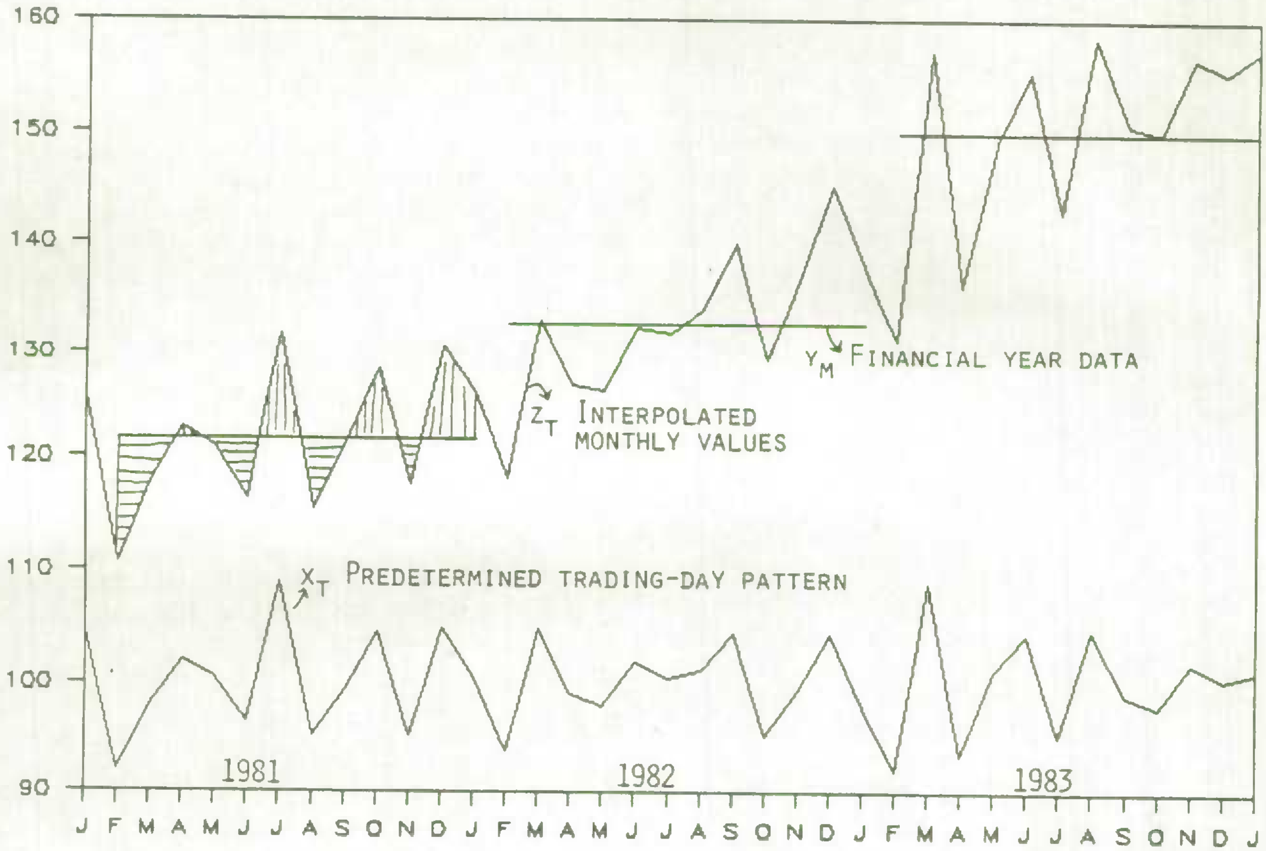


Figure 1.3: Trading-day interpolation from a constant trading-day pattern

1.4 Interpolating Daily Values from Weekly Data

As mentionned in Section 1.1 of Part 1, statistical agencies must sometimes convert bundles of weekly values into monthly values. This problem may also be casted in the framework of interpolation and benchmarking (Cholette and Baldwin, 1988c). The appropriate objective function to minimize is (1.2) with first differences. Time t now stands for days; and y_m, for the values of the weekly bundles. The appropriate indicator x_t is a repeated sequence of seven constant or evolving daily weights, for example 50% for Monday, 90%, 100%, 140%, 160%, 160% and 0.001% for Sunday. Their interpretation remains the same as in Section 1.3.

The trading-day interpolation process illustrated in Figure 1.4 (a), consists of benchmarking the weekly trading pattern selected (not displayed) to the available bundles values. A constant weekly pattern was chosen: D_1=120%, D_2=120%, D_3=140%,D_4=160%, D_5=160%, D_6=.01% and D_7=.01% (no activity for Saturday and Sunday). The resulting daily values z_t may be used as such or re-aggregated into monthly values.

Weekly bundles are likely to contain trend-cycle, seasonal and irregular fluctuations. (They cannot contain trading-day variations which cancels out on any one week.) The presence of seasonality - especially - makes it unlikely that the next weekly bundle will be in the neighbourhood of the last one. Unfortunately, this is more or less what is assumed in

Figure 1.4 (a). The fourth bundle is implicitly anticipated to be in the neighbourhood of A. Figure 1.4 (b) presents a scenario where the fourth bundle is lower than anticipated. The new interpolated values pertaining to the third bundle contradict those first obtained for that bundle. The revision is specially pronounced at the end and at the start of the third bundle. This illustrates the necessity of forecasting the next bundle in order to have good daily estimates for the current bundle. Another alternative is to ignore the estimates of the current bundle. (Accepted operational delays might allow that.) It may also be possible to ask the respondents to supply gross anticipations of their next bundle, like no change, 10 % increase (or decrease), 20 % increase, 30% increase, etc.

Another issue is the scale of the calculations required by the formal minimization of objective function (1.2). The four bundles of Figure 1.4 (b) cover 4, 4, 5 and 4 weeks respectively for a total of (at least) 119 desired daily values. Multiplications and inversions of matrices with dimensions 119 by 119 are therefore involved. The numerical approximation described in Section 4.2 of Part 1 is in order. A variant of the approximation could be considered. The weekly bundles could be first be disaggregated into weekly values, which are in fact illustrated in Figure 1.4 (a); and then those weekly values, dis-aggregated into daily values.

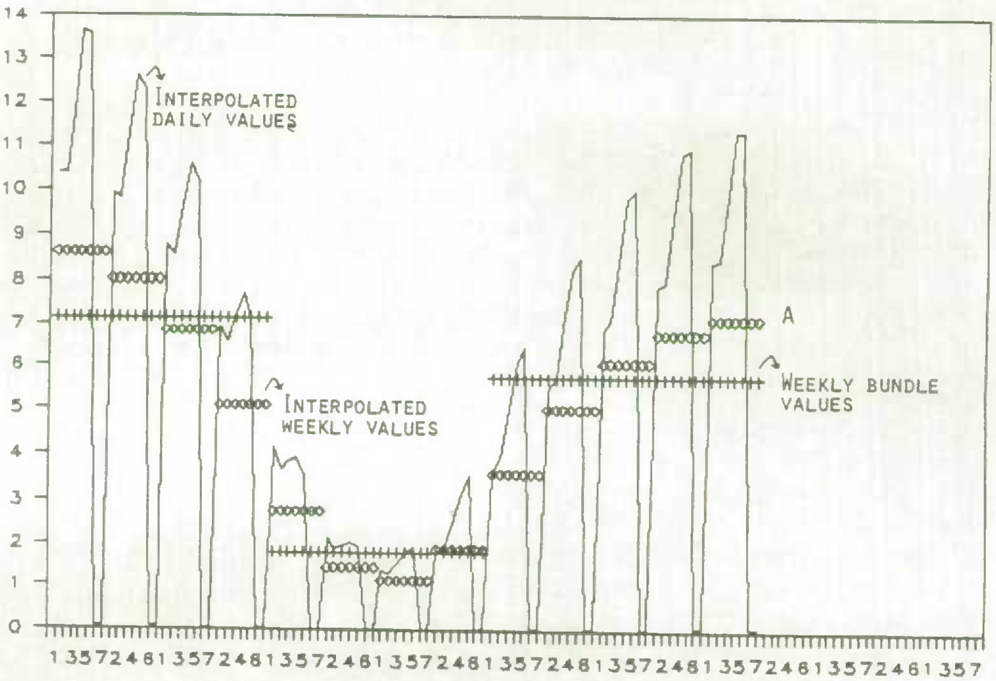


Figure 1.4 (a): Interpolated daily values from three bundles of weekly data, respectively covering 4, 4 and 5 weeks

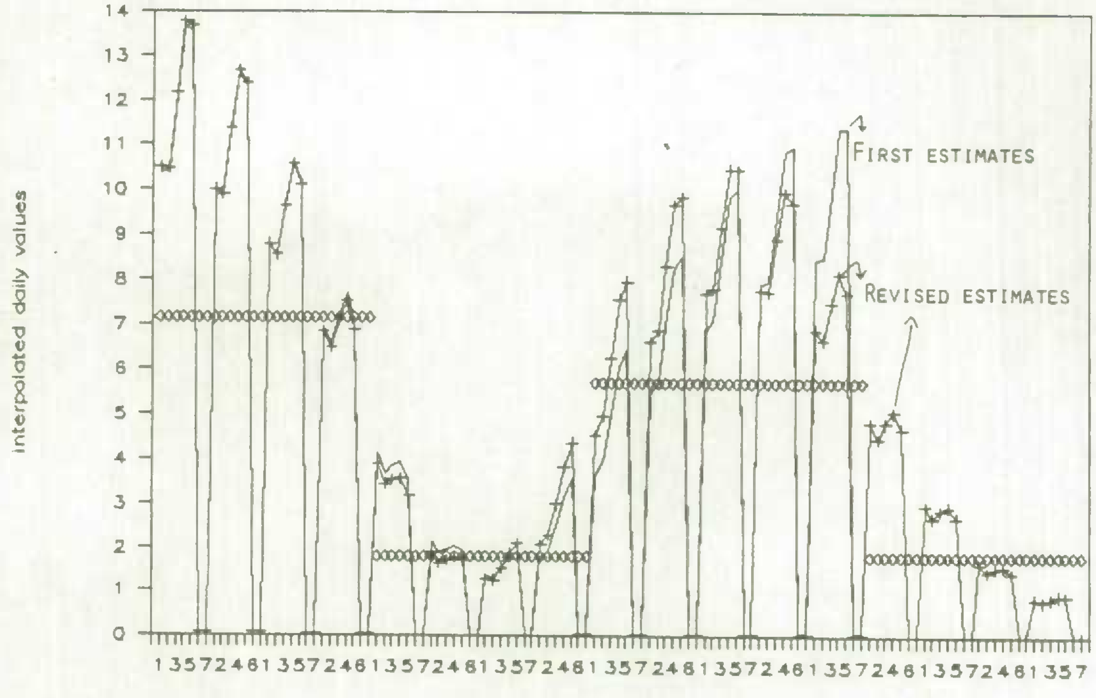


Figure 1.4 (b): Revision to the interpolated daily values when adding a fourth bundle of weekly data

2. INTERPOLATION BY MEANS OF GROWTH RATES

In some situations, the desired "sub-annual" series is derived from annual benchmarks, from a few sub-annual benchmarks and from sub-annual growth rates. In other words, the indicator of the desired sub-annual series z_t takes the form of growth rates r_t (in the neighbourhood of 1.0) to be followed by the desired series between periods t-h and t:

(2.1) $$z_t = r_t z_{t-h} + a_t, \quad t=h+1,..,T$$

or

(2.1') $$z_t - r_t z_{t-h} = a_t, \quad t=h+1,..,T$$

where a_t is a projection error. That error originates from the fact that the growth rates do not project the desired series exactly. In particular, the projections generated from (2.1) alone will not necessarily comply with available benchmarks. Parameter h is usually equal to 4 for quarterly series, i.e. the growth rates describe the movement from the same quarter of the previous year to the quarter considered. Note that the projections errors compound with time. For instance, a large error at time t is essentially repeated at times t+h, t+2h,... (ceteris paribus). The rest of this section will assume quarterly series with same-quarter growth rates (h=4).

The growth rates are determined by the series builder, from relevant information available sub-annually. An example will clarify the situation. The desired quarterly series z_t is the Total Expenditures by Hospitals. The annual values, i.e. the annual benchmarks y_m, are known, but not the quarterly values z_t. For each quarter, the same-quarter growth rates r_t are calculated from a projector series x_t, in the example the quarterly Wages and Salaries series paid by hospitals ($r_t = x_t/x_{t-4}$). It is indeed arguable that the growth in the quarterly expenditures is related to that in the wages: a large proportion of the expenditures consist of the wages and that proportion must be quite constant or change very gradually over time - at least for a given quarter. Subject matter experts are in position to verify that. In order to allow the projector series x_t and desired series z_t to have different seasonal patterns, annual same-quarter growth rates are generally used (instead of quarter-to-quarter). Indeed two series may have different seasonal patterns and identical annual growth rates.

In most situations, the growth rates are some weighted average of the growth measured in a few projector series and related information. The series builder may also change the weights of the average and the blend of related series through time. This possibility makes the growth rate approach very flexible and adaptable to the current operational circumnstances.

The presumed growth rate behaviour of the desired series formalized in equation (2.1) leads to the following objective function:

$$f(z) = g^x \sum_{t=h+1}^{T} (z_t - r_t\, z_{t-h})^2 + \sum_{m=1}^{M} g^y{}_m \{((\sum_{t=\tau_m}^{\rho_m} z_t)/y_m) - 1\}^2 \tag{2.2}$$

$$+ \sum_{k=1}^{K} g^z{}_k \{(z_t/z^d{}_k) - 1\}^2.$$

The first term of (2.2) is the **growth rate criterion**. That criterion minimizes the projection errors of (2.1'). Put differently, it specifies that the desired series should as much as possible behave according to the selected growth rates r_t. The second and third terms are the annual and sub-annual benchmarks satisfaction criteria. Like in Section 2 of Part 1, high values of $g^y{}_m$ and $g^z{}_k$ specify binding benchmarks.

For series derived from growth rates, the sub-annual benchmarks typically form an initial base-year at the start of the series. The specification of the projection problem would in fact be incomplete without sub-annual benchmarks. These determine the seasonal pattern of the desired series (and nothing else does). Indeed, one can show that the projections in (2.1) are implicitely based on 4 initial values $z^d{}_1$, $z^d{}_2$, $z^d{}_3$ and $z^d{}_4$:

$$z_t = (\prod_{k=1}^{i-1} r_{(k-1)4+j})\; z^d{}_j,\; j=1,\ldots,4;\; i=1,2,\ldots;\; t=(i-1)4+j, \tag{2.3}$$

where i stands for the year and j for the quarter considered. The initial base-year values are chosen by the subject matter expert. For instance the quarterly distribution of the projector series x_t is used to distribute the annual benchmarks of the first year. The distributed values may be improved on the basis of additional information.

Mathematically, the sub-annual benchmarks may be located anywhere in the series, namely at the end of the series. In that case, the projections are actually backcasts or retropolations. Such a situation is likely to happen in the development stage of series. Series are often built from the current year backwards.

Figure 2.1 illustrates a case where base-years are available both at the start and at the end of the series. As a result, the seasonal pattern of the interpolated series gradually moves from that given by (the sub-annual benchmarks in) the initial base year to that given by the terminal base-year. The figure also displays the series obtained in the absence of a terminal base-year. Although it also satisfies the annual benchmarks, its reliability deteriorates as the estimates depart from the initial base-year. Namely, its seasonal pattern basically repeats the initial seasonal pattern. One concludes that incorporating the terminal base-year retroactively improves the series, by retro-polating the terminal seasonal pattern.

Another approach to benchmarking series obtained from growth rates is to generate the projections through equation (2.1) and to benchmark them with the modified Denton method for instance. This produces a series consistent with the benchmarks. However, the reliability of the series is not retro-actively improved (ceteris paribus) on the availablity of the terminal base-year (or of any sub-annual benchmark). Furthermore, a level and a seasonal discontinuity is likely to arise between the new terminal base-year and the values preceeding that year.

Appendix G gives the linear algebra solution to the growth rate interpolation problem. More details about this technique can be found in Cholette (1985a)

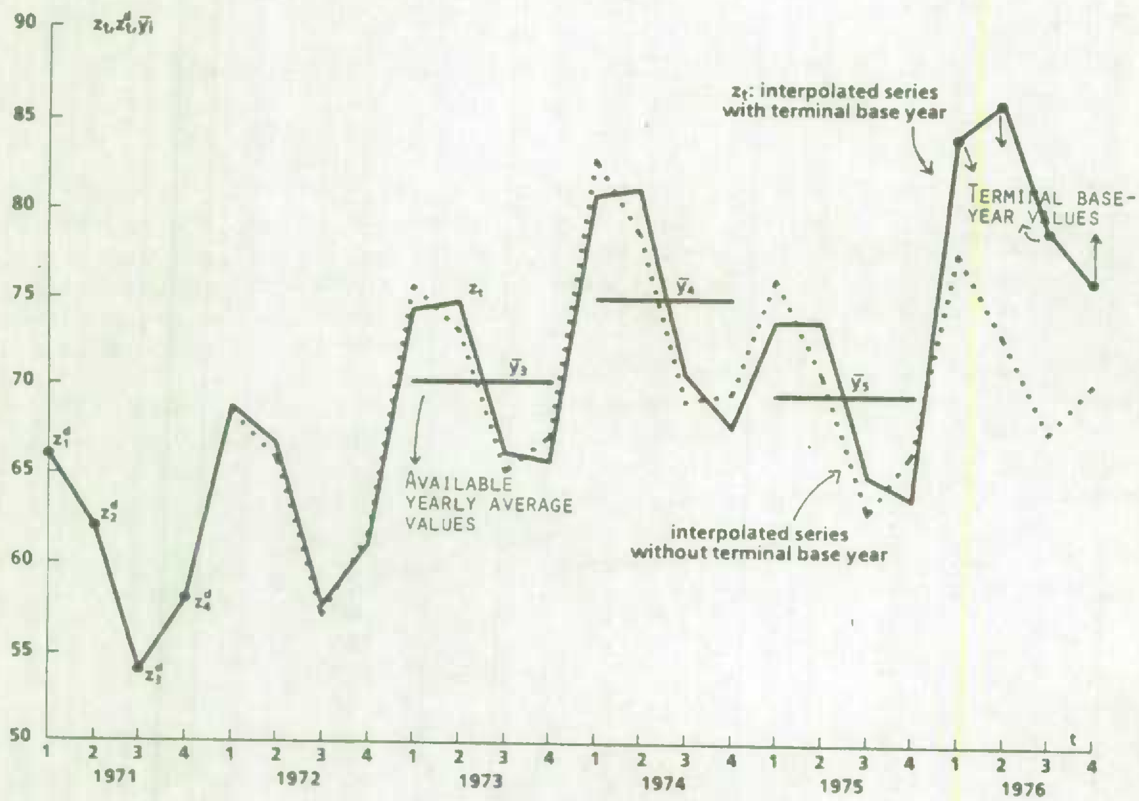


Figure 2.1: Series interpolated from growth rates in the presence 1) of an initial and a terminal base-year and 2) of an initial base year only

3. ARIMA INTERPOLATION OF SERIES

It is also possible to interpolate time series by means of ARIMA criteria. Auto-Regressive Integrated Moving Average models were popularized by Box and Jenkins (1970). These models are primarily used to forecast a time series from its past values. For most socio-economic time series, their short-run forecasting performance is astounding. To our knowledge they were never used for interpolation purposes. But in our opinion, they are worth entertaining, because they would be immensely less laborious than the growth rate approach.

Consider the following trivial quarterly seasonal ARIMA model:

$$z_t - z_{t-1} = z_{t-4} - z_{t-5} + a_t, \quad t=6,\ldots,T \tag{3.1}$$

Variable a_t is a "forecasting" error when considered a priori, i.e. from a previous period of time; and, an "innovation" when considered a posteriori, i.e. after the fact. That model states that the movement from one quarter t-1 to the next quarter t is basically equal to the corresponding movement of the preceeding year. In other words, the quarter-to-quarter movements tend to repeat from year to year. In the absence of any innovation (all $a_t=0$), that behaviour produces a linear trend-cycle with constant seasonal pattern. More precisely, the **eventual forecast function** consists of a linear trend-cycle with constant seasonality. The trivial ARIMA model (3.1), a "(0,1,0) (0,1,0)" in the jargon, describes more than 85% of the variations of most socio-economic time series ($R^2>0.85$).

A more general ARIMA model, the seasonal autoregressive (1,0,0) (1,0,0), is proposed however:

$$(z_t - p_1 z_{t-1}) = p_4(z_{t-4} - p_1 z_{t-5}) + a_t, \quad t=6,\ldots,T \tag{3.2}$$

or equivalently

$$(z_t - p_4 z_{t-4}) = p_1(z_{t-1} - p_4 z_{t-5}) + a_t, \quad t=6,\ldots,T \tag{3.2'}$$

$$z_t - p_1 z_{t-1} - p_4 z_{t-4} + p_1 p_4 z_{t-5} = a_t, \quad t=6,\ldots,T \tag{3.2"}$$

With both the **regular autoregressive parameter** p_1 and the **seasonal autoregressive parameter** p_4 equal to 1, model (3.2) is identically equivalent to the trivial model (3.1). With $p_1=1.0$ and $p_4=1.05$, model (3.2) specifies that the quarter-to-quarter movement tends to increase by 5% from year to year. The resulting behaviour, more precisely of the eventual forecasting function, displays a linear trend-cycle and an increasing seasonal amplitude. With $p_1=1.0$ and $p_4<1.0$, the trend-cycle is linear and the seasonal amplitude decreases. By reasoning in a similar manner with equation (3.2'), one reaches the following conclusions. With parameters $p_1>1.0$ and $p_4=1.0$, the model specifies exponential behaviour of the series with constant seasonal pattern; with $p_1>1.0$ and $p_4<1.0$, exponential behaviour with shrinking seasonal pattern; with $p_1>1.0$ and $p_4>1.0$, exponentioal behaviour with increasing seasonal pattern; etc.

The seasonal autoregressive model (3.2) can therefore describe a wide variety of series behaviours encountered in practice. ARIMA behaviour can be used as a substitue to growth rate behaviour, which leads to the following objective function:

$$f(z) = \sum_{t=6}^{T} (z_t - p_1 z_{t-1} - p_4 z_{t-4} + p_1 p_4 z_{t-5})^2$$

(3.3)

$$+ \sum_{m=1}^{M} g^y{}_m \{((\sum_{t=\tau_m}^{\rho_m} z_t)/y_m) - 1\}^2 + \sum_{k=1}^{K} g^z{}_k \{(z_t/z^d{}_k) - 1\}^2$$

where the regular and the regular autoregressive parameters p_1 and p_4 are known and pre-selected by the series builder.

The first term of the objective function, the ARIMA criterion, minimizes the ARIMA "forecasting" errors of (3.2"). In other words, it specifies that the desired should behave acccording to the seasonal autoregressive model selected. The second and third terms are the benchmark satisfaction criteria, high values of $g^y{}_m$ and $g^z{}_k$ specifying binding benchmarks.

Like the growth rate approach and for the same reasons, ARIMA interpolations requires at least one sub-annual benchmark for each quarter. Figure 3.1 illustrates an interpolation situation: Annual values start in 1978, but quarterly values x_t are available only in 1985. The job of the series builder is to develop that series backwards from 1985, that is to find plausible quarterly values for all the years. A first attempt is made by selecting a seasonal autoregressive criterion with increasing seasonal amplitude, e.g. with $p_1=1.0$ and $p_4=1.05$. Such a seasonal pattern is required to avoid negative interpolated values at the start of the series. The four original sub-annual values x_t of 1985 were specified as non-binding benchmarks; and the annual values y_m, as binding benchmarks. Figure 3.2 displays the resulting series. Figure 3.3 presents a second attempt, in which the expert imposes a different seasonal pattern at the start of the series, by means of binding sub-annual benchmarks.

As illustrated by both examples, ARIMA interpolation produces smooth series with no irregularity. The interpolated values may be used as such or as the basic structure of the series to be improved by the series builder. The philosophy behind such an approach is: first establish was is more "predictable" about the series; second, if needed, incorporate deviations from the basic structure using subject matter expertise. Selecting those two values is much less laborious than calculating growth rates for each period of time.

A by-product of ARIMA interpolation consists of forecasts. In Figures 3.1 to 3.3, the 1986 forcasts are done so that they lead to a pre-selected sub-annual benchmark value $z^d{}_{36}$. That capability lends itself to scenario analysis.

Note that ARIMA interpolation - as specified by objective fonction (2.3) at least - is inappropriate for monthly series with trading-day variations. Indeed classical ARIMA models cannot reliably model such fluctuations.

The method presented in this section is a particular case of that proposed by Guerrero (1987). In the latter, the ARIMA parameters and the weights attributed to the benchmarks are estimated instead of chosen by the series builder. Such estimation requires the availability of the desired for a sufficient length of time.

Objective function could incoroporate the movement preservation criterion specified only over selected segments of the series

The solution to ARIMA interpolation is developped in Appendix G.

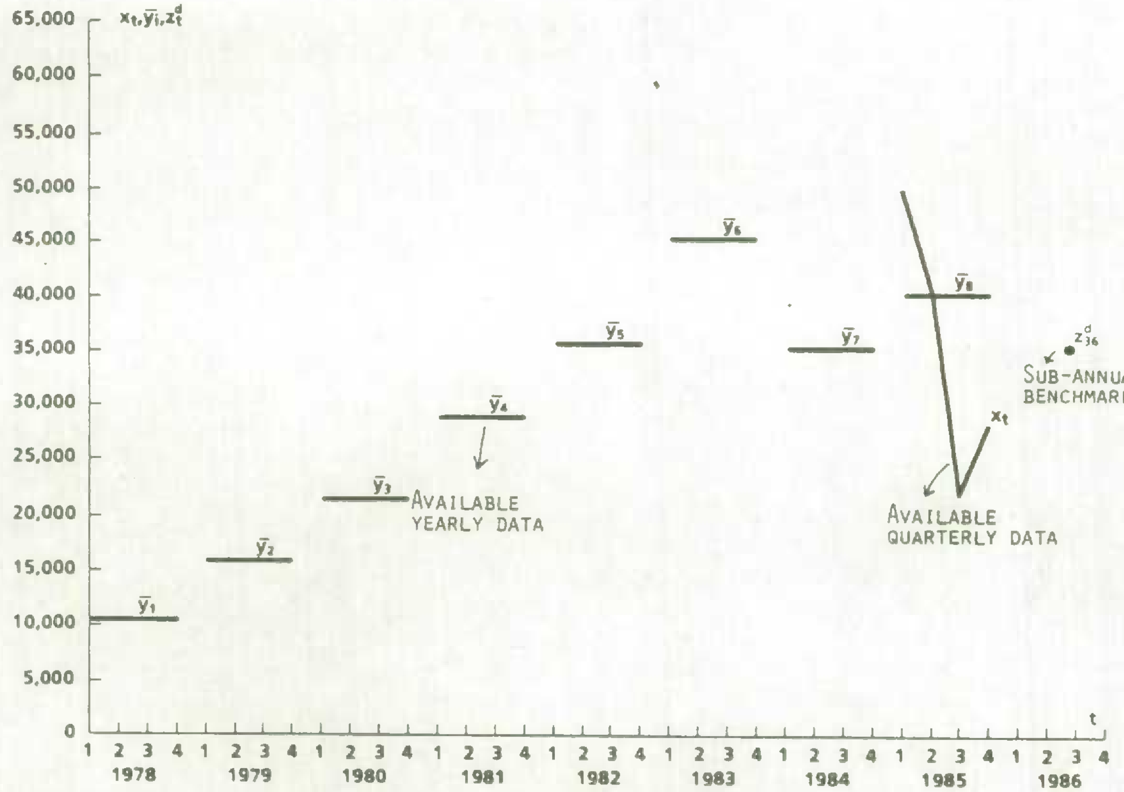


Figure 3.1: ARIMA Interpolation situation

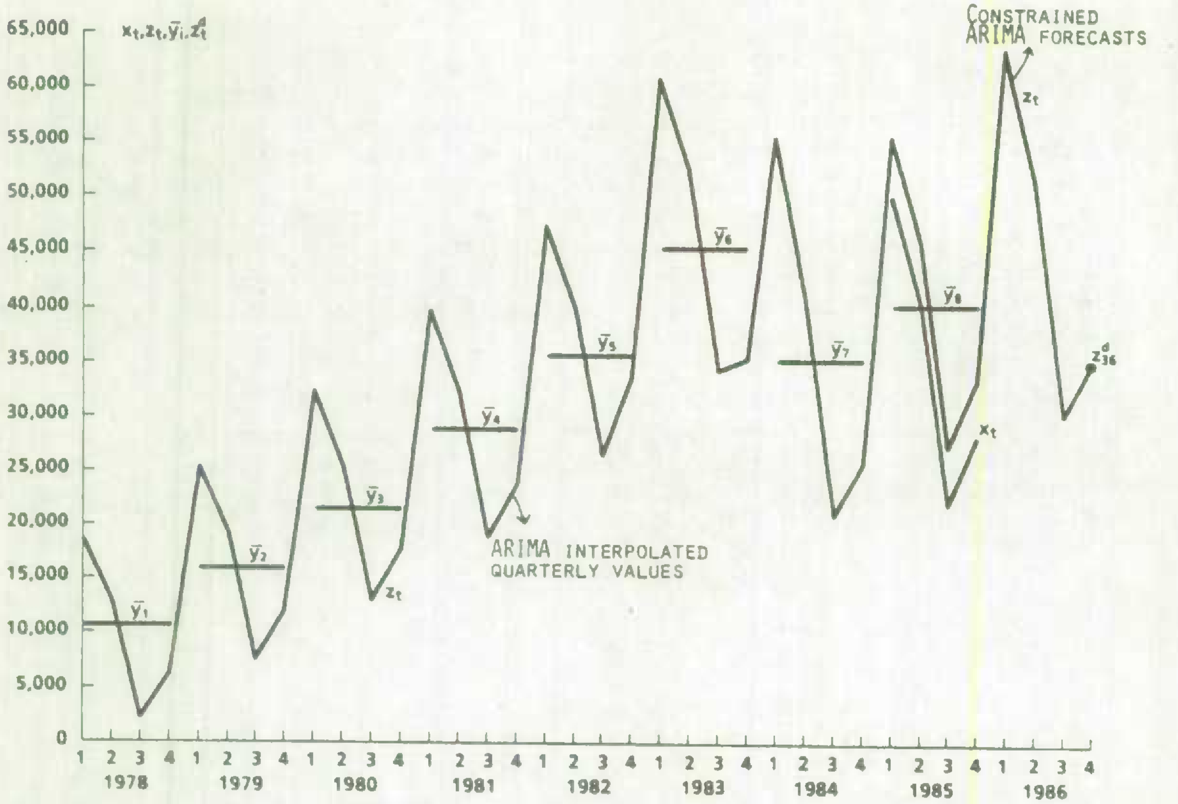


Figure 3.2: Retropolated seasonal pattern by means of an seasonal autoregressive model with p_1=1.0 and p_4=1.07

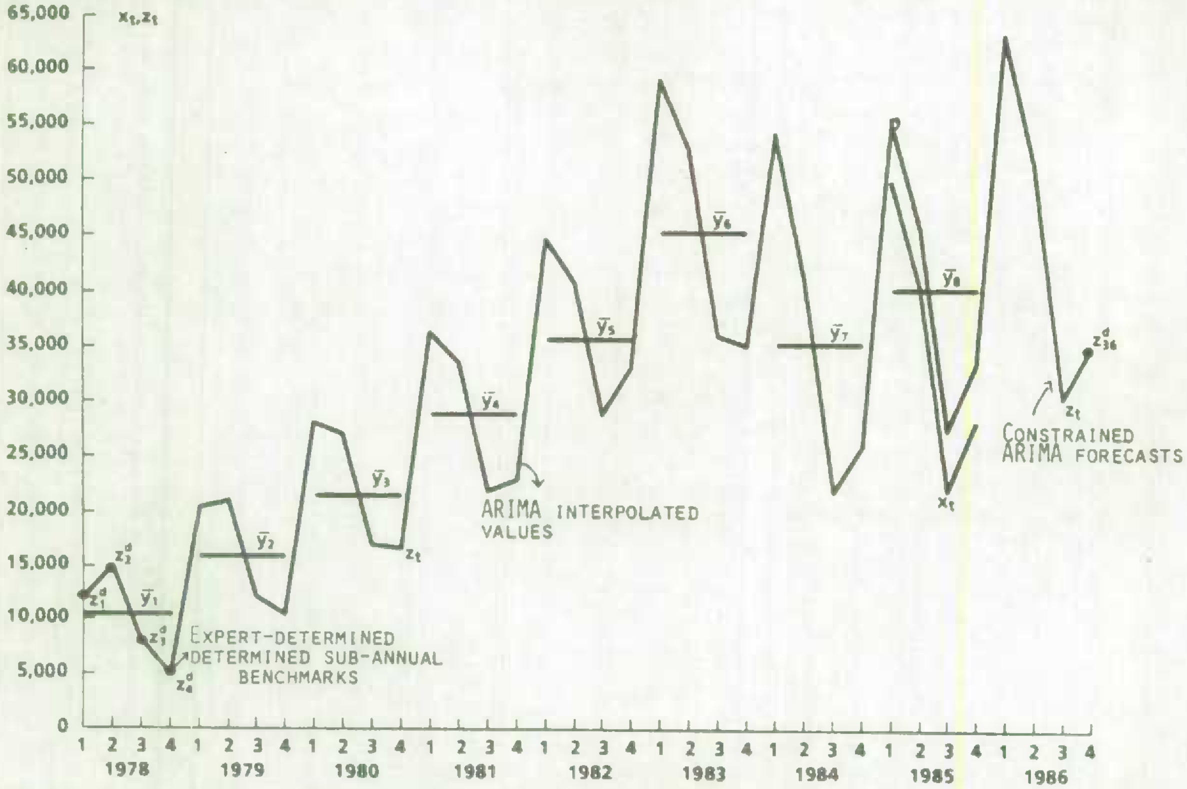


Figure 3.3: ARIMA interpolations between two seasonal patterns

4. INTERPOLATION FROM RELATED SERIES

The growth rate interpolation method, presented in Section 2 of Part 2, described how series may be interpolated from growth rates and from benchmarks. The growth rates had been derived by the series builder from related series. The rates were in turn used as an imput to the growth rate interpolation method, on the ground that the desired series should more or less display the same growth as the related series. A case can be made to directly incorporate the related series in the interpolation process. This section enquires into that possibility.

4.1 The Regression Model Approach

As specified by Chow and Lin (1971), Somermeyer et al. (1976), Fernandez (1981), Wilcox (1983) and Litterman (1983), the desired "sub-annual" series z_t is related to series x_{1t}, x_{2t}, ..., x_{Qt}, which are available sub-annually, by means of a regression model:

$$z_t = \sum_{q=1}^{Q} x_{qt} \gamma_q + u_t, \quad t=1,...,T \tag{4.1}$$

Since z_t is not available sub-annually, the parameters γ_q of the regression are estimated on the annual benchmarks y_m of z_t (as the regressand) and on the annual sums of x_{tq} (as the regressors), by means of Generalized Least Squares. The estimated coefficients are then applied on the sub-annual values x_{qt} to obtain the the desired sub-annual values z_t. The benchmarks y_m are automatically satisfied, because of a technical reason having to do with the variance-covariance matrix V used in the Generalized Least Squares estimation.

In order to avoid a singular variance-covariance matrix V however, all these authors effectively minimize the size of the first correction $(z_1 - (\Sigma\, \gamma_q\, x_{q1}))$to the regression fitted values $(\Sigma\, \gamma_q\, x_{q1}))$, as in the original Denton (1971) benchmarking method. This mis-specification is not considered important by the authors, because they are concerned with long series. Nowadays, statistical agencies are concerned with short series or at least short series interval. This mis-specification may seriously affect the parallelism between the interpolated and the regression fitted values.

As shown by Sanz (1981) and Fernandez (1981), the Chow and Lin and similar regression approaches may equivalently be specified by means of quadratic minimization, just as the methods presented in this document. It is then possible to avoid the aformentionned specification error by means of the following objective function:

$$f(z) = g^x \sum_{t=2}^{T} \{\Delta[z_t - (\sum_{q=1}^{Q} x_{qt}\, \gamma_q)]\}^2 + \sum_{m=1}^{M} g^{y}{}_m\, \{(\sum_{t=\tau_m}^{\rho_m} z_t) - y_m\}^2 \tag{4.2}$$

$$\sum_{k=1}^{K} g^{z}{}_k\, \{z_t - z^d{}_k\}^2.$$

The parameters γ_q are linearly estimated in the objective function along with z_t. The first term specifies that the corrections made to the regression fitted values $\Sigma \gamma_q x_{qt}$ (sum. over q) are as constant as possible. In other words, the interpolated values are to be as parallel as possible to the fitted values yielded by the regression. The second and third terms are the annual and sub-annual benchmark satisfaction criteria.

It is then possible to used the regression approach to interpolation without distorting the movement of the interpolated series at the start of the series. Other considerations are now examined.

For Q and γ_1 both equal to 1, that is when the related series x_{1t} is actually the original sub-annual series x_t to be benchmarked, objective function (4.2) is identically equivalent to the additive variant of benchmarked presented in the first Part of this document.

For Q equal to 1 and γ_1 not equal to 1 (e.g. 1.10), objective function (4.2) specifies a mixed proportional and additive benchmarking variant. Part of the adjustment consists of changing the level of the original series x_t by a factor γ_1 (e.g. 1.10). That proportionality factor is a constant parameter for the whole series interval considered. Experience shows that deterministic models, with time invariant parameters, are not appropriate for socio-economic time series. As pointed by Kendall (1973), time series do not behave according to a given algebraic relation over a sustained interval of time; and, adding new observations to a deterministic model is also likely to cause revisions to the parameters. This revision translates into revisions to all the interpolated values, including those in the distant past. This property of the regression approach is quite unacceptable to statistical agencies and to users of the data.

For Q greater than 1, the relation between the desired and the related series is specified to behave according to a deterministic econometric model with constant parameters. This deterministic character leads to the same criticism. Another aspect of the specification is more questionable however. Objective function (4.2) implicitly assumes that the (net) seasonal pattern of the related series is the same as that of the interpolated series. On both these accounts, the growth rate interpolation, presented above in Section 2, is preferable: The relation is not deterministic because growth rates vary from period to period; and (as explained) the seasonal pattern of the interpolated series may be different from that of the related series.

Because it is cast in the more familiar framework of regression analysis, the econometric specification of interpolation from related series may seem rational and appealing. For practical reasons, we advocate stochastic specifications, like the ARIMA and the growth rate approaches and - when applicable - the modified Denton approach. As explained in Section 4 (of Part 1), the latter method entail an automatic stabilization of the estimates after a few years. This section remains nevertheless usefull to put the latter approaches into perspective.

Before concluding this section, an extreme case of interpolation from related series is reported. That case illustrates the philosophical, more precisely the epistemological, implications of interpolation and to some extent of benchmarking.

4.2 Purely Artificial Data

Some time series published by some statistical agnecies are totally artificial or "model-based". Such series have no annual and no sub-annual benchmarks. They are only the result of an assumed econometric relation for the socio-economic variable considered. For instance, according to economic theory, the national production p_t of country at time t is a function of the number of employed workers e_t and of the quantity of capital (equipment) k_t with which the workers are producing:

$$p_t = p(e_t, k_t) \tag{4.3}$$

In this equation, production and employment are relatively easy to measure. The data for the stock of capital however are very hard to obtain, because they pertain to a variable (investment) cumulated over many decades - and even centuries - of economic activity. Since national accounts systems originated during World War II and thereafter, capital statistics are rather imprecise. Another reason is the difficulty to keep track of obsolete equipment being prematurely replaced by more advanced technology. It is then relatively tempting to fabricate capital stock series. This may be achieved in the following manner.

From equation (4.3), one can easily express capital in terms of the more easily measurable variables production and employment:

$$k_t = p^{-1}(e_t, p_t) \tag{4.4}$$

By selecting a production function (4.3) and its parameters, based on economic theory or on observation of a limited sector of the economy, the inverse function (4.4) is ready to use to create artificial capital data: Substituting the available employement and production on the righ-hand side of the equation yields the desired capital artificial estimates.

The anecdotal part is the following. A famous professor in a famous Canadian university asked his students to estimate various production functions from production, labour and capital data of countries of their choice. The goal of the exercise was to unravel the production theories which were better supported by reality. Much to her disbelief, one of the students found that the Cobb-Doublas production function totally "explained" production in one country. In other words, the estimated function (4.3) perfectly described the behaviour of production with no residual. The professor found those unprecedented results very interesting, until he realized that the data for capital in that country were a total fabrication. The intelligent student had inadvertently reconstructed the manner in which the capital data had been generated - namely by means of (4.4)!

This anecdote and similar occurences have profound epistemological consequences. On the one hand, students, academics and researchers use the data made available to them by statistical agencies in order to determine which socio-economic theories are supported by facts. On the other hand, the "data", supposed to reflect the facts, have actually been generated by statisticians by postulating some theories. Obviously, such data do not represent factual information, but the thinking and the opinions of statisticians about the facts. The data consequently confirm the theories akin to those used in calculating them; and invalidate other theories - in particular any new theory!

In some countries - especially the developpoing ones -, a large proportion of time series are the product of models assumed to represent the socio-economic system. However, to some extent the same situation prevails in developped countries: Central statisticians decide which theories and which "related" series will be used for interpolations. In international conferences, researchers present the result of years of investigation and triumphantly claim that the "data" significantly support such and such an hypothesis, such and such a theory. It is not uncommon then to see a witty statistician attending the conference reply - much to the disbelief of the flabbergastered academic -: "Of course, this is precisely the theory we used in generating the data"!

Statisticians using interpolation methods must be aware and make their users aware of this sort of philosophical considerations.

CONCLUSION

This document proposed a family of methods for benchmarking and interpolating time series. Part 1 generalized the modified Denton (1971) benchmarking approach in order to deal with various situations, namely: the reference periods of the benchmarks may vary from occasion to occasion; their pattern of availability may be irregular; they may be unreliable. The approach was also generalized to benchmark systems of series subject to aggregation (e.g. regional and industrial) constraints.

Part 2 presented methods for interpolating time series. It was found that calendarization (e.g. conversion of financial year data into conventional year values) may be successfully addressed by some variants of benchmarking. Methods to interpolate series by means of growth rates and of ARIMA models were also introduced.

All the methods presented in this document are based on minimizing some quadratic criteria under constraints. They are numerical as opposed to statistical. Statistical model-based methods, like those by Hillmer and Trabelsi (1987) and Guerrero (1987), incorporate the stochastic properties of the series in the estimation process and provide confidence intervals (i.e. reliability measures) for the estimates as a by product. The lack of such intervals - especially - is the weak point of numerical methods.

Numerical methods, on the other hand, require relatively less time series expertise and are massively applicable - even at low levels of aggregation. At those levels, the low signal to noise ratio and the ignorance of the stochastic properties often preclude the use of statistical methods. Nunerical methods are then the only way to incorporate subject matter expertise, that is the intimate knowledge by series builders of the socio-economic processes and variables involved. That expertise is sometimes the only information available apart from the benchmarks. The numerical methods proposed in this document are all designed to systematically incorporate subject matter knowledge. Depending on the method considered, that knowledge takes the form of sub-annual benchmarks, of growth rates, of seasonal, ARIMA and other patterns to be displayed by the series. Taking a specific form, the expertise thus remains documented and may be discussed, and the estimation process becomes replicable on a large scale.

Further developments to numerical benchmarking and interpolation methods are desirable: development of measures of reliability of the estimates, development of further measures of assesment of the benchmarking operation, development of approaches to approximate simultaneous benchmarking, development of criteria for selecting indicator series - may be along the lines of Friedman (1962).

ACKNOWLEDGEMENT

I wish to publically thank my colleague Normand Laniel for his thorough and dedicated examination and discussion of most of the material in this document.

APPENDIX A: Vectors and matrices used in the different variants of benchmarkding and interpolation

All the variants of benchmarking and interpolation discussed in this document are based on a common notation. Some variants involve only a subset of the matrices defined here.

Vector Y contains the M "annual" benchmarks available for the series considered. The number of annual benchmarks M may be lower than the total number of years I in the series or in the series interval considered. In other words, some years may have no benchmark.

$$Y' = [\ y_1\ y_2\ \dots\ y_M\ ] \tag{A.1}$$

Vector Z_d contains K "sub-annual" benchmarks. Some series may have no such benchmarks.

$$Z_d' = [z^d{}_1\ z^d{}_2\ \dots\ z^d{}_K\ ] \tag{A.2}$$

Vector Z contains the T desired "sub-annual" values.

$$Z' = [z_1\ z_2\ \dots\ z_T]$$

Vector X contains the T original unbenchmarked sub-annual values.

$$X' = [x_1\ x_2\ \dots\ x_T] \tag{A.3}$$

The following diagonal matrices have diagonal elements which are the inverse of the elements of vectors already introduced:

$$\underset{T \text{ by } T}{X^{-1}} = \begin{bmatrix} 1/x_1 & 0 & 0 \dots & 0 \\ 0 & 1/x_2 & 0 \dots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & . & 1/x_T \end{bmatrix} \tag{A.4}$$

$$\underset{K \text{ by } K}{Z_d{}^{-1}} = \begin{bmatrix} 1/z^d{}_1 & 0 & 0 \dots & 0 \\ 0 & 1/z^d{}_2 & 0 \dots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & . & 1/z^d{}_K \end{bmatrix} \tag{A.5}$$

$$\underset{M \text{ by } M}{Y^{-1}} = \begin{bmatrix} 1/y_1 & 0 & 0 \dots & 0 \\ 0 & 1/y_2 & 0 \dots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & . & 1/y_M \end{bmatrix} \tag{A.6}$$

Matrix D_1 is the matrix first difference operator; and D_2, the second difference operator:

$$\text{(A.7)} \qquad \underset{(T-1) \text{ by } T}{D_1} = \begin{bmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 & 0 & 0 \\ . & . & . & . & & . & . & . \\ . & . & . & . & & . & . & . \\ . & . & . & . & & . & . & . \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & -1 \end{bmatrix}$$

$$\text{(A.8)} \qquad \underset{(T-2) \text{ by } T}{D_2} = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ . & . & . & . & & . & . & . \\ . & . & . & . & & . & . & . \\ . & . & . & . & & . & . & . \\ 0 & 0 & 0 & 0 & \dots & 1 & -2 & 1 \end{bmatrix}$$

Matrix D_r is the growth rate deviation operator, which calculates $z_t - r_t z_{t-h}$ for $t=h+1,\dots,T$. For quarterly series ($J=4$) with annual growth rates ($h=J$), the matrix is as follows:

$$\text{(A.9)} \qquad \underset{(T-4) \text{ by } T}{D_r} = \begin{bmatrix} -r_5 & 0 & 0 & 0 & 1 & 0 \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & -r_6 & 0 & 0 & 0 & 1 \dots & 0 & 0 & 0 & 0 & 0 \\ . & . & . & . & . & \dots & . & . & . & . & . \\ . & . & . & . & . & \dots & . & . & . & . & . \\ . & . & . & . & . & \dots & . & . & . & . & . \\ 0 & 0 & 0 & 0 & . & \dots & -r_T & 0 & 0 & 0 & 1 \end{bmatrix}$$

Matrix D_A is the ARIMA deviation operator. It calculates deviations from the seasonal autoregressive model selected by the series builder. For a quarterly series ($J=4$), the matrix is as follows.

$$\text{(A.10)} \qquad \underset{(T-5) \text{ by } T}{D_A} = \begin{bmatrix} f_1f_4 & -f_4 & 0 & 0 & -f_1 & 1 & 0 \dots & 0 & 0 & 0 & 0 & 0 & . \\ 0 & f_1f_4 & -f_4 & 0 & 0 & -f_1 & 1 \dots & 0 & 0 & 0 & 0 & 0 & . \\ . & . & . & . & . & . & \dots & . & . & . & . & . & . \\ . & . & . & . & . & . & \dots & . & . & . & . & . & . \\ . & . & . & . & . & . & \dots & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & . & . & \dots & f_1f_4 & -f_4 & 0 & 0 & -f_1 & 1 \end{bmatrix}$$

For a non-seasonal series a regular autoregressive model of order 2 (say) may be specified as follows.

$$\text{(A.10')} \qquad \underset{(T-2) \text{ by } T}{D_A} = \begin{bmatrix} -f_2 & -f_1 & 1 & 0 & 0 \dots & 0 & 0 & 0 & 0 \\ 0 & -f_2 & -f_1 & 1 & 0 \dots & 0 & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & . & 0 & -f_2 & -f_1 & 1 \end{bmatrix}$$

Matrix B_y is the "annual" sum operator. It calculates the sums of z_t over periods τ_1 to ρ_1, τ_2 to ρ_2, etc.

$$\text{(A.11)}\qquad \underset{M\ \text{by}\ T}{B_y} = \begin{array}{c}
\begin{array}{ccccccccccccccccc} \text{columns:} & & & & \tau_1 & & & \rho_1 & & & \tau_2 & & & \rho_2 & & & \end{array}\\
\begin{bmatrix} 0&0&\dots&0&1&1&\dots&1&0&\dots&0&0&\dots&0&0&0&\dots\\ 0&0&\dots&0&0&0&\dots&0&0&\dots&1&1&\dots&1&0&0&\dots\\ \vdots&\vdots& &\vdots&\vdots&\vdots& &\vdots&\vdots&\vdots&\vdots&\vdots& &\vdots&\vdots&\vdots& \\ & & & & & & & & & & & & & & & & \end{bmatrix}
\end{array}$$

For instance, in the case of a 4-year quarterly series, starting in a first quarter and with fiscal year annual benchmarks covering from the second quarter to the first quarter of the following year, the matrix is

$$\text{(A.11')}\qquad \underset{3\ \text{by}\ 16}{B_y} = \begin{bmatrix} 0&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&1&1&1&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&1&1&1&1&0&0&0 \end{bmatrix}$$

Matrix B_z could be termed a selection operator. Applied to Z, it selects the the t_1th, the t_2th,..., and the t_Kth value. Note that it is a particular case of B_y with $\rho_k = \tau_k$.

$$\text{(A.12)}\qquad \underset{K\ \text{by}\ T}{B_z} = \begin{array}{c}
\begin{array}{ccccccccccccccc} \text{columns:} & & & & t_1 & & & & t_2 & & & & t_3 & & \end{array}\\
\begin{bmatrix} 0&0&\dots&0&1&0&\dots&0&0&0&\dots&0&0&0&\dots\\ 0&0&\dots&0&0&0&\dots&0&1&0&\dots&0&0&0&\dots\\ 0&0&\dots&0&0&0&\dots&0&1&0&\dots&0&1&0&\dots\\ \vdots&\vdots& &\vdots&\vdots&\vdots& &\vdots&\vdots&\vdots& &\vdots&\vdots&\vdots& \end{bmatrix}
\end{array}$$

The following matrices are annual sum and selection operators used in an alternative specification of proportional benchmarking:

$$\text{(A.13)}\qquad \underset{M\ \text{by}\ T}{B^*_y} = \begin{array}{c}
\begin{array}{ccccccccccccc} \text{columns:} & & & \tau_1 & & \rho_1 & & & \tau_2 & & \rho_2 & & \end{array}\\
\begin{bmatrix} 0&\dots&0&x_{\tau 1}&\dots&x_{\rho 1}&0&\dots&0&\dots&0&0&\dots\\ 0&\dots&0&0&\dots&0&0&\dots&x_{\tau 2}&\dots&x_{\rho 2}&0&\dots\\ \vdots& & &\vdots& &\vdots&\vdots& &\vdots& &\vdots&\vdots& \\ & & & & & & & & & & & & \end{bmatrix}
\end{array}$$

$$\text{(A.14)}\qquad \underset{K\ \text{by}\ T}{B^*_z} = \begin{array}{c}
\begin{array}{cccccccccccccc} \text{columns:} & & & t_1 & & & & t_2 & & & & t_3 & & \end{array}\\
\begin{bmatrix} 0&\dots&0&x_{t1}&0&\dots&0&0&0&\dots&0&0&0&\dots\\ 0&\dots&0&0&0&\dots&0&x_{t2}&0&\dots&0&0&0&\dots\\ 0&\dots&0&0&0&\dots&0&0&0&\dots&0&x_{t3}&0&\dots\\ \vdots& & &\vdots&\vdots& &\vdots&\vdots&\vdots& &\vdots&\vdots&\vdots& \end{bmatrix}
\end{array}$$

where subscript τm, ρm and tk are to be read as τ_m, ρ_m and t_k.

APPENDIX B: Formal solution of the additive variant of the proposed benchmarking method

In matrix algebra, the objective function (2.1) (of Part 1) for additive benchmarking writes:

$$F(Z) = g_x\,(Z - X)'D'D\,(X - Z) + g_y\,(B_y Z - Y)'(B_y Z - Y) + g_z\,(B_z Z - Z_d)'(B_z Z - Z_d) \tag{B.1}$$

where the vectors and matrices are defined in Appendix A. Matrix D is usually the first difference operator of equation (A.7). In some applications, it may be appropriate to set D equal to the second difference operator (A8). Symbols g_x, g_y and g_z are scalars. They could also be diagonal matrices of dimension T by T, M by M and K by K respectively. This would allow the specification of some annual benchmarks (for instance) as binding and some as non-binding. Performing the matrix multiplications in (B.1) yields:

$$\begin{aligned} F(Z) = {} & g_x\,(Z'D'D\,Z - 2\,Z'D'D\,X + X'D'D\,X) \\ & + g_y\,(Z'B_y{}'B_y\,Z - 2\,Z'B_y{}'Y + Y'Y) \\ & + g_z\,(Z'B_z{}'B_z\,Z - 2\,Z'B_z{}'Z_d + Z_d{}'Z_d) \end{aligned} \tag{B.2}$$

The values of z_t which minimize this hyper-parabola are required. At the minimum, the derivative with respect to the unknowns z_t are equal to zero. This leads to the normal equations:

$$\begin{aligned} dF/dz = {} & 2\,g_x\,D'D\,Z - 2\,g_x\,D'D\,X \\ & + 2\,g_y\,B_y{}'B_y\,Z - 2\,g_y\,B_y{}'Y \\ & + 2\,g_z\,B_z{}'B_z\,Z - 2\,g_z\,B_z{}'Z_d = 0 \end{aligned} \tag{B.3}$$

The solution to this linear system of equations is:

$$Z = (g_x\,D'D + g_y\,B_y{}'B_y + g_z\,B_z{}'B_z)^{-1}\,(g_x\,D'D \mid g_y\,B_y{}' \mid g_z\,B_z{}') \begin{bmatrix} X \\ Y \\ Z_d \end{bmatrix} \tag{B.4}$$

The second order conditions for a minimum are satisfied, because the matrices D'D and B'B in (B.2) are positive semidefinite.

APPENDIX C: Formal solution of the proportional variant of the proposed benchmarking method

In matrix algebra, objective function (2.3) (of Part 1) for proportional benchmarking writes:

$$
\begin{aligned}
F(Z) = \; & g_x \, Z' X^{-1\prime} D' D \, X^{-1} Z \\
(C.1) \qquad & g_y \, (Y^{-1} B_y \, Z - \iota_y)' (Y^{-1} B_y \, Z - \iota_y) \\
& + g_z \, (Z_d^{-1} B_z \, Z - \iota_z)' (Z_d^{-1} B_z \, Z - \iota_z)
\end{aligned}
$$

where the vectors and matrices are defined in Appendix A. Vectors ι_z and ι_z of dimensions M by 1 and K by 1 contain values equal to 1. Matrix D is usually the first difference operator of equation (A.7). In some applications, it may be appropriate to set D equal to the second difference operator (A.8). Symbols g_x, g_y and g_z are scalars. They could also be diagonal matrices of dimension T by T, M by M and K by K respectively. This would allow the specification of some annual benchmarks (for instance) as binding and some as non-binding. Performing the matrix multiplications in (C.1) yields:

$$
\begin{aligned}
F(Z) = \; & g_x \, Z' X^{-1\prime} D' D \, X^{-1} Z \\
(C.2) \qquad & + g_y \, (Z' B_y' Y^{-1\prime} Y^{-1} B_y \, Z - 2 \, Z' B_y' Y^{-1\prime} \iota_y + \iota_y' \iota_y) \\
& + g_z \, (Z' B_z' Z_d^{-1\prime} Z_d^{-1} B_z \, Z - 2 \, Z' B_z' Z_d^{-1\prime} \iota_z + \iota_z' \iota_z)
\end{aligned}
$$

The values of z_t which minimize this hyper-parabola are required. At the minimum, the derivative with respect to the unknown z_t are equal to zero. This leads to the normal equations:

$$
\begin{aligned}
dF/dZ = \; & 2 \, g_x \, X^{-1\prime} D' D \, X^{-1} Z \\
(C.3) \qquad & + 2 \, g_y \, B_y' Y^{-1\prime} Y^{-1} B_y \, Z - 2 \, g_y \, B_y' Y^{-1\prime} \iota_y \\
& + 2 \, g_z \, B_z' Z_d^{-1\prime} Z_d^{-1} B_z \, Z - 2 \, g_z \, B_z' Z_d^{-1\prime} \iota_z = 0
\end{aligned}
$$

The solution to this linear system of equations is unique:

$$
\begin{aligned}
(C.4) \qquad Z = \; & (g_x \, X^{-1\prime} D' D \, X^{-1} + g_y \, B_y' Y^{-1\prime} Y^{-1} B_y + g_z \, B_z' Z_d^{-1\prime} Z_d^{-1} B_z)^{-1} \\
& (g_y \, B_y' Y^{-1\prime} \iota_y + g_z \, B_z' Z_d^{-1\prime} \iota_z)
\end{aligned}
$$

The second order conditions for a minimum are satisfied because all the squared matrices in (C.2) are positive semidefinite.

As pointed out by Laniel (1986), the problem of proportional benchmarking may be specified in terms of the corrections $c_t = z_t / x_t$

$$
\begin{aligned}
F(C) = \; & g_x \, C' D' D \, C \\
(C.5) \qquad & + \; g_y \, (Y^{-1} B^*_y \, C - \iota_y)' (Y^{-1} B^*_y \, C - \iota_y)
\end{aligned}
$$

$$+ \; g_z \, (Z_d^{-1} B^*_z \, C - \iota_z)' (Z_d^{-1} B^*_z \, C - \iota_z)$$

where B^*_y and B^*_z are defined in Appendix A. Following the same steps as in (C.1) to (C.4), the solution for the corrections is derived:

$$C = (g_x \, D'D + g_y \, B^*_y{}' Y^{-1}{}' Y^{-1} B^*_y + g_z \, B^*_z{}' Z_d^{-1}{}' Z_d^{-1} B^*_z)^{-1} \quad \text{(C.6)}$$
$$(g_y \, B^*_y{}' Y^{-1}{}' \iota_y + g_z \, B^*_z{}' Z_d^{-1}{}' \iota_z)$$

The benchmarked values are then simply $z_t = x_t * c_t$. This formulation gives exactly the same results as the one presented above and does away with matrix X^{-1}. The matrix to be inverted is also numerically better conditioned for inversion, because its elements are in the neighbourhood of 1.

APPENDIX D: Approximation for Benchmarking Individual Flow Series.

Let b_k, $k=1,...,K$, denote all benchmarks, whether annual or sub-annual; $\tau_k,...,\rho_k$, their reference periods; and d_k, the corresponding discrepancies, whether annual or sub-annual, additive or proportional. (K is now the total number of benchmarks.) The approximation consists of the following steps:

1) First the middle points t^*_k of the reference periods of the benchmarks b_k are found:

$$(1) \qquad t^*_k = (\tau_k + \rho_k)/2, \; k=1,...,K$$

For instance a benchmark covering periods 3 to 6 has middle point equal to 4.5 ((3+6)/2); a benchmark covering periods 11 to 15, 13; a (sub-annual) benchmark covering periods 1 to 1, 1.

2) Second initial middle values d^*_k are assigned to the middle points. In the proportional model, the values assigned are simply those of the corresponding discrepancies.

$$(2) \qquad d^*_k = d_k, \; k=1,...,K$$

In the additive model, the middle value is divided by the number of periods covered by the discrepancy.

$$(2') \qquad d^*_k = d_k \,/\, (\rho_k - \tau_k + 1), \; k=1,...,K$$

For instance, for a discrepancy equal to 80 covering periods 3 to 6, the middle value is 20, i.e. 80/(6-3+1).

3) Initial gross corrections c^g_t are found by linearly interpolating between the middle values by means of equations (4.1), where d_k is replaced by the middle values:

$$(3) \qquad c^g_t = d^*_k + (t-t_k) * [(d^*_{k+1} - d^*_k)/(t^*_{k+1} - t^*_k)], \; t^*_k \le t \le t^*_{k+1}, \quad k=1,..,K-1$$

For the periods preceeding and following the first and last middle values, the gross corrections simply repeat the first and last middle values:

$$(3'') \qquad \begin{aligned} c^g_t &= d^*_1, \quad t \le t^*_1 \\ c^g_t &= d^*_K, \quad t \ge t^*_K \end{aligned}$$

Figure D.1 (a) displays the gross corrections, with the discrepancies averaged over their reference periods (i.e. the discrepancies divided by the number of periods they cover). The middle values are located at the apexes of the broken curve.

4) The fourth step consists of adjusting the gross corrections so that they spread the observed discrepancies (i.e. so that the binding benchmarks are satisfied). This step yields initial adjusted gross corrections c^a_t.

In the multiplicative case, the gross corrections are multiplied by the original sub-annual series to obtain a gross benchmarked series. Residual proportional discrepancies between the benchmarks and the gross benchmarked series are calculated. The corrections are then multiplied by the residual proportional discrepancies.

$$(4) \qquad c^a_t = c^g_t * [b_k / (\sum_{t=\tau_k}^{\rho_k} c^g_t\, x_t)], \quad \tau_k \le t \le \rho_k,\ k=1,\ldots,K$$

For instance a benchmark covering periods 3 to 6 has value 1050, the sum of the gross benchmarked series over the same interval is 1000. The residual proportional discrepancy is then 1.05. The gross correction curve is then multiplied by 1.05 for periods 3 to 6. The same is done for each discrepancy.

In the additive case, the adjustment is achieved by adding to the gross corrections the average difference between the discrepancy and the sum of the gross corrections over the reference periods covered.

$$(4') \qquad c^a_t = c^g_t + [d_k - (\sum_{t=\tau_k}^{\rho_k} c^g_t)]/(\rho_k - \tau_k + 1)], \quad \tau_k \le t \le \rho_k,\ k=1,\ldots,K$$

For example a discrepancy covering periods 3 to 6 (say) is 80, and the sum of the gross corrections for the same periods is 60. The difference is then 20 and the average difference is 5 (20/(6-3+1)). The quantity 5 is then added to the gross corrections for the periods calculated in (4').

This did not adjust the periods not covered by any discrepancies. For the periods preceeding the first and following the last discrepancies, the adjusted corrections merely repeat the first and the last adjusted gross correction calculated in (4) or (4').

$$(4") \qquad \begin{aligned} c^a_t &= c^a_{\tau 1}, \quad t \le \tau_1 \quad (\text{read } \tau_1) \\ c^g_t &= c^a_{\tau K}, \quad t \ge \rho_K \quad (\text{read } \tau_K) \end{aligned}$$

For the periods between two discrepancies (but not covered), the adjusted corrections are linear interpolation between the closest adjusted corrections on each side.

$$(4"') \qquad c^a_t = c^a_{\rho k} + (t-\rho_k) * [(c^a_{\tau k+1} - c^a_{\rho k})/(\tau_{k+1} - \rho_k)], \quad \rho_k < t < \tau_{k+1},\ k=1,\ldots,K-1$$

The adjusted gross corrections resulting for this step are displayed in Figure D.1 (b), along with the gross corrections of step 3. Over each interval covered by a discrepancy, the two curves are parallel. The discontinuities between contiguous intervals will be eliminated in step 6.

Steps 1 to 4 provide initial values to the iteration process to begin. The fifth step is a sequence of three sub-steps 5a, 5b and 5c, repeated untill the benchmarking constraints are satisfied. In practice, no more than five iterations are necessary.

5a) The middle points are assigned new middle values. These are equal to the values of the adjusted gross correction at the middle points. When the middle point lies between two periods (e.g. t^*_k=6.5), the middle value is the average of the two linear extrapolations of the corrections on each side.

$$d^*_k = c^a_{t^*k} \quad \text{for } t^*_k \text{ integer}$$

(5a)

$$d^*_k = [c^a_{t'k} + c^a_{t'k+1} + (c^a_{t'k} - c^a_{t'k-1})/2 + (-c^a_{t'k+1} + c^a_{t'k+2})/2] / 2 \quad \text{for } t^*_k \text{ not integer}$$

(read subscript $t'k\pm x$ as $t'_k\pm x$)

where t'_k is the integer part of t^*_k.

5b) Improved gross corrections c^g_t are calculated by linearly interpolating between the new middle values (exactly as in step 3). The first improved gross correction are displayed in Figure D.1 (c), along with the initial adjusted gross corrections of step 4.

5c) The improved gross corrections are in turn adjusted to allocate the discrepancies. This is achieved exactly as described in step 4.

Sub-steps 5a to 5c are repeated five times (say). The resulting final adjusted corrections of step 5c are displayed in Figure D.1 (d) (along with the final corrections of step 7). The curve satisfies the benchmarking constraints. However it does not quite satisfy the parallelism criterion - especially at the apexes where it changes direction drastically. In order for the benchmarked series to be parallel to the original, the corrections should as smooth and constant as possible. The next step corrects the situation.

6) The sixth step consists of smoothing the final adjusted corrections with a moving average. For monthly series benchmarked to annual benchmarks, a simple 7-term moving average is used; for quarterly series, a weighted 3-term moving average with weights 1/4, 2/4 and 1/4; and for daily series benchmarked to weekly benchmarks, a simple 3-term moving average. (More precisely the length of the average depends on the number of time periods referred to by the benchmark. The lenght is equal to the integer part of half the number of time periods plus 1. However if that length is even, it is increased by 1. That length varies, if the benchmarks refer to different number of time periods. It varies linearly avoiding even numbers.)

7) The final corrections are obtained by adjusting the smoothed corrections of step 6 to comply with the benchmarking constraints, exactly as in step 4 except c^s_t is used instead of c^g_t. The final approximate corrections obtained are displayed in Figure D.1 (d) with the final

adjusted corrections of step 5c. The desired benchmarked series is equal to the sum or to the product of the final corrections and of the original sub-annual series.

In order to convincingly illustrate the proposed approximation technique, an unfavourable and unlikely benchmarking situation was chosen. The discrepancies of Figure D.1 behave erratically, their reference periods are irregular: they respectively cover 1, 3, 4 and 5 periods and none covers periods 2 and 3. One concludes that the approximation allows for both annual and sub-annual benchmarks and for very general benchmarking situations. However, the procedure - as designed at least - will not work when the reference periods of some benchmarks overlap, in particular when the reference period of a sub-annual benchmark is embedded in that of an annual benchmark. Such special situations can be handled by the formal minimization.

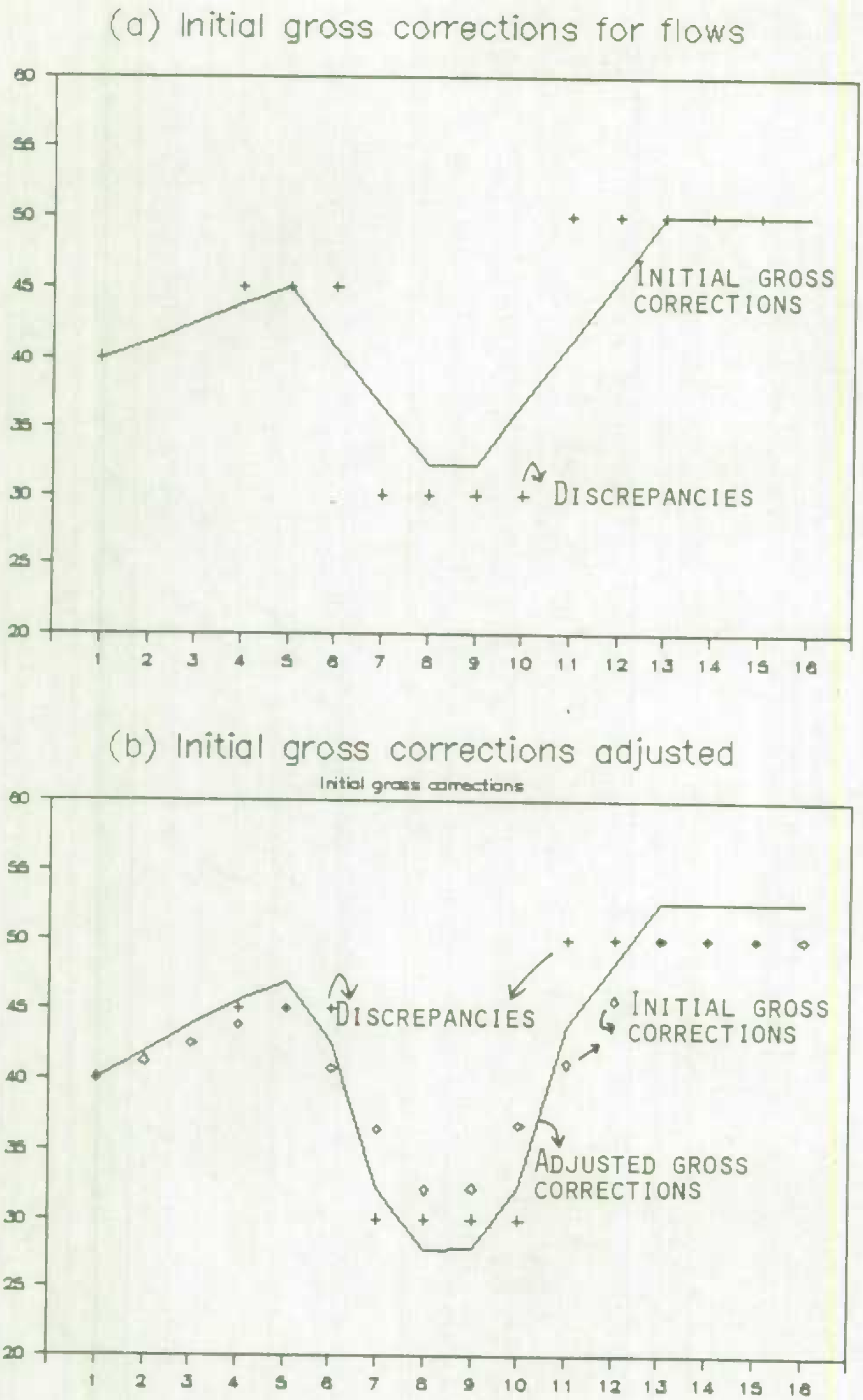
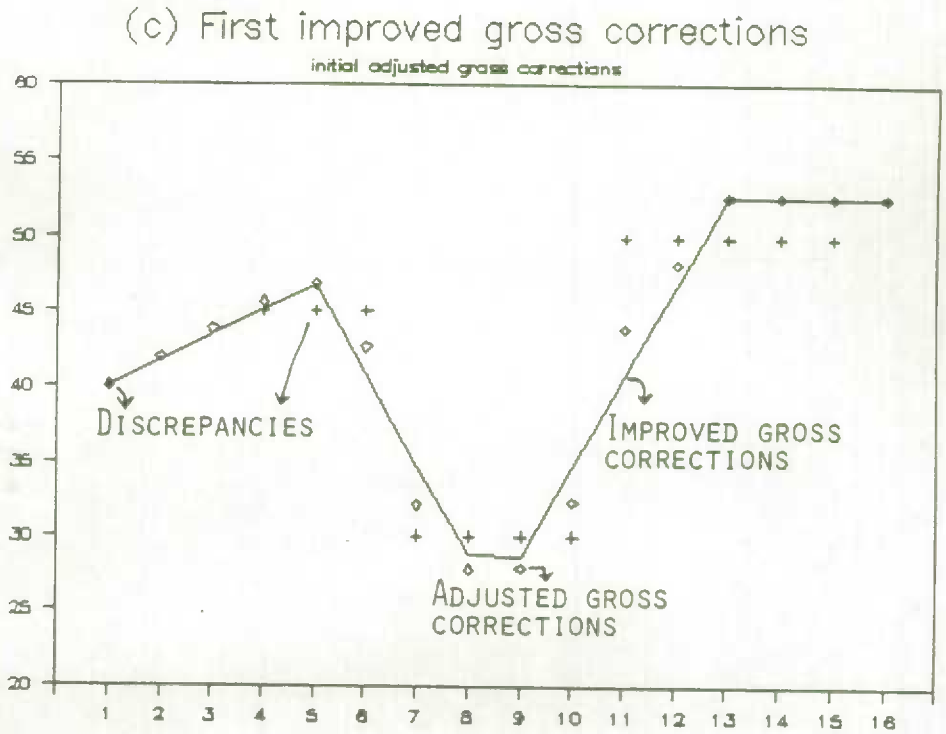


Figure D.1: Approximated additive corrections for a flow series with binding annual and sub-annual benchmarks



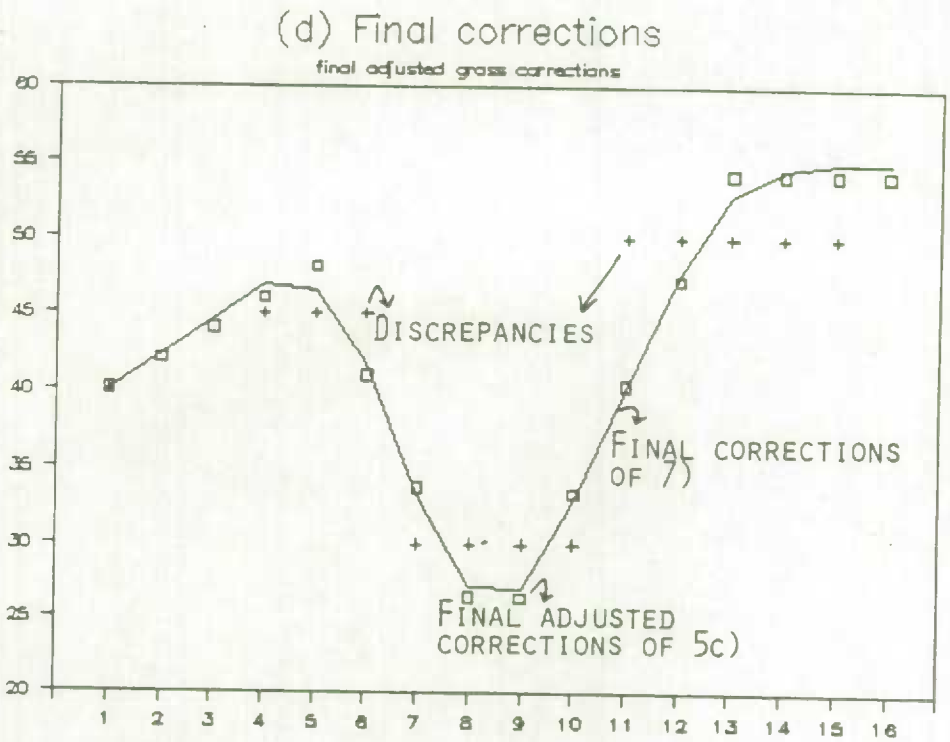


Figure D.1: - continuation

APPENDIX E: Solution to the simultaneous benchmarking problem

Using the correction approach suggested at the end of Appendix C, the objective function for the two-way classification simultaneous benchmarking problem of Section 5 (Part 1) is

$$F(C) = g_x\, C'D'D\, C + g_y\, (Y^{-1}B^*_y\, C - \iota_y)'(Y^{-1}B^*_y\, C - \iota_y) + G\, (B^*_n\, C)'(B^*_n\, C) + G\, (B^*_r\, C)'(B^*_r\, C) \tag{E.1}$$

where B^*_y is defined in Appendix A. Matrix D is now block diagonal difference operator of each component series, each bloc having dimensions T by T. Matrices B^*_n and B^*_r are similar to B^*_y. They are such that when multiplied by the corrections they generate the industrial and the regional discrepancies. When applicable, the grand total aggregation constraint is made implicit in one of them. The solution is:

$$C = (g_x\, D'D + g_y\, B^*_y{}'Y^{-1\prime}Y^{-1}B^*_y + G\, B^*_n{}'B^*_n + G\, B^*_r{}'B^*_r)^{-1}\,(g_y\, B^*_y{}'Y^{-1\prime}\iota_y) \tag{E.2}$$

The benchmarked values are then simply $z_t = x_t * c_t$.

APPENDIX F: Formal solution of the growth rate and the ARIMA interpolation problems

In matrix algebra, the objective functions (2.2) or (3.3) (of Part 2) write:

$$\text{(F.1)}\quad \begin{aligned} F(Z) = {} & g_x\, Z'D'D\, Z \\ & + g_y\, (Y^{-1}B_y\, Z - \iota_y)'(Y^{-1}B_y\, Z - \iota_y) \\ & + g_z\, (Z_d{}^{-1}B_z\, Z - \iota_z)'(Z_d{}^{-1}B_z\, Z - \iota_z) \end{aligned}$$

where the vectors and matrices are defined in Appendix A. Matrix D is the growth rate deviation operator matrix (A.9) or the ARIMA deviation operator (A.10). Symbols g_x, g_y and g_z are scalars. They could also be diagonal matrices of dimension T by T, M by M and K by K respectively. This would allow the specification of some annual benchmarks (for instance) as binding and some as non-binding. Performing the matrix multiplications in (F.1) yields:

$$\text{(F.2)}\quad \begin{aligned} F(Z) = {} & g_x\, Z'D'D\, Z \\ & + g_y\, (Z'B_y{}'Y^{-1}{}'Y^{-1}B_y\, Z - 2\, Z'B_y{}'Y^{-1}{}'\iota_y + \iota_y{}'\iota_y) \\ & + g_z\, (Z'B_z{}'Z_d{}^{-1}{}'Z_d{}^{-1}B_z\, Z - 2\, Z'B_z{}'Z_d{}^{-1}{}'\iota_z + \iota_z{}'\iota_z) \end{aligned}$$

The values of z_t which minimize this hyper-parabola are required. At the minimum, the derivative with respect to the unknown z_t are equal to zero. This leads to the normal equations:

$$\text{(F.3)}\quad \begin{aligned} dF/dZ = {} & 2\, g_x\, D'D\, Z \\ & + 2\, g_y\, B_y{}'Y^{-1}{}'Y^{-1}B_y\, Z - 2\, g_y\, B_y{}'Y^{-1}{}'\iota_y \\ & + 2\, g_z\, B_z{}'Z_d{}^{-1}{}'Z_d{}^{-1}B_z\, Z - 2\, g_z\, B_z{}'Z_d{}^{-1}{}'\iota_z = 0 \end{aligned}$$

The solution to this linear system of equations is unique:

$$\text{(F.4)}\quad \begin{aligned} Z = {} & (g_x\, D'D + g_y\, B_y{}'Y^{-1}{}'Y^{-1}B_y + g_z\, B_z{}'Z_d{}^{-1}{}'Z_d{}^{-1}B_z)^{-1} \\ & (g_y\, B_y{}'Y^{-1}{}'\iota_y + g_z\, B_z{}'Z_d{}^{-1}{}'\iota_z) \end{aligned}$$

The second order conditions for a minimum are satisfied because all the squared matrices in (F.2) are positive semidefinite. Note that the objective function and its solution are identical to that of proportional benchmarking, except matrix X^{-1} is absent.

BIBLIOGRAPHY ON BENCHMARKING AND INTERPOLATION

- Alba, E. de (1979), "Temporal Disaggregation of Time Series: a Unified Approach", New Mexico State university, Statistical Laboratory, Technical Report Series No 32.

- Alba, E. de (198?), "Temporal Disaggregation of Time Series: a Bayesian Analysis", Graduate School of Business, University of Chicago.

- Baldwin, A. (1980) "Extensions of Quadratic Minimization Methods", Gross National Product Division, Statistics Canada, Research Paper.

- Bassie, B.L. (1939), "Interpolation Formulae for the Adjustment of Index Numbers," Proceedings of the Annual Meetings of the American Statistical Association.

- Bell, W.R., Hillmer, S.C. (1983), "Modelling Time Series with Calendar Variations, Journal of the American Association, Vol. 78, pp. 526-534.

- Bishop, Y.M.M, Fienberg, S.E., Holland, P.W. (1975), Discrete Multivariate Data Analysis: Theory and Practice, M.I.T. Press.

- Brackstone, G.J., Rao, J.N.K. (1979), "An Investigation of Raking Ratio Estimators", The Indian Journal of Statistics (Sankya), Series C, Vol.41, Part2, pp 97-114

- Brackstone, J.J. (1987), "Issues in the Administrative Records for Statistical Purposes", Survey Methodology, Vol. 13, No. 1, pp. 29-44; "Utilisation des dossiers administratifs à fes fins statistiques", Techniques d'Enquête, Vol. 13, No 1, pp. 35-52.

- Boot, J.C.G., Feibes, W., Lisman, J.H.C. (1967), "Further Methods of Derivation of Quarterly Figures from Annual Data", Applied Statistics, Vol. 16, no. 1, pp. 65-75

- Boot, J.C.G., Feibes, W. (1967), "The Glejser Extension", Applied Statistics, Vol. 16, No. 2.

- Bournay, J., Laroque, G. (1979), Réflexions sur la méthode d'élaboration des comptes trimestriels", Annales de l'I.N.S.E., Vol. 36, pp. 3-30.

- Box, G.E.P., Jenkins, G.M. (1970), Time series Analysis, Forecasting and Control, Holden-Day.

- Brown, R.G. (1963), Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice-Hall.

- Bustros, J. (1986), "Reporting Arrangements Reference Period", Statistics Canada, Sub-Project Team 4.

- Chatfield, C. (1975), The Analysis of Time Series: Theory and Practice, Wiley.

- Cholette, P.A. (1978),"A Comparison and Assessment of Various Adjustment Methods of Sub-Annual Series to Yearly Benchmarks", Time Series Research and Analysis, **Statistics Canada**, Research Paper 78-03-001B;
"Comparaison et évaluation de quelques méthodes d'ajustement de séries infra-annuelles aux repères annuels", Séries chronologiques, **Statistique Canada**, document de recherche 78-03-001B.

- Cholette, P.A. (1979), "Adjustment Methods of Sub-Annual Series to Yearly Benchmarks", **Proceedings of the Computer Science and Statistics, 12th Annual Sumposium on the Interface**, J.F. Gentleman Editor, University of Waterloo, pp. 358-366.

- Cholette, P.A. (1982), "Prior Information and ARIMA Forecasting", **Journal of Forecasting**, Vol. 1, pp. 375-383.

- Cholette, P.A. (1984), "Adjusting Sub-Annual Series to Yearly Benchmarks, **Survey Methodology**, Vol. 10, pp. 35-49;
"L'ajustement des séries infra-annuelles aux repères annuels", Techniques d'enquêtes, Statistique Canada, Vol. 10, no. 1, pp. 39-53.

- Cholette, P.A. (1985a), "Estimating Sub-Annual Series from Growth Rates and Annual Benchmarks", Time Series Research and Analysis, **Statistics Canada**, Research Paper RASC-85-018B;
"L'estimation des séries infra-annuelles à partir de taux de croissance et de jalons annuels", Séries chronologiques, **Statistique Canada**, Cahier de travail RASC-85-018B

- Cholette, P.A. (1985b), "Linking Heterogeneous Segments of Seasonal Time Series", Proceedings of the Business and Economic Statistics Section, **American Statistical Association**, pp. 178-183.

- Cholette, P.A., Baldwin, A. (1988), "Coverting Annual Data from a Fiscal Year Basis to a Calendar Year Basis" Time Series Research and Analysis, **Statistics Canada**, forthcoming.

- Cholette, P.A., Piche, R. (1986), "Views on Benchmarking and Interpolation", Time Series, **Statistics Canada**, Working Paper no. TRSA-86-018E

- Cholette, P.A. (1987a), "Correcting the Reference Periods of Annual and Quarterly data", **Statistics Canada**, Time Series Research and Analysis Division, Working Paper No. TRSRA-87-010E.

- Cholette, P.A., Higginson, J. (1987), "Converting Bundles of Weekly Data into Seasonal Monthly Values", **Statistics Canada**, Time Series Research and Analysis Division, Research Paper forthcoming.

- Cholette, P.A., Higginson, J. (1987), "A Note on Aggregating Data Referring to Different Financial Years", Time Series Research and Analaysis Division, Research Paper, ask for number

- Cleveland, W.S., Devlin, S.J. (1980), "Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphic Methods", **Journal of the American Association**, Vol. 75, pp. 487-496.

- Cleveland, W.S., Devlin, S.J. (1982), "Calendar Effects in Monthly Time Series: Modelling and Adjustment", **Journal of the American Association**, Vol. 77, pp. 520-528.

- Cleveland, W.S., Grupe, M.R. (1981), "Modelling Time Series when Calendar Effects are Present", in: Arnold Zellner, ed., **Proceeding of the Conference on Applied Time Series Analysis of Economic Data**, U.S. Department of Commerce, Bure of the Census, pp. 57-67.

- Chow, G.C., Lin, An-loh (1971), "Best linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series", **Review of Economics and Statistics**, Vol. 53, No. 4, pp. 372-375.

- Cohen, K.J. et al. (1971) "Autoregressive approaches to the disaggregation of Time Series Data", **Applied Statistics**, Vol. 20, pp 119-129.

- Cohen, K.J., Padberg (1974), "Regressive approaches to Temporal Disaggregation" **Proceedings of the Social Statistics Sections**, American Statistical Association.

- Dagum, E.B. (1977), "Comparison of Various Interpolation Procedures for Benchmarking Economic Time Series", **Séries chronologiques, Statistique Canada**, document de recherche 77-05-006E.

- Dagum, E. (1980), **The X-11-ARIMA Seasonal Adjustment Programme**, Statistics Canada, Cat. 12-564E; aussi disponible en français.

- Denton, F.T. (1971), "Adjustment on Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization", **Journal of the American Statistical Association**, Vol. 66, No. 333, pp. 99-102.

- Elhawary-Rivet, N. "A Comparison of Two Benchmarking Methods", **Statistics Canada**, Industry Measures and Analysis Division, Working Paper 87:01.

- Feibes, W. (1968), "Interpolation of Time Series", Ph. D. Thesis, **State University of New York at Buffalo**, Department of Mathematics.

- Fernandez, R.B. (1981), "A Methodological Note on the Estimation of Time Series", **Review of Economic and Statistics**, Vol. 63, pp. 471-476.

- Fienberg, S.E. (1977), **The Analysis of Cross-Classified Data**, M.I.T. Press.

- Friedman, M. (1962), "The Interpolation of Time Series by Related Series", **Journal of the American Statistical Association**, Vol. 57, No. 300, pp. 729-757.

- Glejser, H. (1966), "Une methode d'évaluation des données mensuelles à partir d'indices trimestriels ou annuels", **Cahier Economiques de Bruxelles**, No. 19, 1er trimestre, pp. 45-64.

- Ginsburg, V.A. (1973), "A Further Note on the Derivation of Quarterly Figures Consistent with Annual Data", **Applied Statistics**, Vol. 22, No. 3, pp. 368-374.

- Guerrero, V. M. (1987) "An ARIMA-Based Solution to Temporal Disaggregation Problem, **Banco de México**, D.F.

- Haberman, S.J. (1974), **The Analysis of Frequency Data**, University of Chicago Press.

- Harvey, A.C. (1984), "A Unified View of Statistical Forecasting Procedures", **Journal of Forecasting**, Vol. 3, pp. 245-275.

- Harvey, A.C., Pierce, R.G. (1984) "Estimating Missing Observations in Econonomic Time Series", **Journal of the American Association**, Vol. 79, pp. 125-131.

- Hillmer, S.C., (1982), "Forecasting Time Series with Trading-Day Variation", **Journal of Forecasting**, Vol. 1, pp. 385-395.

- Hillmer, S. C., Trabelsi, A. (1987), "Benchmarking of Economic Time Series", **Journal of the American Statistical Association**, Vol. 82, pp. 1064-1071.

- Kendall, M.G. (1973), **Time Series**, Haffner

- Kohn, R., Ansley, G.F. (1986), "Estimation, Prediction and Interpolation for ARIMA Models with Missing Data", **Journal of the American Statistical Association**, Vol. 79, pp. 125-131.

- Koopmans, L.H. (1974), **The Spectral Analysis of Time Series**, Academic Press.

- Laniel, N. (1986), "Adjustment of Economic Production Sub-Annual Series to Corresponding Annual Series", Business Survey Methods, **Statistics Canada**, Working Paper no. BSMD-86-006E

- Laniel, N. (1987), "Benchmarking of Business Surveys: Problem Discussion and Strategy to Derive a Solution", **Statistics Canada**, Business Survey Methods Division, July 1987.

- Leser, C.E.V. (1961), "A Simple Method of Trend Construction", **Journal of the Royal Statistical Society, Series B**, pp. 97-1-7

- Leser, C.E.V. (1963), "Estimation of Quasi Linear Trend and Seasonal Variation", **Journal of the American Statistical Association**, Vol. 58, pp. 1033-1043.

- Lisman, J.H.C., Sandee, J. (1964), "Derivation of Quarterly Figures from Annual Data", **Applied Statistics**, Vol. 13, No. 2, pp. 87-90.

- Litterman, R.B. (1983), "A Random Walk, Markov Model for the Distribution of Time Series", **Journal of Business and Economic Statistics**, Vol. 1 No. 2, pp., 169-173.

- Little, R.J.A., Rubin, D.B. (1987), **Statistical Analysis with Missing Data, John Wiley and Sons.**

- Liu, L.-M., (1986), "Identification of Time Series Models in the Presence of Calendar Variation", **International Journal of Forecasting**, Vol. 2, No. 3, pp. 357-372.

- Liu, L.-M., (1980), "Analysis of Time Series with Calendar Effects", **Management Science**, Vol. 26, No. 1, pp. 106-112.

- Monsour, N.J., Trager, M.L. (1979), "Revision and Benchmarking of Business Time Series", **Proceedings of the Business and Economic Statistics Section, American Statistical Association.**

- Nasse, P. (1973), "Le système des comptes nationaux trimestriels", Annales de l'I.N.S.E.E., Vol 14, pp. 119-165.

- Nijman, T.e., Palm, F.C. (1986), "The Construction and Use of Approximation for Missing Quarterly Observations: A Model-Based Approach, **Journal of business and Economic Statistics**, Vol. 4, pp. 47-57

- Pfefferman, D., Fisher, J. (1980), "Festival and Working Days Prior Adjustments in Economic Time Series", in: O.D. Anderson, **Time series**, North-Holland, New York.

- Raveh, A. (1986), "On Measures of Monotone Association" **The American Statistician**, Vol. 40, No. 2, pp. 117-123

- Sanz, R. (1981), "Metodos de desagregacion temporal de series economicas", **Banco de Espana**, Servicio de Estudios, Serie de estudios economicos no 22; also available in English under title "Temporal Disaggregation Methods of Economic Time Series".

- Schlicht, E. (1981), "A Seasonal Adjustment Principle and a a Seasonal Adjustment Method Derived from this Principle", **J.A.S.A.**, Vol. 76, No 374, pp. 374-378.

- Shiskin, J., Young, A.H., Musgrave, J.C. (1967), **The X-11 Variant of the Census Method II Seasonal Adjustment Program**, U.S. Bureau of the Census, Technical Paper No. 15.

- Somermeyer, W.H., Jansen, R., Lauter, A.S. (1976), "Estimating Quarterly Values from Annually Known Variables in Quarterly Relationships", **Journal of the American Statistical Association**, Vol. 71, No. 355, pp. 588-595.

- Taillon, J. (1987), "Balance d'un tableau avec contraintes linéaires: Application aux Flux financiers canadiens désaisonnalisés", **Statistique Canada**, Division des Mesures et de l'analyse des industries, document de recherche.

- Theil, H. (1972). **Statistical Decomposition Analysis**, North-Holland.

- Vangrevelinghe, G. (1966), L'évolution à cour terme de la consommation des ménages, Etudes et Conjoncture, I.N.S.E.E., pp. 59-102.

- Wilcox, J.A. (1983), "Disaggregating Data Using Related Series", Journal of Business and Economic Statistics, Vol. 1. No. 3, pp. 187-191.

- Young, A.H. (1965), "Estimating Trading-Day Variation in Monthly Economic Time Series", U.S. Bureau of the Census, Technical Paper No. 12.

- Young, A.H. (1968), "Linear Approximations to the Census and BLS Seasonal Adjustment Methods", J.A.S.A., Vol. 63, No 322, pp. 445-471.

c. 2