Working Paper No. TSRA-88-019E
Time Series Research and Analysis Division
Methodology Branch

CONVERTING SETS OF WEEKLY DATA INTO SEASONAL MONTHLY VALUES

by

Pierre A. CHOLETTE[1] and Norma B. CHHAB[2]

Revised Version September 1988

A condensed version of this document
was submitted for publication to
Applied Statistics

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

(613) 951-0050

[2]Banco Central de la Republica Argentina
Gerencia de Investigaciones
y Estadisticas Economicas
Reconquista 266
1003 BUENOS AIRES
Republica Argentina

(541) 394-9991

La mensualisation
des chiffres hebdomadaires

- résumé -

Ce document présente une méthode pour transformer en valeurs mensuelles des chiffres couvrant un nombre variable de jours ou de semaines, typiquement de 4 ou 5 semaines consécutives. La méthode consiste à interpoler de tels chiffres en valeurs quotidiennes et à réagréger celles-ci en valeurs mensuelles (ou hebdomadaires). L'interpolation est en fait une application particulière d'une variante appropriée de la méthode d'étalonnage de Denton (1971): un profil quotidien, c'est-à-dire une suite hebdomaraire de pondérations quotidiennes relatives, est ajusté de sorte à se conformer aux chiffres disponibles.

- summary -

This paper presents a method to convert data referring to a varying number of days or weeks - typically 4 or 5 weeks - into monthly values. The method consits of interpolating daily values from such data and of re-aggregating these values into monthly (or weekly) values. The interpolation is in fact a particular application of a appropriate variant of the Denton (1971) benchmarking method: a daily pattern, i.e. a weekly sequence of relative daily weights, is adjusted to comply to the available data.

KEYWORDS: Interpolation, Benchmarking, Weekly Data, Calendarization, Quadratic minimization

1. INTRODUCTION

Much of the monthly data published by statistical agencies actually originate in the form of **weekly bundles of data**, that is of data covering a number of weeks of activity. In the Canadian retail and wholesale trade sectors for instance, many respondent companies send their data in bundles covering four or five weeks. The bundles typically cover 4, 4 and 5 weeks, 4, 4 and 5 weeks, and so forth 12 times per year; or 4 weeks, 13 times per year. It is then the task of the central statistician to adjust such data to reflect months. The problem, illustrated in Section 2, is especially accute when bundles end in the middle of a month. This paper provides statisticians with a rigourous method to perform that operation.

Calendarization, which is the generic term for the operation, is viewed as a benchmarking problem. Benchmarking usually consists of adjusting sub-annual measurements of a socio-economic variable so that they comply with annual, more reliable and separately obtained measurements of the variable. For the application considered, the "sub-annual" series is a pre-determined **daily trading pattern**, that is a weekly sequence of daily weights (Young, 1965); and the benchmarks are the available weekly bundles of data. This particular application of benchmarking results into interpolated daily values. The actual calendarization consists of aggregating the interpolated daily values into the desired monthly estimates. Section 3 adapts the benchmarking approach by Denton (1971) and Cholette (1984, 1987b) for the purpose of calendarization.

Section 4 discusses revisions and the obtention of current monthly estimates. Section 5 compares the method proposed in this paper to some procedures presently used in statistical agencies to calendarize weekly bundles of data. Section 6 discusses implementational issues, like computational considerations, and is followed by a summary and conclusion in Section 7. The structure of the document is somewhat inspired by Ehrenberg (1982), in that results are presented as early as possible and followed by methodological details.

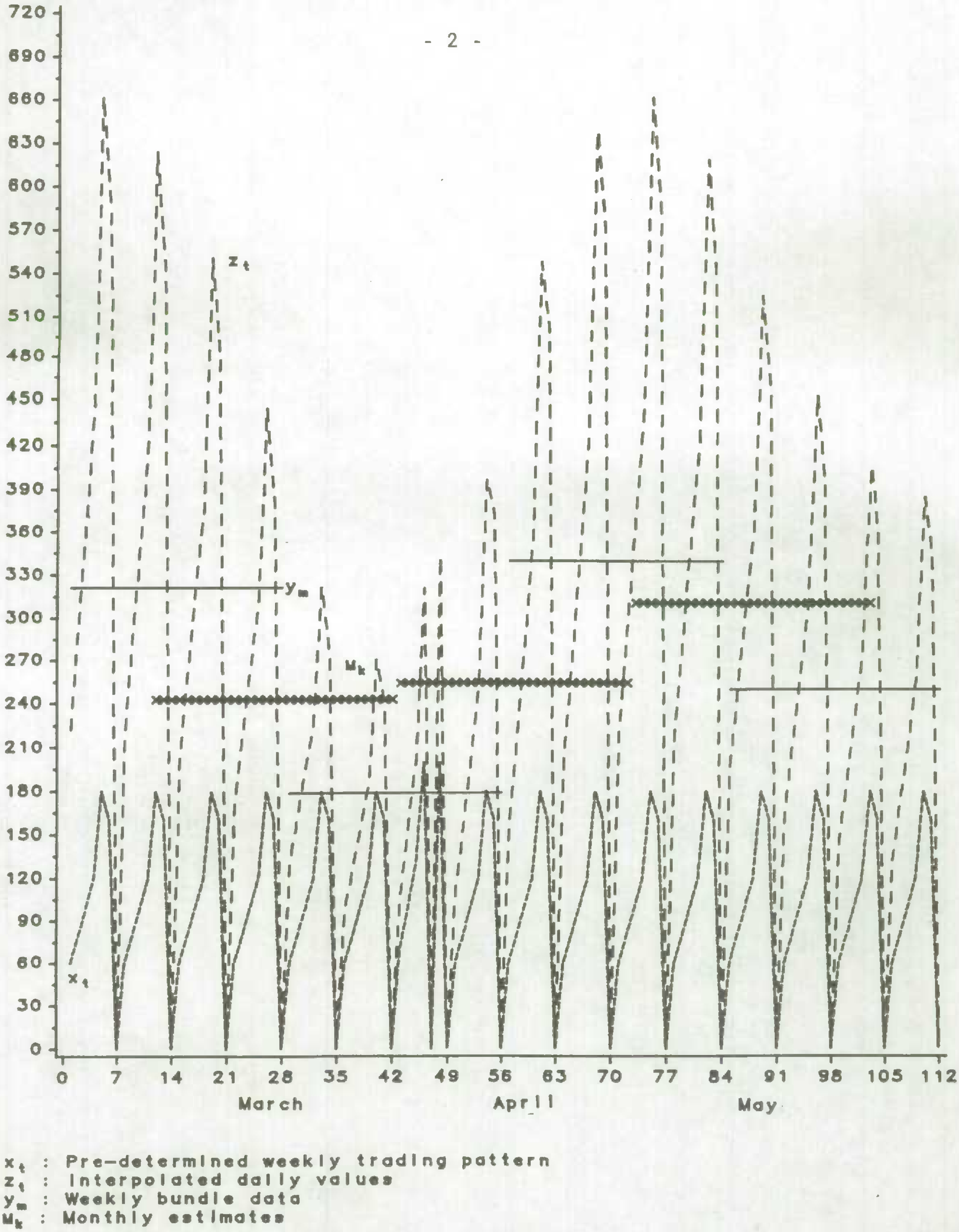


Figure 1: Converting Bundles of Weekly data into Monthly Values: (1) Interpolating daily values z_t from the weekly bundles y_m and a daily pattern x_t ; (2) Recombining the interpolated daily values into the desired monthly estimates $y^c{}_k$

2. CONCEPTS AND DEFINITIONS AND ILLUSTRATION OF THE APPROACH

Figure 1 displays four **bundles of weekly data** y_m for a flow series. (Flow are such that the monthly values are the sum of the daily values in the month.) The weekly bundles are represented by their **daily average values** (9000/28, 5000/28, 9500/28 and 7000/28) over the periods they cover, that is over their **reference periods**. This practice makes them graphically comparable to the interpolated daily values and also indicates their reference periods (by projection on the horizontal axis). The first bundle covers four weeks ranging from February 18 to March 17 (reference periods 1 to 28); the second, from March 18 to April 14 (periods 29 to 56); the third, from April 15 to May 12 (57 to 84); and fourth from May 13 to June 9 (85 to 112). Those weekly bundles have to be converted in monthly values for March, April and May.

Figure 1 also displays a pre-determined **daily trading pattern** x_t, which basically consists of a repeated weekly sequence of 7 **daily weights**: 60, 80, 100, 120, 180, 160 and 0.001. The weights are chosen to reflect 40% less activity on Mondays than on average of the week; 20% less, on Tuedays; 30% more, on Wednesdays; and so forth; and negligible activity on Sundays. The **interpolated daily values** z_t are obtained by keeping them as proportional as possible to the selected daily pattern x_t and perfectly consistent with the available bundle data y_m. (More details in Section 3.) The desired **monthly estimates** M_k, also represented by their daily averages, are derived by simply taking the monthly sums of the daily interpolations. Figure 1 shows that there is a considerable difference between the original bundle values and the monthly estimates.

As mentionned, the daily pattern reflects the relative importance of the days of a week considered. It is generally more convenient to choose the daily weights so that their weekly average is 100%, as in Figure 1. However this is not a pre-requisite. If available, one should preferably use daily measurements of the socio-economic variable considered, which for some reason do not comply with the more reliable bundle values. The daily pattern may also be borrowed from a related socio-economic variable. By extension, the daily pattern does not have to be stable. In Figure 1 for instance, the normal Friday activity of Good Friday (period 47) is specified to take place on the preceeding days and on the Easter Saturday: For the week of Easter, the daily weights are 70, 100, 130, 200, 0.001 (Good Friday), 200 and 0.001 instead of 60, 80, 100, 120, 180, 160 and 0.001.

The daily pattern is appropriate to incorporate **subject matter expertise** into the calendarization procedure. That expertise is the intimate knowledge by series builders of the processes and variables, which influence the socio-economic variable considered; and of the operational circumstances, in which the data is collected. For instance, the subject matter expertise may simply consists of knowing that the manufacturing firms considered are closed during the week-ends. In this case, valid daily weights would be 140% for Mondays to Fridays (700%/5) and 0.0001% for Saturdays and Sundays. (The method does not allow zero weights.) Calendarization may be performed at low levels of dis-aggregation; and at that level, subject matter expertise may be the only daily information available.

3. PROPOSED CALENDARIZATION METHOD

This Section presents the mathematics of the proposed calendarization method just illustrated. A general solution is presented first; and then, a simplified solution applicable to stock series.

3.1 General Solution

The interpolation of daily estimates is achieved by means of constrained mathematical optimization. Namely the following objective function is minimized

$$f(z) = \sum_{t=2}^{T} (z_t/x_t - z_{t-1}/x_{t-1})^2, \qquad (3.1)$$

subject to constraints

$$\sum_{t=\tau_m}^{\rho_m} z_t = y_m, \quad m=1,..,M. \qquad (3.2)$$

Variable z_t stand for the interpolated T daily values to be estimated; and x_t for the daily pattern. Parameter M (not to be confoused with M_k) is the number of weekly bundles y_m in the **series interval** considered. Parameters τ_m and ρ_m indicate the reference periods of the bundle values (e.g. in Fig. 1, $\tau_1=1$, $\rho_1=28$, $\tau_2=29$, $\rho_2=56$, $\tau_3=57$, $\rho_3=84$, $\tau_4=85$ $\rho_4=112$).

The objective function (3.1) specifies the **proportional movement preservation principle**: The proportion of the desired interpolated values z_t to the daily pattern values x_t is to remain as constant as possible. In other words the proportion z_t/x_t should change as gradually as possible from one period to the next. (Although defendable per se, the proportional movement preservation criterion is also a linear approximation of the non-linear growth rate preservation criterion.) The extent to which movement preservation is achievable is determined by the constraints (3.2). These specify that the interpolated daily values must sum to the bundle data over their reference periods.

The solution to this optimization problem is developed in Appendix A. The daily estimates are weighted averages of the bundle data:

$$z_t = \sum_{m=1}^{M} w_{t,m}\, y_m \quad t=1,...,T. \qquad (3.3)$$

The calculation of the weights w_{tm} require matrix algebra operation..The desired monthly estimates M_k are then simply the monthly sums of the daily values.

The interpolation can also be achieved with the additive variant of the movement preservation principle, i.e. parallelism. The divide signs (/) in objective function (3.1) are replaced by minuses. The movement in z_t is then to be as close as possible to that in x_t. The solution developped in Appendix A also encompasses the additive variant. That variant would be appropriate (and so would the proportional), when the daily pattern consists of daily measurements of the socio-economic variable considered and has the same order of magnitude as the desired interpolated values. If this is not the case, the additive variant is unadvisable, for instance when the daily pattern consists of percentages like in Figure 1.

The objective function of benchmarking is sometimes based on second proportional differences. For calendarization at least, we advocate first differences for the vast majority of socio-economic time series, especially when dealing with sub-annual series (Cholette and Baldwin, 1988) like in this paper.

3.2 Simplified Solution for Stock Series

Stock series pertain to levels of a variable at a particular date (e.g. population, unemployment series.) For the purpose of this paper, stock series are those whose monthly values correspond to one of the daily values. For instance, in Canada the monthly values of the bank rate refer to the value on the last Wednesday of the month. The method presented in the previous section works: The reference periods of the bundles are simply set so that $\tau_m = \rho_m$. The daily estimates are then weighted averages of the fiscal year data, as in equation (3.3). However, the solution exactly reduces to:

$$z_t = x_t * p_t, \tag{3.4}$$

where

$$p_t = d_m + (t-\tau_m)*[(d_{m+1}-d_m)/(\tau_{m+1}-t_m)], \quad \tau_m \le t \le \tau_{m+1}, \; m=1,\ldots,M-1, \tag{3.5}$$

$$p_t = d_1, \; t \le \tau_1; \quad p_t = d_M, \; t \ge \tau_M. \tag{3.6}$$

The desired daily estimates z_t are equal to the corresponding daily pattern value x_t times a correction factor p_t. These corrections are linear interpolations between the **proportional discrepancies** d_m on each side of a time period considered. These discrepancies are the ratios of the bundle value to the applicable daily pattern value, $d_m = y_m/x_{\tau m}$ (read subscript τm as τ_m). For the periods preceeding the first and following the last discrepancies, the corrections repeat the values of first and last discrepancies. The calendarized values are simply equal to the applicable sub-annual values of z_t (e.g. all the last Wednesdays of each month). The solution for stock series completely avoids matrix algebra operations. The solution also works for the additive variant, if the multiply sign is replaced by the add sign in (3.4) and if the discrepancies are redefined as $d_m = y_m - x_{\tau m}$ (read τ_m)

Solution (3.4) to (3.6) is akin to some of the interpolation methods described by Friedman (1962).

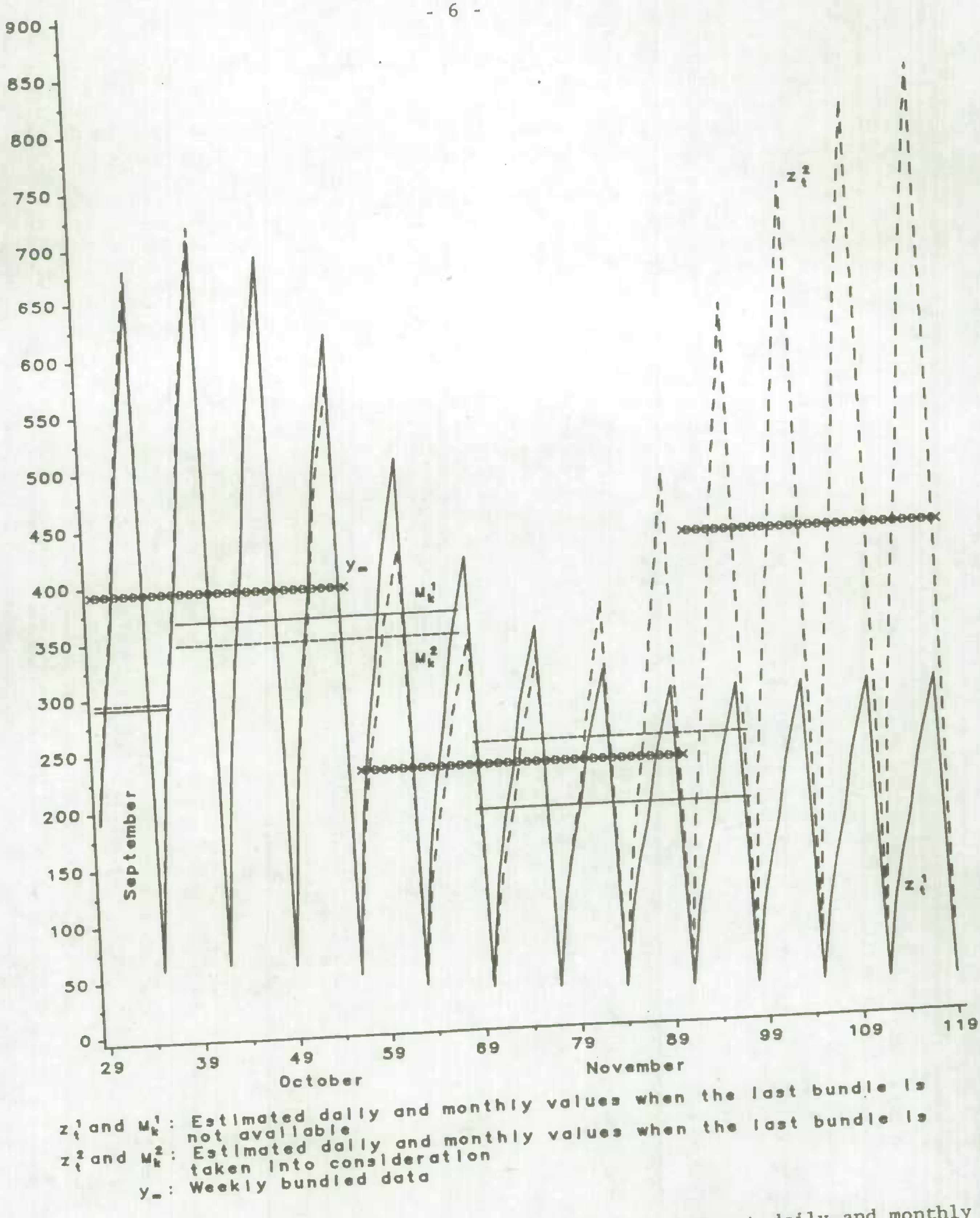


Figure 2: Revisions occurring to the current and recent daily and monthly estimates $z^1{}_t$ and $M^1{}_k$ on adding a new bundle of weekly data

4. CURRENT ESTIMATES AND REVISIONS

The calendarization method proposed can produce daily and monthly estimates for time periods not covered by any weekly bundle of data. This is illustrated in Figure 2 by the daily and monthly estimates z^1_t and M^1_k obtained when the last bundle (pertaining to periods 92 to 119) is not yet available, i.e. not considered in the estimation. The last 7 days of November are then not covered by any bundle. The corresponding daily estimates - and to some extent the November estimate - are then extrapolations, as opposed to interpolations. Extrapolations are subject to heavy revisions. Indeed, in the figure, the extrapolated (and the last interpolated) daily values implicitly assume that the daily average of the next weekly bundle will be lower then the previous one. More generally at the end of series, the method always anticipates the next bundle daily average to be lower, if the last one was lower than the previous one (case illustrated); and higher, if the last one was higher than the previous one. In the presence of seasonal, trading-day and trend-cycle fluctuations, this kind of systematic anticipation is the wrong one, at the sub-annual level at least.

Figure 2 illustrates the revisions to the daily and monthly estimates when the last bundle pertaining to periods 92 to 119 becomes available, i.e. is considered in the estimation. The daily estimates go from z^1_t to z^2_t. The September revised estimate is now 8845 instead of 8725 originally, i.e. 1.4% higher; the October estimate, 10743 instead of 11365, i.e. 5.5% lower; and the November extrapolated estimate 7430 instead of 5664, i.e. 31.2% higher. The extrapolated estimates (an perhaps the last interpolated estimate) should therefore not be used. In other words, daily and monthly estimates not embedded in the reference periods of the weekly bundles should be ignored. This implies delays in the production of the monthly estimates by the statistical agency. These delays may be circumvented by forecasting the next one or two bundles and by calendarizing the bundle series extended by the forecasted bundles. (This practice has proved to reduce revisions in seasonal adjustment, Dagum, 1980.) Even a very ordinal forecast of the next bundle (e.g. lower, equal or greater) would reduce the huge 31.2% revision of November in Figure 2.

4.1 Forecasting Bundles

Perhaps an ordinal bundle forecast can be obtained by asking the respondents for a gross anticipated value for their next bundle (e.g. +10% increase with respect to the current bundle, no change, -20%).

An alternative to generate gross - and possibly not so gross - anticipations is to forecast the next bundle. If all the bundles y_m over the last few years cover 4 weeks (i.e. there are 13 bundles per year), forecasted bundles are obtainable by ARIMA modelling (Box and Jenkins, 1970). Simple ARIMA models could provide suitable forecasts, namely the following seasonal autoregressive model $(1,0,0)\,(1,0,0)_{13}$:

$$(4.1) \qquad y_m - \phi\, y_{m-1} = \Phi(y_{m-13} - \phi\, y_{m-14}) + a_t, \quad \phi \text{ and } \Phi > 0.$$

For ϕ and Φ equal to 1, the model states that the change between two bundles y_m and y_{m-1} of one year tends to be equal to the change between the corresponding y_{m-13} and y_{m-14} of the previous year. In other words, the change between two bundles tends to repeat from year to year, reflecting constant seasonal amplitude. Similarly, for $\phi=1$ and $\Phi=1.05$, the model states that the change between two bundles tends to increase by 5% from year to year, reflecting increasing seasonal amplitude; for $\phi=1$ and $\Phi=0.95$, to decrease by 5%, reflecting shinking seasonal amplitude. The model thus covers a wide variety of series behaviour, which can be assessed and monitored by subject matter experts. The simple model (4.1) also explains a very large portion of the variance of most socio-economic time series.

The parameters of the seasonal autoregressive model (4.1), or of any other ARIMA model, may be estimated. (Some programmes like SCA by Liu and Hudak, 1983, do allow autoregressive parameters greater than 1.) However since only gross forecasts are required, it could be sufficient and certainly more practical to select the parameter values, on the basis of the implied behaviour of the series (as discussed); and on the basis of whether the forecasts reduce the revisions. The forecasts corresponding to model (4.1) are given by:

$$(4.2) \qquad y^f_m = \phi\, y_{m-1} + \Phi y_{m-13} - \phi\Phi\, y_{m-14}.$$

More than one forecast may be generated by substituting forecasts to y_{m-k} in (4.2) when y_{m-k} is not available. The bundle series to be calendarized is then extended by one or more bundles.

If the bundles cover 4, 4 and 5 weeks in a repetitive manner, the same strategy is applicable, except the seasonal autoregressive model is of order 12 instead of 13: $(1,0,0)(1,0,0)_{12}$. The forecasts are then given by:

$$(4.3) \qquad y^f_m = \phi\, y_{m-1} + \Phi y_{m-12} - \phi\Phi\, y_{m-13}.$$

If the reference periods of the bundles do not conform to any repetitive pattern, the bundles may be converted (using the method proposed) into 4-weeks bundles, instead of months, for the time periods embedded in the original bundles. The first forecasting strategy then become applicable to the estimated 4-week bundles. One or two 4-week bundles are forecasted, which are not (totally) embedded in the reference periods of the original bundles. If the first forecasted bundle partially overlaps with the last original bundle, both bundles are specified as constraints. In other words, the interpolated daily values satisfy both overlapping bundles. (The proposed method does not prevent bundles with overlapping reference periods.)

4.2 Specifying the behaviour of the monthly values in the objective function

The most direct way to reduce revisions of the last monthly estimates would be to specify the seasonal and trading-day behaviour to be followed by the monthly values M_k in the objective function, much in the same way as for the daily values. Let ξ_k stand for the product of a pre-determined monthly seasonal pattern s_k and of a pre-determined monthly trading-day pattern ζ_k:

$$\xi_k = s_k \zeta_k = s_k \left(\sum_{t=\eta_k}^{\theta_k} x_t \right) / n_k, \tag{4.4}$$

where ξ_k is then a known monthly seasonal-trading-day pattern. Parameters η_k and θ_k are the reference periods of month k. Parameter n_k is set equal to the number of days in the month, when the length-of-month effect is assigned to the seasonal component s_k. Parameter n_k is set equal to the average number of days in the year 30.4375 (365.25/12), when that effect is assigned to the trading-day component. (In the absence of seasonality, i.e. $s_k=1$, , one must select 30.4375. For more details refer to Young, 1965, and to Bell and Hillmer, 1982.) The values of the seasonal and the trading-day components, s_k and ζ_k, may be carried over from the previous year (for which it would generally have been calculated). Indeed the intra-year movement of time series is dominated by seasonal and trading-day variations; and by definition, these change very little from year to year.

The pre-determined seasonal-trading-day pattern is explicitely incorporated in the following objective function:

$$\begin{aligned} f(z) &= \sum_{t=2}^{T} (z_t/x_t - z_{t-1}/x_{t-1})^2 \\ &+ \sum_{k=1}^{K} \left[\left(\sum_{t=\eta k}^{\theta_k} z_t \right) / \xi_k - \left(\sum_{t=\eta k-1}^{\theta_{k-1}} z_t \right) / \xi_{k-1} \right]^2, \quad \theta_K \le T. \end{aligned} \tag{4.5}$$

This modified objective function is still minimized subject to constraints (3.2). The first term specifies that the daily values z_t are to display the pre-selected daily pattern x_t; and the second, that the last K monthly values are to display the pre-determined seasonal-trading-day pattern ξ_k. Both criteria are specified by means of the proportional movement preservation principle. Note that the reference period of the last month may not be embedded in the reference periods of the weekly bundles ($\theta_K \le \rho_M$ or $\theta_K > \rho_M$). Parameter K is the number of months at the end of the series interval, for which the seasonal-trading-day criterion is specified. It is not desirable to specify that criterion for more than two months (at the start or) at the end of the series. This would tend to create a tautology between the seasonal-trading-day factors supplied to calendarization and those obtained in seasonal adjustment.

The formal solution to that modified calendarization method is provided in Appendix B. Unfortunately it is computationally much more expensive than the solution for the original problem of Section 3. Since only gross anticipations are desired, perhaps the following approximation to minimizing (4.5) would be sufficient:

1) Calendarize with the variant of Section 3 and obtain monthly values embedded in the reference periods of the weekly bundles.

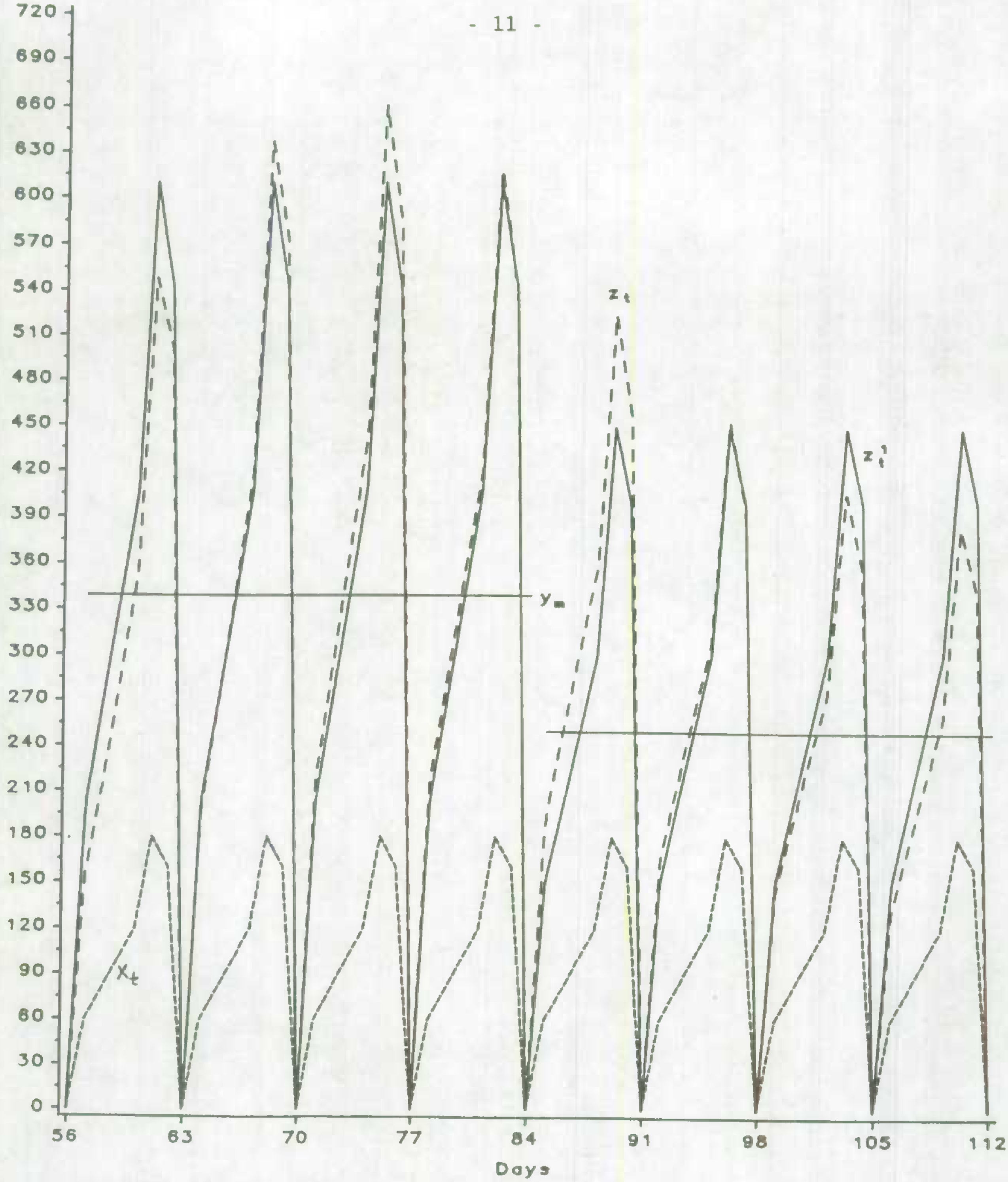
2) Forecast one or two monthly values by applying to the last monthly value in 1) the corresponding growth rates in the pre-determined seasonal-trading-day pattern.

3) Recalendarize (with variant of Section 3) the bundle series extended with the forecasts obtained in 2) treated as a bundle. As explained in Section 4.1, the reference periods of bundles may overlap.

4.3 Revision Policy

This section raised the issue of revisions. More attention was paid to the extrapolated estimates (e.g. November). However, the interpolated estimates near the end of the series, namely October and September in Figure 2, are also subject to substantial revision. The last monthly value obtained (with or without forecasting) should be recalculated and revised on the availability of the next bundle. This improves its reliability. The estimate for the month considered could be revised again on the availability of an extra bundle, because the reliability of the estimates improves as they become more central in the series interval. That second revision is less imperative however: as bundles are added to the series, the revisions become negligible, as illustrated in Figure 2, and may be ignored. The first revision is imperative; the second, desirable; and the number of extra revisions is more a matter of revision policy of the statistical agency.

This section cautionned against revisions of current and recent calendarized estimates and suggested avenues that could reduce the size of the revisions. This caution and theses avenues are applicable to other calendarization methods used in statistical agencies.



x_t : Pre-determined weekly trading pattern
z_t : Interpolated daily values according to the method proposed
y_m : Weekly bundle data
z_t^1 : Interpolated daily values implicit in a method commonly used by statistical agencies

Figure 3: Comparison of the proposed calendarization method to another method sometimes used in statistical agencies

5. CALENDARIZATION PROCEDURES CURRENTLY USED

As implied in the Introduction, some procedures are now currently being applied by central statisticians to convert weekly bundles of data into monthly values. Some existing procedures are now reviewed and compared to the method proposed in this paper.

5.1 Period Adjustment Procedure

One procedure is described and referred to as period adjustment by Miller (1986). The estimate M of a month depends on only one bundle value y. Therefore, the bundle has to overlap as much as possible with the target month. In order to facilitate the procedure, the statistical agency does not accept bundles ending in the middle of months or not including important events like Christmas. The respondent is instructed to delay sending such bundles and to include one or two extra weeks therein. The procedure is also based on pre-determined daily weights x_t.

1) The actual sum of the daily weights over the reference periods of a bundle is calculated, for instance:

$$\sum_{t=1}^{28} x_t. \tag{5.1}$$

2) The desired sum of the daily weights over the reference periods of the target month is calculated, e.g.

$$\sum_{t=1}^{31} x_t. \tag{5.2}$$

3) The monthly value M is the product of the bundle values and the ratio of the desired to the actual sum of weights, i.e.

$$M = y \left[\left(\sum_{t=1}^{31} x_t\right) / \left(\sum_{t=1}^{28} x_t\right)\right]. \tag{5.3}$$

In order to compare it with the method proposed in this paper, the period adjustment procedure is now described in terms of implicit interpolated daily values. The procedure is equivalent to interpolate daily values; and to extrapolate daily values, if the target month is not all included in the bundle (like in the example chosen). These estimated daily values z_t are equal to the product of the daily pattern x_t and of the average daily value (e.g. y/28) of the bundle considered, for instance.

$$z_t = (y / 28)\ x_t,\ t=1,\ldots,31. \tag{5.4}$$

The monthly value M is the sums of the interpolated daily values, e.g.

$$M = \sum_{t=1}^{31} z_t. \tag{5.5}$$

The equivalence between (5.5) and (5.3) is proven by substituting (5.4) into (5.5):

$$M = \sum_{t=1}^{31} (y / 28)\, x_t = y/28 \left(\sum_{t=1}^{31} x_t \right) = y \left[\left(\sum_{t=1}^{31} x_t \right) / \left(\sum_{t=1}^{28} x_t \right) \right].$$

(This proof assumed that the daily weights x_t have a weekly sum of 7; and therefore a sum of 28 over 28 days. If this not the case, the algebra is a bit more complicated but the same result is achieved, as the standardization occurs in the ratio in 3).)

This interpolation perspective reveals the assumptions of the period adjustment procedure:

i) The seasonal and trend-cycle variations are constant, that is inexistent, within a bundle. Indeed, the interpolations in (5.4) are merely a rescaled daily pattern (by factor y/28).

ii) The extrapolated daily values (for days 29 to 31 in the example) also imply that the assumption holds between the bundles. Indeed the average daily value of the next bundle is anticipated to be equal to the current average daily value (Y/28).

The only variations allowed in the resulting monthly estimates are then the trading-day variations and length-of-month variations. For all socio-economic variables, this assumption is clearly unacceptable. Note that assumption ii) is contradicted by the bundles if they vary!

5.2 Modified Period Adjustment Procedure

In a variant of the procedure just described, referred as modified period adjustment procedure in this paper, the monthly estimates depend on one or two bundles. Consequently the modified procedure does not require overlapping between the bundles and the target months. The procedure is also described in terms of implicit interpolated daily values. Over one bundle, the implicit daily values, represented in Figure 3 by curve $z^1{}_t$, are the product of the (standardized) daily pattern and of the bundle daily average. In other words they are also obtained by equation (5.4) applied to each of the bundles. The monthly estimates are simply the monthly sums of the interpolated daily values, regardless of the bundle they pertain to. The interpolated values in the sum pertain to one bundle, if the target month is all included in the bundle; or to two bundles, if the target overlaps two bundles.

The assumptions of the modified period adjustment procedure are then:

i) The seasonal and trend-cycle components are constant, that is inexistent, within a bundle.

ii) These components change abruptly between one bundle and the next. This is illustrated in Figure 3: the daily values $z^1{}_t$ display a repetitive movement with constant level and amplitude within a bundle, and those level and amplitude change in the next bundle.

For socio-economic variables, the first assumption is unacceptable for the reasons already stated for the period adjustment procedure. The second assumption is also unacceptable. Indeed, the effect of climatic and of instutional seasons (e.g. Christmas) unfolds gradually from week to week and from day to day. Similarly business cycles develop gradually. (If these movements could be assumed to behave abruptly, it would still be questionable that the jumps occurs exactly between the bundles!) However, the assumption is somewhat off-set in a monthly estimate which incorporates daily values from different bundles. The seasonal and trend-cycle components in the monthly estimate is then an average of those components in the bundles, which smooth the abruptness described by assumption ii). The optimality of that smoothing remains very questionable. With the proposed calendarization method on the other hand, the interpolated daily values incorporate the effect of seasons and business cycles in a smooth and gradual manner, within a bundle and between bundles, as illustrated in Figure 3 by curve z_t.

The assumption of graduality is realistic for socio-economic variables. With very few exceptions, socio-economic agents require time to react to events: Suppliers and buyers are bound together by contracts specified over time; employers and employees are bound by collective agreements or at least some kind of loyalty; projects take time to complete; consumers cannot interupt their consumption of certain goods and services overnight; etc.

The two calendarization methods considered produce substantially different monthly estimates: The September, October and November estimates corresponding to (some of) the daily values displayed in Figure 3, are respectively 8093, 7725 and 9125 with the proposed method against 7525, 7625 and 9626 with the modified period adjustment procedure. The two sets of estimates also display different seasonal-trend-cyclical movements.

Both period adjustment procedures examined in this section yield results identical to the method proposed in this paper, only if the bundle daily averages are constant for the whole series. The latter method is superior to both period adjustment procedures considered in this section. That superiority lies in the untenable assumptions of the ad hoc procedures. In other words, it should not be necessary to carry out tests to prove that.

6. IMPLEMENTATION

This section discusses implementational issues of calendarization, namely the level of aggregation at which calendarization should be performed and computational considerations.

In principle, calendarization should be carried out at the highest level of aggregation for which the respondents (e.g. companies) have common reference periods for their weekly bundles and a common daily pattern. This situation could occur in regulated financial sectors. Calendarizing at the higher levels of aggregation should maximize the reliability of the estimates, and save on computational resources. In practice, the knowledge of the daily pattern may be less accurate at higher levels of aggregation.

The Appendix A shows the matrix operations required by calendarization according to the method proposed in this paper. The number of rows of matrices V and B′ depend on the total number of days considered. However, those matrices are known algebraically and do not have to be stored in computer memory. One may simply calculate their elements when needed. The size of matrix to be inverted (BVB′) depends on the number of successive weekly bundles of data considered. If a 20-year series contains 13 bundles per year, the dimension of the matrix would still be 260 by 260, and its inversion would be quite prohibitive.

However, the following is known from the benchmarking experience (Cholette, 1984) - and verifiable for calendarization. For bundles 3 to 258, applying the method on the 260 bundles produces practically identical daily values as applying the method on moving 5-bundle intervals. Very specifically, that moving 5-bundle application consists of the following steps. The application of the method to bundles 1 to 5 yields the daily estimates for bundles 1 to 3; the application to bundles 2 to 6 yields the daily estimates for bundle 4 (the middle bundle); the application to bundles 3 to 7; for bundle 5; and so forth; and the application to bundles 256 to 260 yields the daily estimates for bundles 258 to 260. That moving average-type implementation tremendously reduces the scale of calculations. Furthermore, the results become final, i.e. not subject to revision, after 2 bundles. These points argue in favour of implementing calendarization over moving 5-bundle series intervals. A case could be argued for a 7-bundle moving interval. Longer intervals are useless.

The calculations entailed by this moving average-type implementation are still considerable. In the same example, the number of rows of matrices V and B′ is 140 (5 times 28); and the size of the matrix to be inverted (BVB′) is 5. Although V and B do not have to be stored, the calculation of their elements requires computing time. The calculations can be further reduced in the following manner.

1) The application of the method to bundles 1 to 5 yields the daily estimates for bundles 1 to 3. (This step does not save on the calculations) The last daily estimate z_{84} (84=28*3) of bundle 3 is set aside for the second application of the method.

2) The second application is over the last day of bundle 3, whose estimate was set aside, and over bundles 4, 5 and 6. That last daily estimate of bundle 3 is specified as a 1-day bundle y_1 with reference

period $\tau_1 = \rho_1 - 1$. This forces the daily interpolations to start from the last daily estimate of bundle 3. This second application yields the estimates for bundle 4. The last daily estimate of bundle 4, z_{29} in the application interval considered (and z_{112} in the series), is set aside for the third application.

3) The third application is over the last day of bundle 4 and over bundles 5, 6 and 7 and over the last day of bundle 4. The last daily estimate of bundle 4 is specified as a 1-day bundle y_1 with reference period $\tau_1 = \rho_1 - 1$. This forces the daily interpolations to start from the last daily estimate of bundle 4. This third application yields the estimates for bundle 5.

4) The process is repeated for bundles 6, 7 and 8; and so forth; and finally, for bundles 258 to 260.

This moving average implementation with "links" reduces the number of rows of matrices V and B′ to 85 in the example chosen (except for step 1); and the size of BVB′, to 4. This implementation also better insures continuity of the interpolations between bundles and is the one we recommend. Our experimentations indicate that this implementation yields almost identical results as estimating the daily values over the whle series. For that reason, the length of the moving average can be set to 3 with very little penalty, which reduces the number of rows to 57. This implementation has been programmed in the SAS/IML language (Rheaume and Cholette, 1988). The coding is available from the authors.

The calendarization approach presented in this paper can also generate weekly values from the weekly bundles without estimating daily values. Time t then stands for weeks instead of days, and x_t is an appropriate indicator. (With a constant (trivial) indicator for instance, the resulting weekly interpolations are as smooth and constant as allowed by the constraints (3.2).) This suggests an even more economical variant of the latter implementation:

1) Interpolate weekly values from the weekly bundles (without interpolating daily values), in the manner just described.

2) Interpolate daily values from the interpolated weekly values (in the same manner).

3) Combine the daily values into the desired monthly values.

The number of rows of matrices V and B′ is now seven times smaller in step 1) (with the same example); and four times smaller in step 2). The resulting monthly values should be very close to those obtained by directly interpolating the daily values from weekly bundles. Further experimention is needed in that regard.

The last two implementations suggested - the last one especially - makes it feasible to perform calendarization on micro-computers; and, makes it conceivable to integrate the operation into (micro) Computer Assisted Telephone Interview systems.

7. DISCUSSION AND CONCLUSION

A method for the transformation of bundles of weekly data, referring to more than one week, into monthly estimates was presented. Appropriate calendarization - or the lack thereof - obviously impacts on the quality of time series produced by statistical agencies. Calendarization conditions all the other statistical processes applied thereafter: seasonal adjustment, integration into accounting frameworks (e.g. the National Accounts), econometric modelling, forecasting, etc. Despite that, we failed to encounter any published reference on the subject.

The "ad hoc" procedures examined in Section 5 are based on untenable assumptions. On the contrary, the method proposed is based on realistic assumptions, namely that seasonal and trend-cycle fluctuations develop in a gradual manner, which establishes its superiority. It would be relevant to carry out tests to determine the sensitivity of its monthly estimates to the pre-selected daily pattern. If the tests show little sensitivity, one may use a rather approximate daily pattern, or borrow the daily pattern from a related socio-economic variable for which it is known. Strong sensitivity could motivate the statistical agency to temporarily collect daily information, as this was done in Israel to estimate the effect of moving festivals (Morris and Pfeffermann, 1984, p. 256).

The approach to calendarization presented in this paper is an application of the proportional variant of the Denton (1971) benchmarking method generalized by Cholette (1987b). The relevant generalization is that the reference periods of the benchmarks can vary from occasion to occasion, i.e. from one bundle to the next. (Also the initial condition $z_0 = x_0$ of Denton is absent.) In a benchmarking context, the values of the indicator x_t are sub-annual measurements of the socio-economic variable of interest. That variable is also measured yearly by y_m, in a more reliable and independent manner. Benchmarking consists of adjusting x_t, in such a way that it sums to y_m over the reference periods of the y_m and in such a way that the movements of x_t are preserved. With the calendarization method proposed, the x_t is a predetermined daily pattern which is benchmarked to the weekly bundles considered as "benchmarks"; and the benchmarked values z_t correspond to the daily interpolations.

The benchmarking method used here is also a particular case of that of Hillmer and Trabelsi (1987), under the following conditions. i) No ARIMA model is specified (derived) for the benchmarked series. ii) A (proportional) random walk is specified for the "survey errors". iii) The stochastic variance of the benchmarks is very small compared to that of the unbenchmarked series. Furthermore, the theorems developped by Hillmer and Trabelsi become applicable. Namely the benchmarked series provides Minimum Mean Square Estimates of the interpolated daily values with computable confidence intervals.

As mentionned in Section 6, the approach presented in this paper can also be used to generate weekly values from the weekly bundles without estimating daily values. If one needs both daily estimates and the corresponding weekly sums as well as the monthly values, one should bear in mind that the calendarization operation is not transitive: Disaggregating the weekly bundles into weekly values, and then disaggregating the latter into daily values does not yield the same daily values (consequently the same monthly estimates) as disaggregating the bundles into daily values.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. D. Pfeffermann, of the Hebrew University of Jerusalem, for a valuable comments on the method presented in this paper.

Appendix A: Solution of the Calendarization Problem

In matrix algebra, objective function (3.1) is written:

$$F(Z) = Z'X^{-1\prime}D'D\,X^{-1}\,Z. \qquad \text{(A.1)}$$

Vector Z contains the T desired daily values z_t. Matrix D is the first difference operator:

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

Matrix X^{-1} is diagonal. Its diagonal elements are the inverse of the daily pattern values:

$$\underset{T\text{ by }T}{X^{-1}} = \begin{bmatrix} 1/x_1 & 0 & 0 & \dots & 0 \\ 0 & 1/x_2 & 0 & \dots & 0 \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 0 & 0 & . & & 1/x_T \end{bmatrix}. \qquad \text{(A.3)}$$

(For computational reasons, the diagonal may have to be standardized, by multiplying it by the average of the x_t's for instance. Only the relative values of this matrix matter.)

Constraints (3.2) write:

$$B\,Z = Y. \qquad \text{(A.4)}$$

Vector Y contains the weekly bundle values:

$$Y' = [\, y_1 \quad y_2 \quad \dots \quad y_M \,]. \qquad \text{(A.5)}$$

Matrix B is a sum operator. Its values are arranged in such a way that when B is multiplied by vector Z, the bundle sums of Z are obtained:

$$
\underset{M \text{ by } T}{B} = \begin{array}{c} \text{columns:} \\ \\ \\ \\ \\ \\ \end{array}
\begin{array}{cccccccccccccc}
 & & & \tau_1 & & & \rho_1 & & & \tau_2 & & & \rho_2 & \\
[\,0 & \ldots & 0 & 1 & 1 & \ldots & 1 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \ldots\,] \\
[\,0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 1 & 1 & \ldots & 1 & 0 \ldots\,] \\
[\,. & & . & . & . & & . & . & & . & . & & . & .\,] \\
[\,. & & . & . & . & & . & . & & . & . & & . & .\,] \\
[\,. & & . & . & . & & . & . & & . & . & & . & .\,] \\
[& & & & & & & & & & & & &]
\end{array}. \qquad (A.6)
$$

Parameters τ_m and ρ_m stand for the reference periods of each bundle (for instance τ_m=1,29,57,...; ρ_m=28,56,84,...). For stock series, $\tau_m = \rho_m$.

Objective function (A.1) is minimized subject to the constraints (A.4). This is accomplished by the Langrangian augmented objective function:

$$F(Z,\Lambda) = Z'X^{-1\prime}D'D\,X^{-1}\,Z - 2\Lambda'(B\,Z - Y), \qquad (A.7)$$

where Λ is a M by 1 vector containing the Lagrange multipliers. The values of Z which minimize this hyper-parabola are required. At the minimum, the derivative with respect to the unknowns Z and Λ are equal to zero. This leads to the normal equations:

$$
\begin{aligned}
dF/dZ &= 2\,X^{-1}\,D'D\,X^{-1}\,Z - 2\,B'\,\Lambda &= 0 \\
dF/d\Lambda &= \quad -\,2\,B\,Z \qquad + 2\,Y &= 0
\end{aligned} \; . \qquad (A.8)
$$

The solution to this system of equation is

$$
\underset{(T+M)\text{ by }1}{\begin{bmatrix} Z \\ \\ \Lambda \end{bmatrix}} = \underset{(T+M)\text{ by }(T+M)}{\begin{bmatrix} X^{-1\prime}D'D\,X^{-1} & -B' \\ & \\ B & 0 \end{bmatrix}^{-1}} \underset{(T+M)\text{ by }1}{\begin{bmatrix} 0 \\ \\ Y \end{bmatrix}} = \begin{bmatrix} W_0 & W \\ & \\ W^{\lambda}{}_0 & W^{\lambda}{}_y \end{bmatrix} \begin{bmatrix} 0 \\ \\ Y \end{bmatrix}. \qquad (A.9)
$$

For more details see Boot et al (1967) and Denton (1971). Performing the inversion in (A.9) by parts yields the solution for Z:

$$\underset{T\text{ by }1}{Z} = V\,B'\underset{M\text{ by }M}{(B\,V\,B')^{-1}}\,Y = \underset{T\text{ by }M}{W}\,Y, \qquad (A.10)$$

where matrix V is a very close approximation of the inverse of singular matrix $V^{-1} = X^{-1\prime}D_1{}'D_1\,X^{-1}$ of (A.9). The approximation of matrix V, inspired from Bournay and Laroque (1979), is known algebracically:

$$
\underset{T\text{ by }T}{V} = \begin{bmatrix}
1\ x_1x_1 & \alpha\ x_1x_2 & \alpha^2\ x_1x_3 & \ldots & \alpha^{T-1}x_1x_T \\
 & & & & \\
\alpha\ \ x_2x_1 & 1\ x_2x_2 & \alpha\ \ x_2x_3 & \ldots & \alpha^{T-2}x_2x_T \\
 & & & & \\
\alpha^2\ x_3x_1 & \alpha\ x_3x_2 & 1\ \ x_3x_3 & \ldots & \alpha^{T-3}x_3x_T \\
. & . & . & & . \\
. & . & . & & . \\
. & . & . & & .
\end{bmatrix}, \qquad (A.11)
$$

where α lower but as close to 1.0 as possible (e.g. 0.9999999). The inversion required in (A.10) is that of a M by M matrix instead of a T+M by T+M in (A.9) (M being the number of weekly bundles considered). The elements of V B' can easily be expressed algebraically. More details on this approximation can be found in Cholette (1988b).

The solution imply Z = W Y. In other words the interpolated daily values are weighted averages of the available weekly bundle data.

$$z_t = \sum_{m=1}^{M} w_{t,m}\, y_m. \qquad (A.12)$$

For the additive variant of calendarization, matrix V simplifies to

$$\underset{T \text{ by } T}{V} = \begin{bmatrix} 1 & \alpha & \alpha^2 & \ldots & \alpha^{T-1} \\ & & & & \\ \alpha & 1 & \alpha & \ldots & \alpha^{T-2} \\ & & & & \\ \alpha^2 & \alpha & 1 & \ldots & \alpha^{T-3} \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \end{bmatrix}, \qquad (A.13)$$

and the solution is

$$\underset{T \text{ by } 1}{Z} = X + V\ \underset{M \text{ by } M}{B'(B\,V\,B')^{-1}}\ (Y - B\,X) = X + \underset{T \text{ by } M}{W}\ (Y - B\,X), \qquad (A.14)$$

where Y - B X are the discrepancies between the bundle values and the corresponding sums of the pre-determined daily pattern X. The estimates may be expressed as a linear combination of the of the daily pattern X and of the bundle values Y:

$$Z = L_x X + L_y Y. \qquad (A.15)$$

where L_x=(I-WB) and L_y=W.

Appendix B: Possible solution for current estimates

Objective function (4.5) is minimized subject to the constraints (A.4). This is accomplished by the Langrangian augmented objective function:

$$F^{\lambda}(Z) = Z'X^{-1\prime}D'D\,X^{-1}\,Z + Z'B_s{}'S^{-1}\,D_s{}'D_s\,S^{-1}\,B_s\,Z - 2\Lambda'(B\,Z - Y), \qquad (B.1)$$

where the vectors and matrices Z, X^{-1}, D, Λ, B and Y retain the same definition as in Appendix A. The other matrices S^{-1}, B_s and D_s are defined as follows.

Matrix S^{-1} is diagonal. Its diagonal elements are the inverse of the known seasonal-trading-day pattern values s_k:

$$S^{-1} = \begin{bmatrix} 1/s_1 & 0 & 0 & \dots & 0 \\ 0 & 1/s_2 & 0 & \dots & 0 \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 0 & 0 & . & & 1/s_K \end{bmatrix}, \qquad \text{(B.2)}$$

K by K

(For computational reasons, the diagonal may have to be standardized, by multiplying it by the average of the s_k's for instance. Only the relative values of this matrix matter. However for calibration purposes, the same standardization must be made both to X^{-1} and S^{-1}.)

Matrix B_s is a monthly sum operator similar to B:

$$B_s = \begin{array}{c} \text{columns:} \\ \end{array} \begin{array}{c} \eta_1 \quad\quad\quad \theta_1 \quad\quad\quad \eta_2 \quad\quad\quad \theta_2 \end{array}$$

$$B_s = \begin{bmatrix} 0 \dots 0 & 1 & 1 & \dots & 1 & 0 \dots 0 & 0 & \dots & 0 & 0 & \dots \\ 0 \dots 0 & 0 & 0 & \dots & 0 & 0 \dots 1 & 1 & \dots & 1 & 0 & \dots \\ . & . & . & & . & . & . & & . & . & \\ . & . & . & & . & . & . & & . & . & \\ . & . & . & & . & . & . & & . & . & \\ & & & & & & & & & & \end{bmatrix}, \qquad \text{(B.3)}$$

K by T

where parameters η_m and θ_m stand for the reference periods of the K months considered. For stock series, $\eta_m = \theta_m$. Matrix D_s is the first difference operator of (A.2) except its dimension are K by K.

Following the same steps as in Appendix A, the following solution is obtained:

$$\begin{bmatrix} Z \\ \Lambda \end{bmatrix} = \begin{bmatrix} V^{-1} & -B' \\ B & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ Y \end{bmatrix} = \begin{bmatrix} W_0 & W \\ W^{\lambda}{}_0 & W^{\lambda}{}_y \end{bmatrix} \begin{bmatrix} 0 \\ Y \end{bmatrix}, \qquad \text{(B.4)}$$

T+M by T+M

where V^{-1} is equal to $X^{-1\prime}D'D\,X^{-1} + B_s'S^{-1}\,D_s'D_sS^{-1}\,B_s$.

The matrix to be inverted in (A.9) is rather large. Its size depends on the number of days encompassed by the bundles considered. And the solution (A.10) of Appendix A, where the size of the matrix depended only on the number of bundles (M), is not applicable. Indeed that solution was based on the approximation of V given $V^{-1} = X^{-1\prime}D'D\ X^{-1}$. If an approximation of V given $V^{-1} = X^{-1\prime}D'D\,X^{-1} + B_s'S^{-1}\,D_s'D_sS^{-1}\,B_s$ were found solution (A.10) would be applicable.

REFERENCES

- Alba, E. de (1988), "Temporal Disaggregation of Time Series: a Bayesian Analysis", **Journal of Business and Economic Statistics**, Vol. 6, No. 2, pp 197-206.

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

- 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

- 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. (170), **Time Series Analysis, Forecasting and control**, Holden-Day.

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

- Cholette, P.A. (1984), "Adjusting Sub-Annual Series to Yearly Benchmarks, **Survey Methodology**, Vol. 10, pp. 35-49.

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

- Cholette, P.A. (1987b), "Benchmarking and Interpolation of Time Series", **Statistics Canada**, Time Series Research and Analysis Division, Working Paper No. TSRA-87-014E.

- Cholette, P.A., Baldwin A. (1988), "Converting Fiscal Year Data into Calendar Year Values", **Statistics Canada**, Time Series Research and Analysis Division, Research Paper No. TSRA-88-012E.; **Journal of Business and Economic Statistics**, forthcoming.

- Cholette, P.A., (1988), "Converting Fiscal Quarter Data into Calendar Quarter Values", **Statistics Canada**, Time Series Research and Analysis Division, Research Paper forthcoming.

- Cholette, P.A. (1988), "Benchmarking Systems of Socio-Economic Time Series", **Statistics Canada**, Time Series Research and Analysis Division, Research Paper No. TSRA-88-017E.

- Cholette, P.A., Higginson, J. (1987), "A Note on Aggregating Data Referring to Different Financial Years", **Statistics Canada**, Time Series Research and Analaysis Division, Research Paper No. TSRA-87-015E.

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

- 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.

- Ehrenberg, A.S.C. (1982), "Writing Technical Papers", **The American Statistician**, Vol. 36, pp. 326-329.

- Friedman, M. (1962), "The Interpolation of Time Series by Related Series", **Journal of The American Statistical Association**, Vol. 57, pp. 729-757.

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

- Liu, L.M., Hudak, G.B. (1983), **The Sca System for Univariate-Multivariate Time Series and General Statistical Analysis**, Scientific Computing Associate, P.O. Box 625, DeKalb, Illinois 60115, U.S.A.

- Miller, M. (1986), "Monthly Wholesale-Retrail Trade Survey: Updates to the Data Capture - Preliminary Edit Specifications", **Statistics Canada**, Business Survey Method Division (internal document).

- Morris, N.D., Pfeffermann, D. (1984), "A Kalman Filter Approach to the Forecasting of Monthly Time Series Affected by Moving Festivals", **Journal of Time Series Analysis**, Vol. 5, No. 4, pp 255-268.

- Rheaume, R.A., Cholette, P.A. (1988) "Users Manual of programme CALENDR for the Calendarization of Socio-Economic Time Series Data, **Statistics Canada**, Time Series Research and Analysis Division, Ref. No. TSRA-88-020E

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