SELECTING LINKAGE TRANSITION INTERVALS TO MINIMIZE TREND-CYCLE AND SEASONAL DISTORTIONS

by

Guy Huot

Methodology Branch

Time Series Research and Analysis Division

Direction de la méthodologie

Division de la recherche et de l'analyse des chroniques

TSRA-88-013E

TIME SERIES RESEARCH & ANALYSIS DIVISION
METHODOLOGY BRANCH

SELECTING LINKAGE TRANSITION INTERVALS TO MINIMIZE TREND-CYCLE AND SEASONAL DISTORTIONS

by

Guy Huot

This is a preliminary version. Do not quote without author's permission. Comments are welcome.

May, 1988

RÉSUMÉ

Cette étude fait l'exposé et l'application d'une nouvelle procédure d'estimation de la période d'interpolation requise pour harmoniser les deux segments d'une même série chronologique lorsque ceux-ci divergent par leur niveau (et non par leur profil saisonnier). Cette procédure consiste à décomposer la série en tendance-cycle, saisonnalité et résidus. Chacune de ces trois composantes décrit une caractéristique distincte du phénomène socio-économique représenté par la série et requière par conséquent sa propre période d'interpolation. Puisqu'une seule période d'interpolation peut être retenue, la solution consiste alors à choisir celle qui minimise la distortion associée à l'estimation de chacune des composantes.

1. Introduction

A well-behaved socio-economic time series may sometimes show a sudden discontinuity in the behavior of its trend-cycle or seasonal components. Such a sudden change usually prompts questions about the structure of the series.

Some key questions are:

(a) Does the observed disruption reflect the true evolution of the phenomenon as described by the time series under study? If so, the discontinuity is part of the pattern of the series. An example is the impact of the severe 1981 recession on the level of the Canadian beneficiaries series which nearly doubled in one year.

(b) If the answer to question (a) is "no", should then the segment of the series available before the discontinuity be brought into agreement with the new segment? This problem of agreement often occurs when underlying concepts are redefined and the survey is redesigned. Then what linkage procedure should be used? What is the best length of transition interval?

This paper addresses the last question only. Thus, it discusses the selection of the best length of transition interval to phase in the discontinuity that occurred when the SEPH frame changed from the Labour Division Master File to the Business Register Master File. "SEPH" stands for the Statistics Canada Survey on Employment, Payroll and Hours. The series analysed is an employment series.

In order to answer the question, the series was decomposed into it's three basic components (trend-cycle, seasonality and irregularity). The analysis consists of finding out the best length of transition interval for each component, and selecting the length that would best suit all three components. The rationale for this approach is twofold: first, the

selection of a transition interval should not be done at the expense of any component of the series. Secondly, the decomposition of the series should be taken into account since both the original and the seasonally adjusted series are published and used for decision-making purposes.

The article is divided into four sections. The second section develops a method for identifying the best length of transition interval. The third section presents the empirical findings for each of the three components. Section four gives the main conclusions of the study.

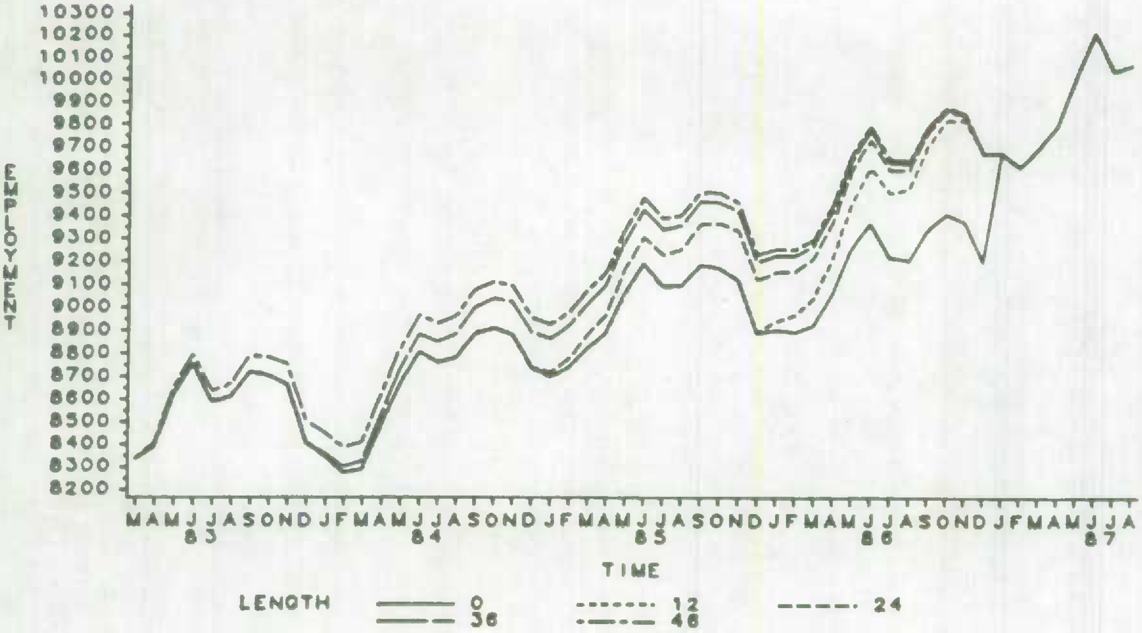
2. Methodology

The employment series analysed, starting March 1983 and ending August 1987, shows a discontinuity in January 1987. It is assumed that at the SEPH start-up in March 1983 no discrepancies existed between the Labour Division Master File and the Business Register Master File. Therefore, 46 months is the maximum length of interval that could be used to phase-in the discontinuity.

Intervals with lengths 6, 12, 18, 24, 30, 36, 42 and 46 months have been considered by Brown (March 1988) for the selection of the best interval. The discontinuity was phased-in linearly over these intervals. Figure 1 shows the original series (represented by the solid line) and the original series revised using intervals with lengths 12, 24, 36 and 46 months.

Brown also computed the coefficient of transition interval (CTI) and the relative variance (RV) statistics for each length of transition intervals. The two statistics suggested a length of 24 months. However, as Brown pointed out, the application of these two statistics to time series may not be appropriate, especially when the series is non-stationary. Without strong evidence in favor of any length, a conservative approach was suggested with a length of 46 months (the second best choice being 24 months).

FIG. 1: EMPLOYMENT SERIES REVISED USING TRANSITION INTERVALS OF LENGTH 0 TO 46 MONTHS (IN THOUSANDS)



The focus of our attention is an approach that:

(1) can be used with time series;

(2) provides valid criteria to compare the adequacy and performance of different lengths of transition interval; and

(3) allows for the selection of the best length of transition interval.

The approach consists of decomposing the series into its three basic components and selecting the length that minimizes trend-cycle and seasonal distortions. The X-11-ARIMA program (Dagum, 1980) will be used to decompose the series.

It is worth emphasizing that the phasing-in and decomposing procedures may require different lengths of series. While a series with an overall length of 56 months of data is enough to phase-in the discontinuity, a minimum of 60 is desirable for the decomposition. Otherwise the components estimated using fewer than 5 years of data may be subjected to revisions when using 5 years or more of data. These revisions occur especially when there is moving seasonality in the series.

3. Empirical Results

The results are based upon the multiplicative X-11-ARIMA decomposition model. Intervals with lengths 0 (original series), 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 42 and 46 months were used. The accuracy of each decomposition was evaluated using the "total quality" indicator built into the program (Lothian and Morry, 1978). A total quality value higher than 1.0 is an indication of unacceptable quality. The value for length 0 was .24. It was .18 for length 3 and .10 for all other lengths. Thus none of the lengths of interval was rejected on the grounds of a poor quality decomposition. The analysis was then carried on at the components level.

3.1 The seasonal component

The evaluation of the lengths of transition interval for the seasonal component is made from:

(1) an F-test for the presence of identifiable seasonality in the original series (F_o);
(2) an F-test for the presence of stable seasonality in the final unmodified seasonal and irregular ratios (F_s);
(3) an F-test for the presence of moving seasonality in the final unmodified seasonal and irregular ratios (F_m);
(4) the relative contribution of the seasonal component to the total variance in the series (RC_s); and

(5) the standard deviation of the month-to-month changes in the trend-cycle and irregular series (SD_{mm}).

The results for these 5 statistics are summarized in table 1.

Table 1: Tests and measures to assess the best length of interval for the seasonal component

Length of interval	F_o	F_s	F_m	RC_s	SD_{mm}
		Series ending in December 1986 (before the discontinuity)			
n.a.	35.6	219.4	.1	89.4	.4
		Series ending in August 1987			
0 month	9.2	51.7	.2	75.4	.8
3 months	8.7	86.2	1.3	78.5	.6
6 months	18.1	223.4	.3	82.3	.5
9 months	29.8	263.7	.0	85.1	.5
12 months	27.9	240.6	.1	84.3	.5
15 months	31.3	261.2	.1	84.2	.4
18 months	36.9	248.4	.1	84.5	.4
21 months	43.4	295.7	.2	85.1	.4
24 months	38.6	271.5	.1	84.9	.4
30 months	35.6	264.6	.0	84.5	.4
36 months	37.2	282.5	.1	85.3	.4
42 months	39.2	277.2	.1	85.4	.4
46 months	39.1	282.9	.0	85.7	.4

The series ending in December 1986 has an F_o value of 35.6. It suggests an easily identifiable seasonal pattern in the original series. However, the F_o value drops to 9.2 when the series is extended to August 1987 without phasing in the discontinuity. As a result of the discontinuity, the structure of the original series has changed thus making the identification of the seasonal pattern more difficult. An 18-month length of transition interval is required for the F_o value to return to 36.

The F_s-test (carried out on the series after removing the trend-cycle component) requires only a 6-month length of transition interval to return to a value of 219.4. Such a short length suggests firstly, a very stable

seasonal pattern and secondly, a discontinuity that had an impact mainly on the trend-cycle component. The 51.7 and 86.2 values indicate that the discontinuity, phased-in with a 0 and a 3-month length, is partly interpreted as residual variation in the F_s-test. It is only when a sufficiently long phase-in is done that the discontinuity is smoothed and identified as part of the trend-cycle and seasonal components. The seasonal pattern is then estimated without distortion.

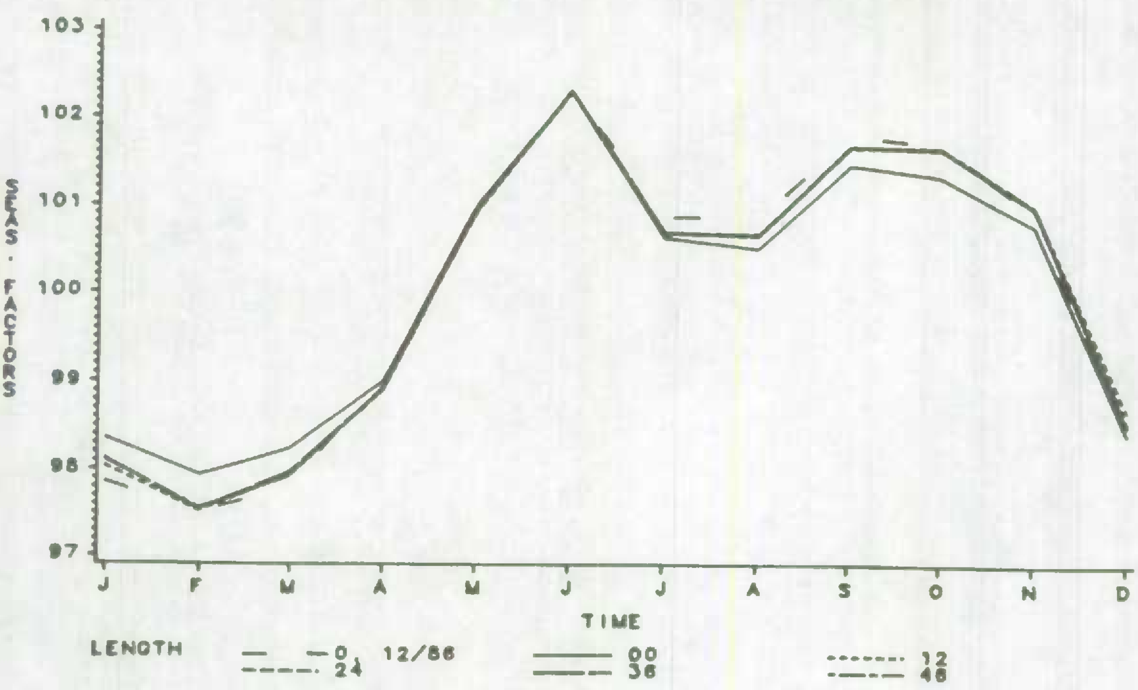
Discontinuities in the seasonal pattern may also exist, as well as in the trend-cycle component. The linear phase-in technique is then inappropriate since it is likely to preserve the seasonal pattern present in the series before the discontinuity. Either another technique should be used (see for instance Cholette, 1984), or the components will be estimated with some distortion.

The F_m-test supports the absence of moving seasonality characterized by gradual changes in the seasonal amplitude. Thus, the components estimated using fewer than 5 years of data are not likely to be subjected to revisions when using more data. Moving seasonality may still be present since the second segment of the series is too short to allow for a powerful test for change in the seasonal amplitude. In general, an adequate test of the seasonal pattern requires a minimum of 12 months of data after the discontinuity.

The three F-tests, used to find out if significant seasonal patterns (stable and moving) are present in the series, rejected none of the lengths of interval. These tests are, however, very useful in identifying the shortest length of interval that will minimize the seasonal distortion. From the point of view of F_o, an 18-month length is required while 9 months are sufficient for F_s. The F_m-test confirms the absence of changes in the seasonal pattern. Thus, as far as change in the seasonal pattern is concerned, the lengths are all equally acceptable.

The RC_s values are measures of the relative contribution of the seasonal component to the total variance of the series. RC_s is not a statistical test. Nevertheless, the decrease in the RC_s values from 89.4 to 75.4 suggests that the discontinuity with no phasing-in has distorted the seasonal pattern making its identification and estimation more difficult. This problem of estimation shows in figure 2 as an overestimation of the seasonal factors in January, February and March and an underestimation in August, September, October and November (length 00). Another finding is that the seasonal factors corresponding to phasing-in lengths of 9 months and more, coincide as shown in figure 2 for lengths 12, 24, 36 and 46 months. Figure 2 also shows that these coincident seasonal factors depart slightly in January, July, August and September from those of the series ending in December 1986 (length 0, 12/86). Thus, although according to

FIG. 2: ESTIMATED SEASONAL FACTORS OF THE EMPLOYMENT SERIES REVISED USING TRANSITION INTERVALS OF LENGTHO TO 46 MONTHS



the F-tests, the series is characterized by a stable seasonal pattern and no identifiable moving seasonality, the decomposition model allowed a slight change in the estimation of the seasonal patterns.

SD_{mm} is a measure of smoothness of the series adjusted for seasonality. It gives an indication of how well the seasonal component is identified and removed from the original series. The smallest SD_{mm} value of .4 is obtained with a phase-in of 15 months and over.

Overall, the results support a transition interval with length of 15 to 18 months for the seasonal component.

3.2 <u>The irregular component</u>

Outliers may be perfectly valid observations. Alternatively, the presence of certain outliers in the irregular component may suggest that the discontinuity was not phased-in properly.

Table 2 displays the outliers corresponding to each decomposition. They were identified by testing the irregular values against a 2.5σ limit estimated from the data (Dagum, 1980). Table 2 reveals a significant association between the position and the number of outliers and the length of the transition intervals. This finding confirms that the discontinuity has to be phased-in using an adequate transition interval, otherwise, it cannot be properly explained by the decomposition model.

From the point of view of the irregular component, the best minimum length ranges between 12 to 15 months. The corresponding outliers, reduced to one or two, are independent of the discontinuity. In fact, they are in the same position as those of the series ending in December 1986 (before the discontinuity).

Table 2: Outliers identified in the original and revised irregular component

Length of interval	months										
	Series ending in December 1986 (before the discontinuity)										
n.a.		Mar84			Mar85						
	Series ending in August 1987										
0 month				Jan85					Nov86	Dec86	Jan87
3 months							Aug86	Sep86		Dec86	
6 months	Apr83	Mar84			Mar85						
9 months			Apr84		Mar85						
12 months		Mar84				Dec85					
15 months		Mar84									
18 months		Mar84			Mar85						
21 months		Mar84									
24 months		Mar84									
30 months		Mar84									
36 months		Mar84									
42 months		Mar84									
46 months		Mar84			Mar85						

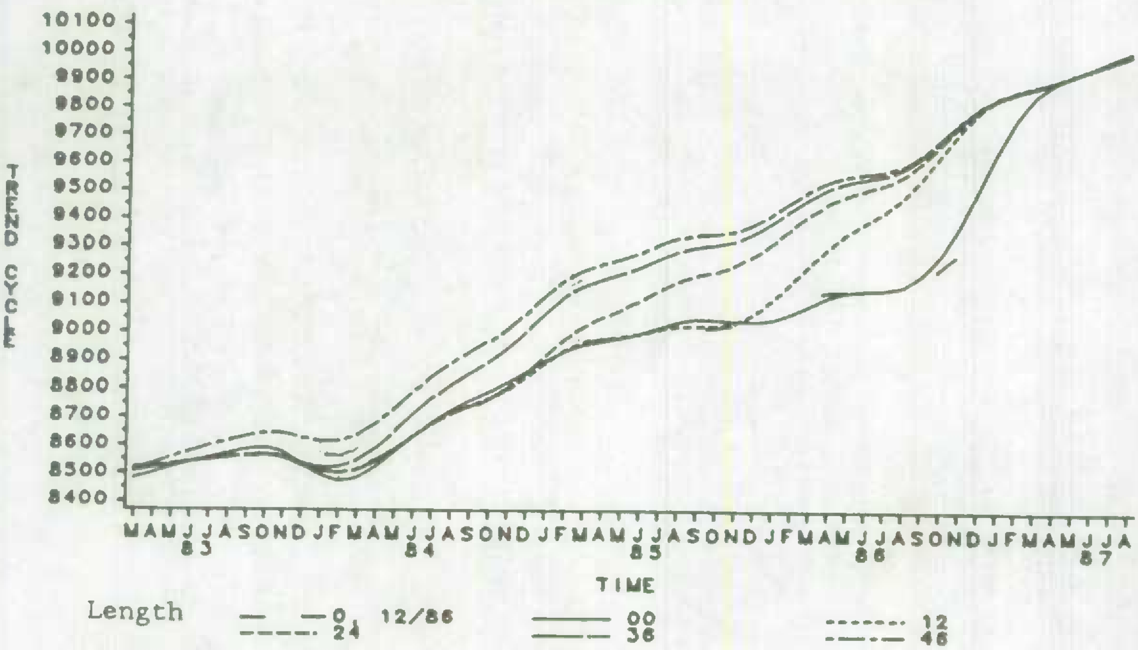
3.3 The trend-cycle component

While the seasonal component is characterized by a yearly movement, the trend-cycle movement is much longer. Thus the two components have different socio-economic meanings and they will likely be affected in different ways by the discontinuity. As a result, the phasing-in may require a different length of transition interval for each component.

Figure 3 shows the estimated trend-cycle component of the series ending in December 1986 (length 0, 12/86), and then ending in August 1987 when the discontinuity is phased-in with lengths 0, 12, 24, 36 and 46 months. There is not much difference between the trend-cycle estimates of the series ending in December 1986, and ending in August 1987 with no phase-in. The only exceptions are November 1986 and December 1986 which were previously

identified as outliers. The explanation for the presence of these outliers is that the discontinuity has introduced an abrupt turning point in the trend-cycle component of the series extended to August 1987.

FIG. 3: ESTIMATED TREND-CYCLES OF THE EMPLOYMENT SERIES REVISED USING TRANSITION INTERVALS OF LENGTH 0 TO 46 MONTHS (IN THOUSANDS)



Of course, this turning point is not valid from a socio-economic point of view since it shows a false upswing in the trend-cycle component of the employment series. Consequently, the selected length of transition interval for the trend-cycle component must:

(1) prevent the occurence of false turning points; and
(2) preserve the movements present in the original trend-cycle component (series ending in December 1986).

A visual inspection of the data shows that the 12-month length should be rejected because it fails the first criterion, and the 24-month length because it fails the second. Hence the best minimum length for the trend-cycle component is 42 monthts. This length is in no way selected at the expense of the other components. Besides visual inspection, a test, based on the degree of parallelism that exists between the original and revised trend-cycle component, could be used to help select the best minimum length of transition interval.

4. Conclusion

The analysis reveals that the selection of the best length of transition interval strongly depends on the patterns followed by the seasonal, irregular and trend-cycle components before the discontinuity, together with the new structure introduced by the redesigned survey. Moreover, these three components are affected in different ways by the discontinuity. As a result, the phasing-in requires a different length of transition interval for each component. Thus, the best length is the one that adequately suits all three components. Accordingly, a 42-month transition interval is recommended for this employment series.

References

Brown, A. (March 10, 1988). "Choice of Transition Interval". Memorandum, Business Survey Methods Divison, Statistics Canada (Unpublished Document).

Cholette, P.A. (1984). Linking Heterogeneous Segments of Seasonal Time Series. American Statistical Association, Proceedings of the Business and Economic Section, 178-183.

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

Lothian, J. and M. Morry (1978). "A Set of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Program". Research Paper, Time Series Research and Analysis Division, Statistics Canada (H-024E, Supplementary Document).