stantial contribution to an important area of time series analysis that has received, and will continue to receive, considerable attention from researchers.

References

Becker, R.A. and Chambers, J.M. (1984). S – An Interactive Environment for Data Analysis and Graphics. Belmont: Wadsworth.

Box, G.E.P., Hillmer, S.C., and Tiao, G.C. (1978). Analysis and Modelling of Seasonal Time Series. In Seasonal Analysis and Economic Time Series, ed. A. Zellner, Washington D.C.: US Department of Commerce, Bureau of the Census, 309–334.

Box, G.E.P., Pierce, D.A., and Newbold, P. (1987). Estimating Current Trend and Growth Rate in Seasonal Time Series. Journal of the American Statistical Association, 82, 276–282.

Cleveland, W.S., Devlin, S.J., and Terpenning, I.J. (1981). The SABL Statistical and Graphical Methods. Murray Hill, NJ: Bell Laboratories.

Cleveland, W.S. and Terpenning, I.J. (1982). Graphical Methods for Seasonal Adjustment. Journal of the American Statistical Association, 77, 52–61.

Cleveland, W.S. and Devlin, S.J. (1988). Locally-weighted Regression: An Approach to Regression Analysis by Local Fitting. Journal of the American Statistical Association, 83, 596–610.

Harvey, A.C. and Todd, P.H.J. (1983). Forecasting Economic Time Series With Structural and Box-Jenkins Models: A Case Study. Journal of Business and Economic Statistics, 1, 299–315.

Harvey, A.C. (1984). A Unified View of Statistical Forecasting Procedures. Journal of Forecasting, 3, 245–275.

Harvey, A.C. and Durbin, J. (1986). The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling (with discussion). Journal of the Royal Statistical Society, Ser. A, 149, 187–227.

Harrison, P.J. and Stevens, C.F. (1976). Bayesian Forecasting. Journal of the Royal Statistical Society, Ser. B, 38, 205–247.

Hillmer, S.C. and Tiao, G.C. (1982). An ARIMA-model-based Approach to Seasonal Adjustment. Journal of the American Statistical Association, 77, 63–70.

Comment

Britt Wallgren and Anders Wallgren[1]

The authors have developed a new method for dealing with an important problem in applied statistics. All statistical agencies report time series and to make these reports intelligible to the users, data must be decomposed so the important components can be described numerically and graphically.

We have not only read and enjoyed their

[1] Statistical Research Unit, Statistics Sweden, S-701 89 Örebro, Sweden.

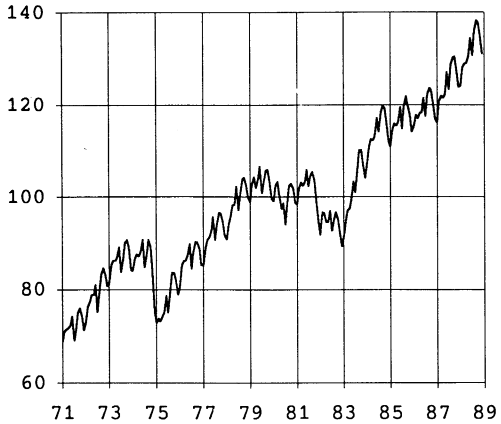

Fig. 1. United States, Original Data.

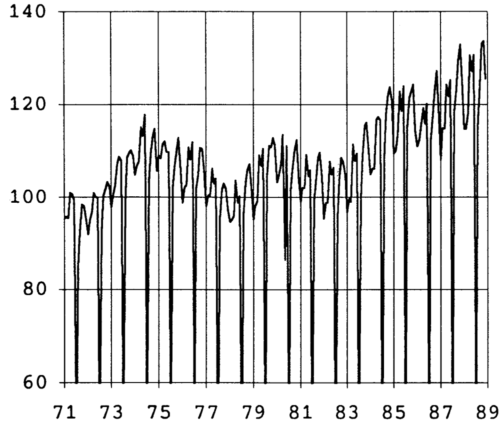

Fig. 3. Sweden, Original Data.

well-structured article but have also tested the STL program (after modifying input/output and compiling). The authors have indeed attained their goal of easy computer implementation and fast computation. Where STL requires seconds, the PC version of X-11-ARIMA requires minutes. However, when we compared the statistical properties of STL and X-11-ARIMA we found that with STL's present limitations, it cannot compete with X-11-ARIMA. If STL is modified (the authors mention two modifications) we believe that the statistical properties of STL will be much improved, and STL will then become an important tool for many applications.

In Section 3.6 the authors state that their primary goal is to estimate the seasonal component and perform seasonal adjustment. The estimate of the trend (trend-cycle) is of secondary interest – it is only a means in estimating the seasonal component. We believe that the statisticans behind X-11 and X-11-ARIMA share this opinion. Dagum (1978) stresses the importance of seasonal adjustment and the X-11 output also gives

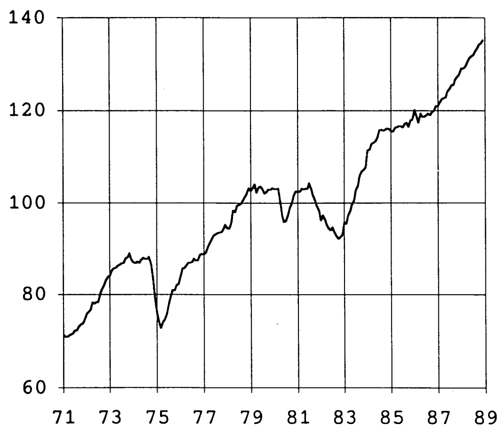

Fig. 2. United States, Seasonally Adjusted Data.

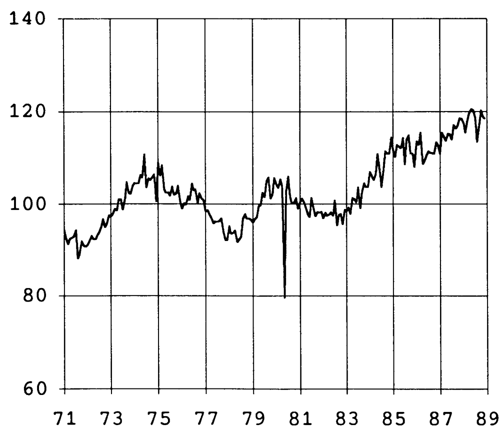

Fig. 4. Sweden, Seasonally Adjusted Data.

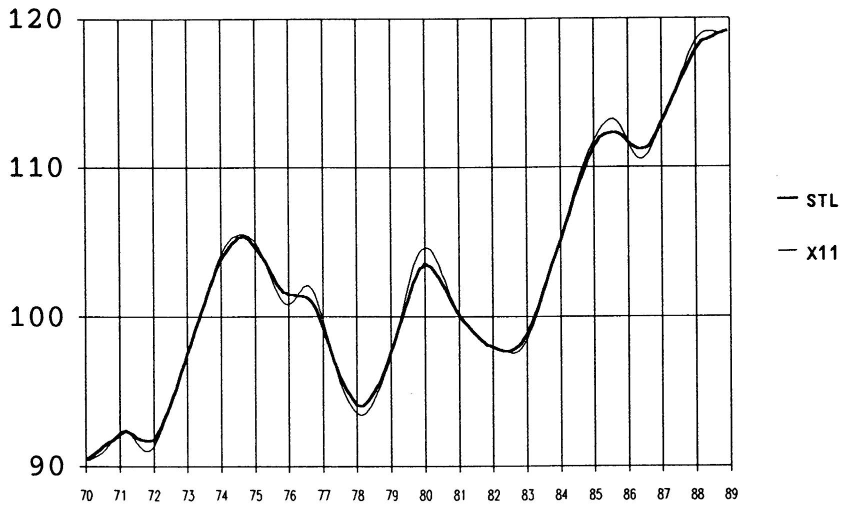

Fig. 5. Trend-cycle Estimates.

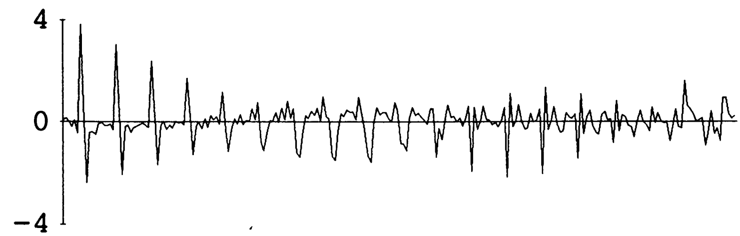

Fig. 6. Seasonal Component, S(STL)-S(X-11).

the impression that seasonal adjustment is the primary goal of the analysis. We do not share this opinion because we generally work with time series whose seasonally adjusted values are very difficult to interpret. In Figures 1–4 monthly data describing industrial production (manufacturing, quantity index, 1980 = 100) in the United States and Sweden are compared.

Both series have been seasonally adjusted with STL. In the Swedish data, the July values fall under the border of Figure 3.

The seasonally adjusted series in Figure 2 clearly shows all important aspects of the progress of U.S. industrial production. For this series, the seasonal adjustment is sufficient. However, the Swedish counterpart in Figure 4 does not give a clear description – it is difficult to identify turning points in the business cycle and it is extremely difficult to determine the actual stage of the cycle. Comparing Figures 1 and 4 we find that seasonally adjusted Swedish data and unadjusted U.S. data are almost equally difficult to interpret. Our conclusion is that for aggregated series from large nations whose random component is small, seasonal adjustment may be the primary goal. However, for disaggregated series from large nations (for instance U.S. unemployed males ages 16–19 used by Cleveland et al. in their Figures 5 and 11), series from small nations and series from business corporations we need decomposition and smoothing methods which smooth both the seasonal and random components.

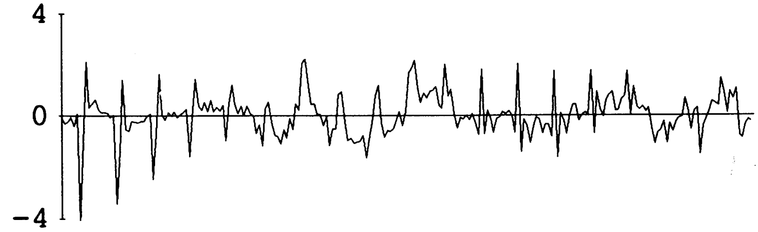

Fig. 7. Remainder, R(STL)-R(X-11).

When comparing different methods of time series decomposition it is important to know whether a method can solve only easy problems or if it also can solve difficult problems which are important to many users.

Easy to decompose
* Series with almost linear trend
* The historical part of the time series
* Series with small irregular fluctuations

Difficult to decompose
* Series with (business) cycles
* The current part of the time series
* Series with large irregular fluctuations

We have compared how STL and X-11-ARIMA decompose the Swedish monthly industrial production (Figure 3) and how these methods solve the difficult problems just mentioned. In Figures 5–7, the estimates of the three components are compared. The decomposition with STL used $n_s = 31$, $n_t = 19$, $n_l = 13$ and robust iterations. With X-11-ARIMA, we used an additive model, automatic ARIMA modelling, and replacement of extremes with the ARIMA model. The trend-cycle was estimated with a 23 term Henderson filter and adjusted for strikes (May 1980). We have tried to follow the recommendations given by Cleveland et al. when choosing the STL parameters and have chosen the X-11-ARIMA parameters to get a corresponding description.

In Figure 5 we see that STL and X-11-ARIMA differ in the decomposition into T and $(S + R)$. Figures 6 and 7 show that the

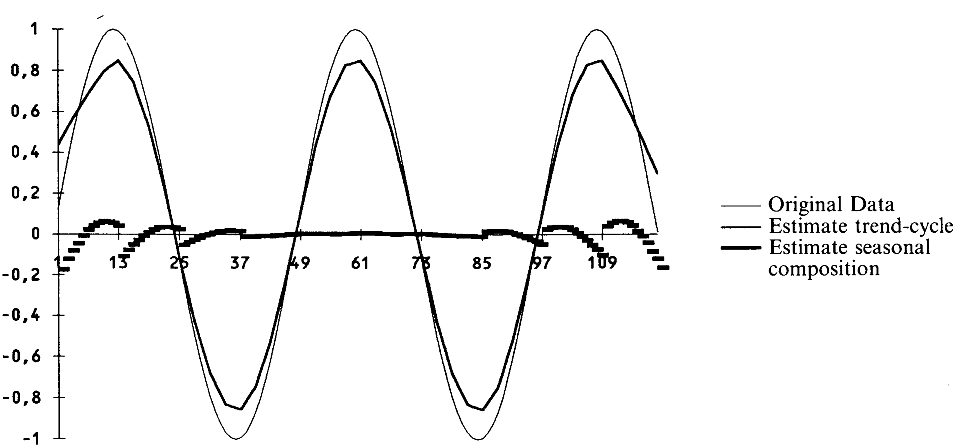

Fig. 8. STL's Decomposition of a Sine Wave.

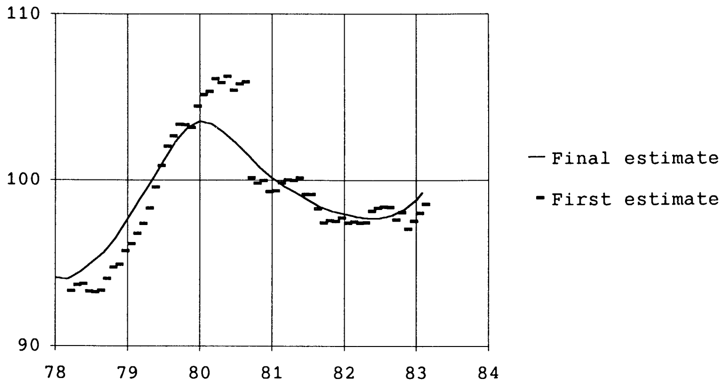

Fig. 9. Trend-cycle Estimates by STL.

two methods also differ in the decomposition into S and R.

The trend-cycle estimate of STL cannot closely follow peaks and troughs (this is also shown in Figure 8). The low-pass filter in Step 3 and the trend smoothing in Step 6 perhaps need to be modified, otherwise business cycles will partly go into ($S + R$) and make the estimation of S difficult. In Section 2.1 Cleveland et al. mention that $d = 2$ is a better choice if the series has substantial curvature. However, the present version of the STL program only allows $d = 1$ and as we understand, the present rules for choosing the parameters also assume that $d = 1$. To be able to smooth series with business cycles, we need an STL program that allows $d = 2$ in Steps 3 and 6 and rules for determining the other parameters. We also suspect that the $n_p \cdot n_p \cdot 3$ moving average in Step 3 smoothes too much in this case.

At the beginning of Section 5.2 Cleveland et al. state that they ignored end effects when they developed rules for choosing the STL parameters. We suspect that this simplification makes STL less useful in practice. It is generally of vital interest to decompose the current part of the series, making end effects important. To test the ability of STL to decompose near the end points we used sine waves of various frequencies. In Figure 8 the STL decomposition of a sine

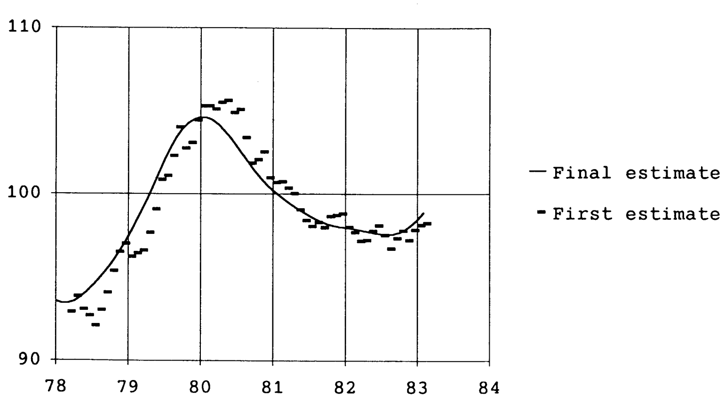

Fig. 10. Trend-cycle Estimates by X-11.

Table 1

	Mean absolute revision		Maximum absolute revision	
	STL	X-11-ARIMA	STL	X-11-ARIMA
First estimate of T	1.1	0.8	4.0	2.4
Second estimate of T	0.9	0.5	3.4	1.7
Third estimate of T	0.7	0.3	2.8	1.1
First estimate of S	0.9	0.5	3.4	1.6
Second estimate of S	0.9	0.5	2.8	1.9
Third estimate of S	1.0	0.6	3.0	2.8

wave is shown with $n_p = 12$, $n_l = 13$, $n_t = 23$, and $n_s = 7$. The series represents ten years of monthly data. We can see that the seasonal component is well estimated in the middle but poorly estimated near the end points. The trend is poorly estimated at turning points and also near the end points. However, X-11 could describe these sine waves without problems – the trend-cycle estimate follows closely peaks and troughs for series with cycles lasting two or more years, no seasonal component is invented, and no bad end effects can be seen.

We have investigated the end point problems with the series in Figure 3. Using data up to March 1978 we get the first preliminary estimates of T and S for this month. When we add one observation we get the first estimates for April 1978 and revised estimates for March 1978. Proceeding in this way up to February 1983 we get 60 first estimates of T and S, 60 estimates which have been revised after one month, 60 estimates revised after two months, etc. These estimates can be compared with the "final" estimates based upon all 228 observations from the period 1970–1988. In Figures 9 and 10 the trend-cycle estimates by STL and X-11-ARIMA are shown.

With STL it is very difficult to detect the turning point at the beginning of 1980. How much the preliminary estimates are revised is shown in Table 1 (index units).

We conclude that X-11-ARIMA is the better tool if you want to decompose the current part of this kind of series. We hope that the STL method will be improved so the first estimates will be more reliable. In Section 6.4 Cleveland et al. mention that STL could be combined with an ARIMA model. We suggest that this combined method be tested and we hope that this will improve the first estimates. STL will then be of great practical use to those who regularly report time series data. The users of statistical reports are always irritated by revisions and any method that can reduce the need for revisions is welcome.

Both STL and X-11-ARIMA have options for automatic detection and treatment of outliers. X-11-ARIMA has two options: (1) observations with extreme ARIMA residuals are replaced with the ARIMA function values and (2) observations are given weights depending on the remainder in the decomposition and the analyst can choose the rules for how outliers are defined. In STL observations are given robustness weights depending on the remainder. The only choice here is between decomposition with or without robust iterations.

In the series describing Swedish industrial

Table 2

	STL	X-11-ARIMA
Observations replaced by ARIMA	–	2
Observations given zero weight	2	15
Observations given full weight	1	192
Observations given partial weight	225	19
Mean weight	0.84	0.89

Table 3

Year	Month	STL		X-11-ARIMA	
		Robustness weights		Weights	
		Final	With data up to August 80	Final	With data up to August 80
79	8	0.42551	0.99701	1.0000	0.1324
79	9	0.38541	0.96859	1.0000	1
79	10	0.82547	0.74193	0.2586	1
79	11	0.91891	0.91105	1.0000	1
79	12	0.68368	0.65445	1.0000	1
80	1	0.96264	0.90703	1.0000	1
80	2	0.99995	0.99403	1.0000	1
80	3	0.75721	0.71478	1.0000	1
80	4	0.96563	0.97118	1.0000	1
80	5	0.00000	0.00000	(replaced by ARIMA-value)	
80	6	0.99936	0.72043	1.0000	1
80	7	0.26013	0.99688	0.9674	1
80	8	0.99982	0.00000	1.0000	1

production there is one known outlier. In May 1980 the series was disturbed by lock-outs and strikes. The outlier options in STL and X-11-ARIMA gave the results in Table 2.

This means that with STL 16% of the data are discarded. We wonder if STL is too sensitive to outliers and discards too much data. This must lead to unreliable estimates of the components. In Figure 11 the weights given by STL and X-11-ARIMA are compared. Black areas indicate reduced weights. Our impression is that "outliers" with reduced weight appear near turning points in the business cycle both with STL and X-11-ARIMA.

If observations near turning points get reduced weight these important turning points will be detected too late. In Table 3 we compare the final weights based on all 228 observations with the weights based on 128 observations up to August 1980.

In STL, the preliminary robustness weight for August 1980 is zero, and the turning point at the beginning of 1980 is not yet detected. This is shown in Figure 9. We can see that both methods change the weights and that the weights and changes are different for each method. We wonder if automatic outlier options should be used at all (in STL and X-11-ARIMA). We prefer a

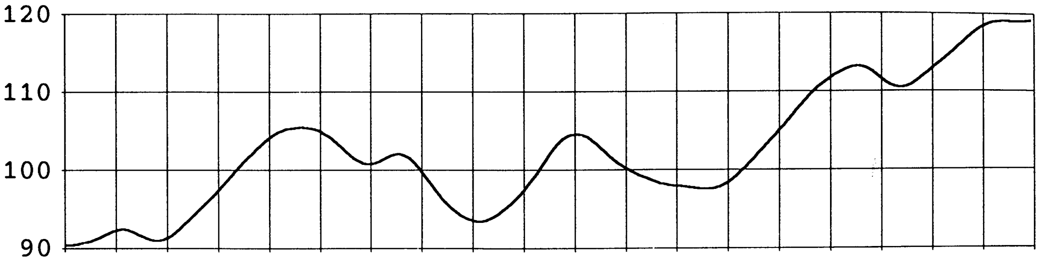

Fig. 11a. Trend-cycle by X-11.

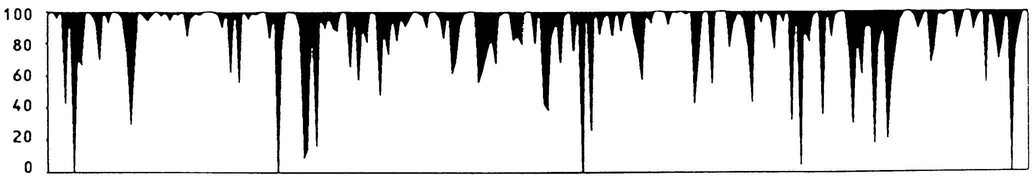

Fig. 11b. Robustness Weights by STL.

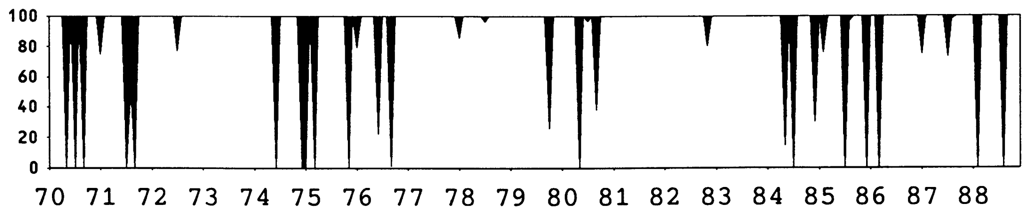

Fig. 11c. Robustness Weights by X-11-ARIMA.

strategy in which the analyst decides which observations should be replaced by ARIMA values. The automatically generated weights may have unknown and undesirable effects on the estimated components.

The present version of STL is too limited. Those who report time series with business cycles or large irregular fluctuations and who want to estimate the current state need more reliable methods. We hope that STL will be developed to eliminate these shortcomings.

Reference

Dagum, E.B. (1978). A Comparison and Assessment of Seasonal Adjustment Methods for Employment and Unemployment Statistics. Background Paper No. 5, National Commission on Employment and Unemployment Statistics, Washington, D.C.