Comment

Stuart Scott[1]

1. Introduction

The authors are to be congratulated for taking a good approach to a difficult problem. Since the 1960's, the standard for seasonal time series decomposition has been the X-11 method developed by Julius Shiskin and his colleagues. Modern time series modeling has been successfully applied in numerous situations, but has not yet supplanted X-11 and its updates. One of the rare attempts to develop alternative nonparametric techniques has been SABL, also developed by W.S. Cleveland and colleagues at Bell Labs. Thus, the Cleveland et al. paper reflects experience from previous investigations into seasonal decomposition as well as development of loess for nonseasonal smoothing over a number of years. I agree with the paper that (1) many features of X-11 have proved to be successful, but may be improved upon and (2) some of the (parametric) modeling techniques may be less flexible than STL in specifying a seasonal component.

I have applied STL to two problems of current interest to me at the U.S. Bureau of Labor Statistics: the use of intervention analysis in seasonal adjustment and trend-cycle estimation. Using recommended filter lengths, STL gives smoother trend and seasonal components than X-11 with typical options. As a quick indication, Figure 1 shows trend components from STL and X-11 ARIMA for Unemployed Men, 16–19 (UM), the second series in Cleveland et al. (see their Figure 5). STL's trend is much simpler and avoids numerous false or short-lived turning points. On the other hand, an analyst may prefer the X-11 ARIMA trend for more closely following the short-term movement in the series.

[1] Bureau of Labor Statistics, 441 G Street, N.W., Washington, D.C. 20212, U.S.A.

2. Intervention Analysis

This section applies STL to intervention analysis of four energy series from the Producer Price Index and two Current Employment Statistics series at the Bureau of Labor Statistics (BLS), studied previously by other methods. Buszuwski and Scott (1988) applied Box and Tiao (1975) methodology with ARIMA modeling to estimate outliers and intervention effects. By assigning these effects to the trend-cycle and irregular components, seasonal adjustment of these series was improved in terms of better diagnostic statistics and, in most cases, better sets of seasonal factors. Criteria for comparing sets of seasonal factors include similarity with factors derived prior to a major intervention period and smoothness of seasonal pattern through the twelve-month cycles. The question to be examined here is whether STL can perform better than X-11 on series with erratic behavior and compete with the ARIMA modeling approach.

Analysis done so far gives mixed results for STL. The accompanying table contains diagnostic statistics from X-11 ARIMA, the stable F, M7, and Q, for all six series with no

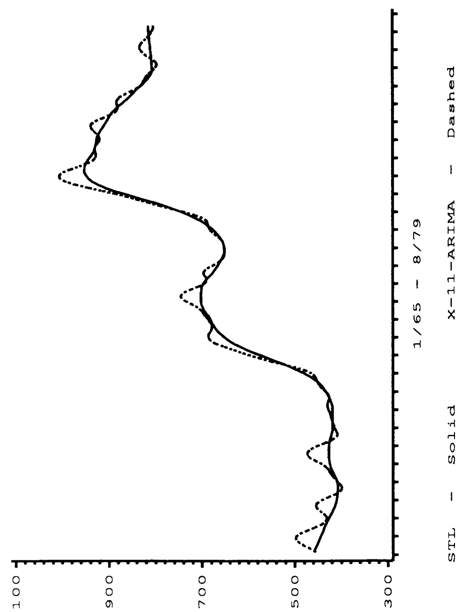


FIG. 1. TREND SERIES, X-11-ARIMA AND STL, UM

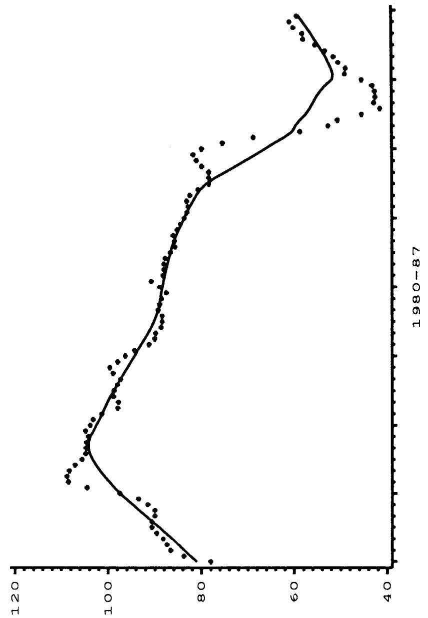


FIG. 2. OBSERVED & TREND SERIES, KJF

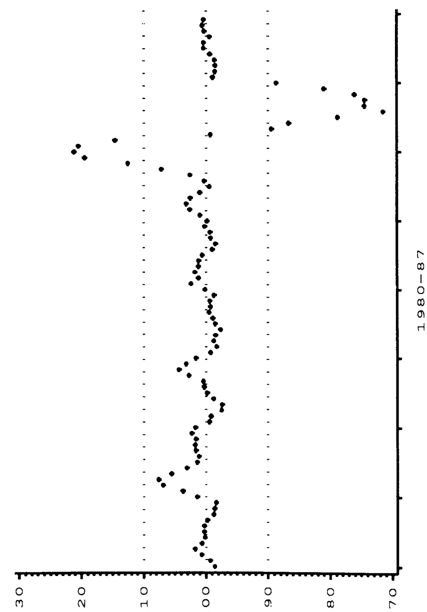


FIG. 3. IRREGULAR COMPONENT, KJF

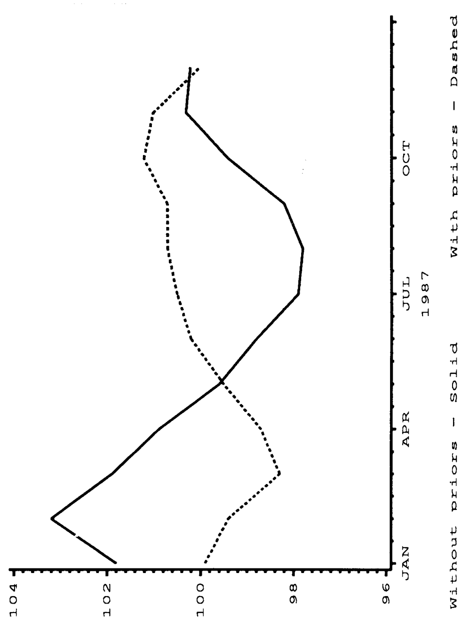


FIG. 4. SEASONAL FACTORS FOR KJF

prior adjustment (NP), with prior adjustment based on ARIMA modeling (AP), and with one or two prior adjustments based on STL (SP and SPR). SP statistics are worse (lower on F, higher on M7 and Q) for four of the six series, but better for the remaining two, Gasoline and Lumber & Wood Products. Where STL performs worse, it is much worse. Again, these results are based on using STL only for prior adjustment for interventions. X-11 and X-11 ARIMA are used for the actual seasonal adjustments.

Let me examine in further detail the results for Kerosene and Jet Fuels (KJF). Any method is likely to have a difficult time estimating the seasonal, since an erratic trend dominates. Figure 2 shows the observed series and STL's trend component. The sharp price drop beginning January 1986 is a candidate for treatment as an intervention. At that time, the oil-producing nations were unable to establish prices firmly and Saudi Arabia drastically increased its production. Figure 3 graphs the irregular component. There was a strong correspondence between low values of the robustness weight and irregulars falling outside the interval (90, 110). The first method adopted for prior adjustment was to use these extreme irregulars for prior adjustment. This is a fairly pure test of STL's identification of extremes versus X-11's. Figure 4 shows the seasonal factors for 1987, with and without prior adjustment. The seasonal patterns differ greatly. As mentioned earlier, X-11 with no prior adjustment yields better diagnostics and better agreement with factors based on 1980–85 data only.

What goes wrong? Referring again to Figure 2, major departures from the trend curve occur systematically in early 1981 and, very strongly, from late 1985 through 1986; in other words, much short-term trend is missed. Last year's paper with Buszuwski discusses in more detail how an intervention can distort the seasonal component. This led to a second attempt at prior adjustment which included adjustment for ramp-shaped interventions, as in Buszuwski and Scott. The filter length choices (n_s, n_t) were changed from (9, 23) to (50, 9). (The large value for n_s corresponds to use of the stable seasonal option in X-11 for this series.) Ramps corresponding to steep sections of the trend curve were estimated. After adjustment for these ramps, the series was decomposed again using STL and large irregulars identified. Prior adjustment combined these two steps, and yielded the dashed series in Figure 5. Figure 6 shows that the seasonal pattern now comes close to matching that with no prior adjustment, and the diagnostics are also much improved.

More basically, STL assumes that the underlying trend is smooth while an intervention may involve an abrupt change or near-discontinuity. McDonald and Owen (1986) tackle this with a locally-weighted regression technique similar to loess, but allowing one-sided filters for fitting at each point, as well as central filters. This also leads to filters of varying sizes. Their technique successfully tracks a sawtooth function, with a discontinuity, with added Gaussian noise. They call their technique split linear fitting and describe it as smoothing with edge detection.

3. Trend Estimation for the 1981–82 Recession

The second application examines trend estimates for three labor force series derived from the U.S. Current Population Survey (the household survey), conducted by the Bureau of the Census for BLS. Trend series derived from STL are compared with trend estimates employing the 13-point Henderson filter. This filter is the default trend filter in the X-11 methodology and the one recently adopted by the Australian Bureau of Statistics for its labor force press release (Trewin 1987).

FIG. 5. OBSERVED & PRIOR ADJUSTED SERIES, KJF

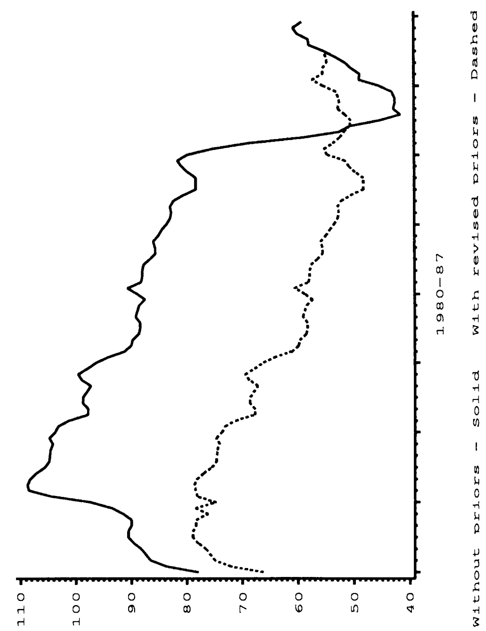


FIG. 6. SEASONAL FACTORS FOR KJF

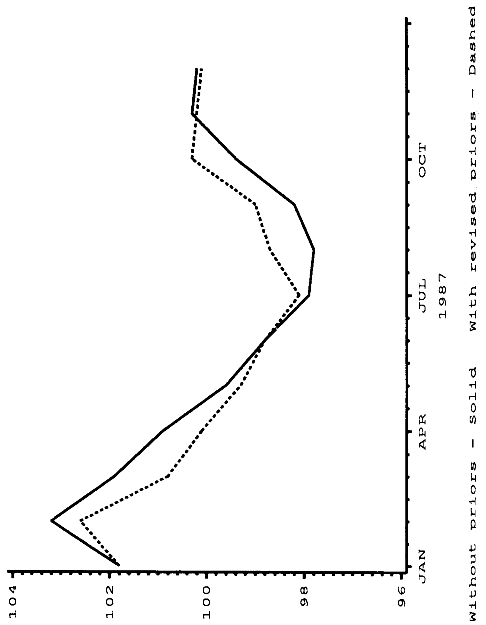


FIG. 7. HENDERSON TREND SERIES, UM

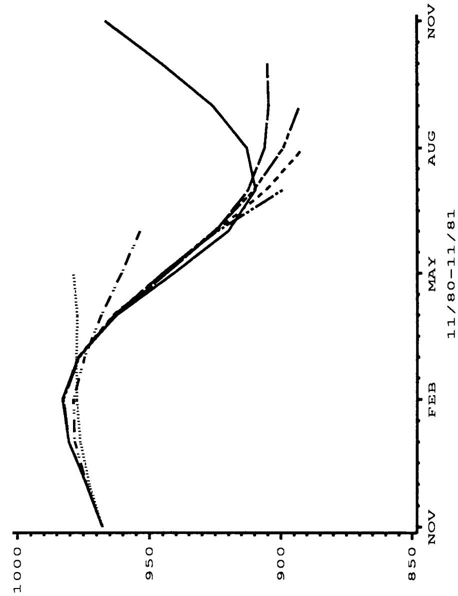


FIG. 8. STL TREND SERIES, UM

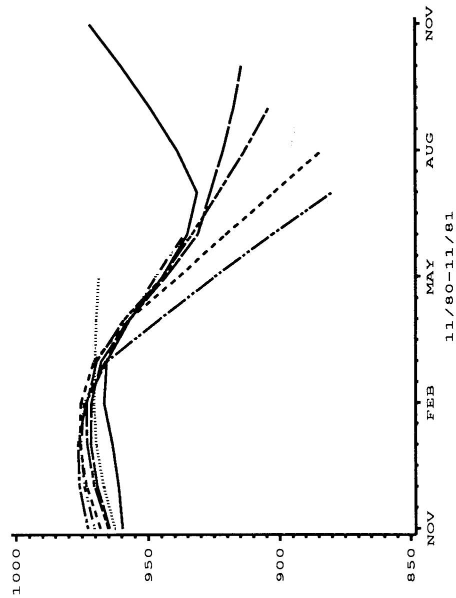
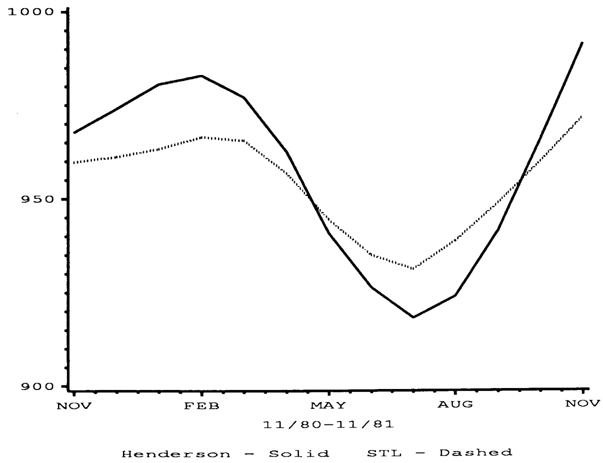


FIG. 9. TREND COMPARISON, UM



Let us consider the traditional time series decomposition, $O = T + S + I$, where O denotes the observed series, T the trend or trend-cycle, S the seasonal, and I an irregular component. For analysts and policymakers interested in short-term economic behavior, the trend-cycle is the component of principal interest, with the other components nuisance terms. There is, however, considerable variation among users in the choice of series for analysis. Many analysts prefer the unadjusted series as the purest numbers from the survey. The civilian unemployment rate, *seasonally adjusted*, is the featured statistic in BLS's monthly press release, *The Employment Situation.* Currently, no trend series are published by BLS. Presumably, this is due to the fact that reasonable general-purpose trend filters cannot be expected to estimate contemporaneously a turn in the series. Furthermore, an individual's notions of the appropriate amount of smoothing for trend may change according to the series or intended purpose. In this sense, the trend-cycle component is probably even more difficult to define than the seasonal.

Turning points for the July 1981 through November 1982 recession, one of seven post-World War II recessions, were examined for Unemployed Men, 16–19 (UM), depicted in Figure 1, Employed Women, 20+, and Nonagricultural Employment, Part-Time for Economic Reasons. For all three series, STL trend curves consistently evolve more gradually and less decisively than X-11 ARIMA's. Figure 7 illustrates the construction of trend series for UM, and forms the basis for comparing methods. The upper graph shows the evolving trend series across a turning point, the onset of the recession. The curves represent Henderson current trend values as of seven ending months, May through November, 1981. Delineation of the July turning point evolves smoothly and gradually in later months. November establishes a definite change in direction and pushes the turning point back to July. In Figure 8, STL responds more slowly. Even in October, the curve has not fully leveled off. Abruptly in November, a turning point in July and August emerges. Figure 9 superimposes the two trend curves through November. Both are satisfyingly smooth, but the Henderson curve estimates a deeper trough.

For comparison, simple loess was applied to the seasonally adjusted series, in the same manner as the Henderson filter. (In this case, only values back to 1/79 were input). With loess, smoothness is set by specifying the fraction of points f. With $f = 0.6$, the trend estimate was similar to STL; with $f = 0.4$, similar to Henderson. With $f = 0.3$, the loess trend identified the July turning point one month earlier than Henderson, and estimated an even deeper trough. The price for this early identification was less smoothness, both in the evolution of the trend and in the "final" trend.

4. Conclusions

The application of locally-weighted regression in STL is appealing for combining flexibility and conceptual simplicity in its filter determination. The flexibility was realized in both applications in these comments, especially the trend estimation. The recommendations leading to very smooth trends seem reasonable for many settings. The use of trend curves to discourage overreaction to series fluctuations is perhaps more appropriate than looking to them for early detection of turning points.

Checking for outliers based on the median absolute deviation of irregulars, rather than a moving standard deviation, makes sense on statistical grounds, and, especially in the intervention analysis, appeared to work better than X-11.

Some of the theoretical arguments for choice of filters seem to break down in the case of interventions. McDonald and Owen's split linear fit for nonseasonal series incorporates a capability for handling near-discontinuities in series such as the 1986 price fall in Kerosene and Jet Fuels.

My comments evaluate STL in limited areas, largely within the X-11 context. I will be interested in more direct comparisons with X-11 for seasonal adjustment, and in additional efforts to refine STL or develop additional analytic tools to accompany it.

5. References

Box, G.E.P. and Tiao, G.C. (1975). Intervention Analysis with Applications to Economic and Environmental Problems. Journal of the American Statistical Association, 70, 70–79.

Buszuwski, J.A. and Scott, S. (1988). On the Use of Intervention Analysis in Seasonal Adjustment. Proceedings of the Business and Economic Statistics Section, American Statistical Association, 337–342.

McDonald, J.A. and Owen, A.B. (1986). Smoothing with Split Linear Fits. Technometrics, 28, 195–208.

Trewin, D. (1987). Estimation of Trends and Time Series Models from Continuing Surveys. Bulletin of the International Statistical Institute 52, Book 2, 443–462.