References

Akaike, H. (1981). Seasonal Adjustment by a Bayesian Modeling. Journal of Time Series Analysis, 1, 1–13.

Box, G.E.P. (1980). Sampling and Bayes' Inference in Scientific Modelling and Robustness. Journal of the Royal Statistical Society, Ser. A, 143, 383–430.

Ozaki, T. and Thomson, J.P. (1987). Seasonal Adjusting of Overseas Balance of Payments Series. Mathematical Statistics Division Working Paper 1987/1, Victoria University of Wellington, New Zealand.

Rejoinder

Robert B. Cleveland, William S. Cleveland, Jean E. McRae and Irma Terpenning

1. Additions to STL

We are very impressed with the energy with which the discussants probed STL and compared it to other procedures. Based on these discussions and our own thoughts we now have a list of features – some of them "bells and whistles" as Trewin aptly puts it – that we would like to see added to STL. The list will be given at the end of our rejoinder. We hope that researchers in the seasonal adjustment community will involve themselves in these additions, just as a number of the discussants already have done.

Nothing in the discussions leads us to believe that the fundamental ideas of STL are unsound. These fundamentals are

- basing all smoothing (with the exception of three equal-weight moving averages) on loess
- basing the robust estimation on iterated weighted least squares
- using the backfitting algorithm, as X-11 does, to iterate between seasonal and trend smoothings
- providing for seasonal cycles of any length period, for missing values, and for a wide and nearly continuous range of amounts of smoothing of the trend and seasonal components
- keeping the system as simple as possible
- designing the software in a modular, easy-to-alter fashion.

Comments of discussants that indicated displeasure with performance characteristics can be addressed either by using STL in a different way or by using some of the additions on our list.

2. Open and Closed Systems

Changing STL or adding to it is straightforward because it is an *open system*: one that can be readily understood and altered if necessary. When STL does not perform as desired in some application, it often is possible to diagnose the problem and find a solution. Furthermore, software implementations can have a simple, modular structure, so modifying or adding to STL software is straightforward.

By comparison, X-11 is a *closed system.* It is so complicated that it is exceedingly difficult to understand its properties and determine solutions to problems. In fact, we conjecture that very few who use X-11 understand fully the details of how it works; we base this on our own experiences and interactions with others in the field. Furthermore, even if one had a solution to a problem of X-11 that affected its internal workings, implementing it in the current code would be exceedingly difficult.

The problem with a closed system, even if it works reasonably well, is that it stifles progress. The only movement forward can occur by treating it as a black box and embedding it in other systems as a unit. The great progress forward of X-11 ARIMA is an example. X-11 has forced an intellectual poverty on us in which our only option is to compare it with other ideas through pure empiricism. While empiricism is an important part of any field, having little ability to reason theoretically makes improvement difficult.

Unfortunately, some of the discussants seem to have viewed STL as a closed system, perhaps conditioned by extensive exposure to X-11. Instead of viewing STL as a set of basic ideas, with details to be viewed flexibly and changed as needed, there was a tendency to view STL as a set of immutable details incapable of allowing alteration or experimentation. The next section documents the chief example.

3. There is No Such Thing as THE Trend Estimate of STL

Several discussants (Gray and Thomson, Wallgren and Wallgren, and Ozaki) viewed the trend component produced by STL with the default value of $n_{(t)}$ as something to be used as a descriptor of trend-cycle. Such a procedure cannot work. First, the purpose of the default value is simply a technical choice to keep the trend and seasonal from competing for the same variation in the series. Getting a desirable trend-cycle component requires a trend post-smoothing: subtract the STL seasonal component and then use the basic loess smoother on the residuals to get a desirable description of trend-cycle. A default parameter for trend smoothing cannot provide a general solution because just what variation should go into the trend-cycle depends on the application and on the goal of the analysis. For example, Findley, Monsell, Shulman, and Pugh (to appear) write: ". . . the concept of trend is use-dependent rather than fixed. For example, analysts seeking long-term trends expect less fluctuating trends than investigators of short-term trends." And Scott states: "an individual's notions of the appropriate amount of smoothing for trend may change according to the series or intended purpose." This is why some discussants found the STL default trend too smooth and others found it too noisy. For example, if the analyst does not want to smooth away the peaks and troughs of low-frequency cycles, which was the case for several discussants, locally quadratic smoothing should be used in this trend post-smoothing; this was noticed by Wallgren and Wallgren.

We are partially at fault here. Our description of this issue was buried in the second paragraph of Section 3.6 of the paper. Also, we plotted and discussed the default trend, not making entirely clear that this was for diagnostic purposes and was not a study of trend-cycle. We also made the mistake of not illustrating the trend post-smoothing; in fact we had such an illustration in

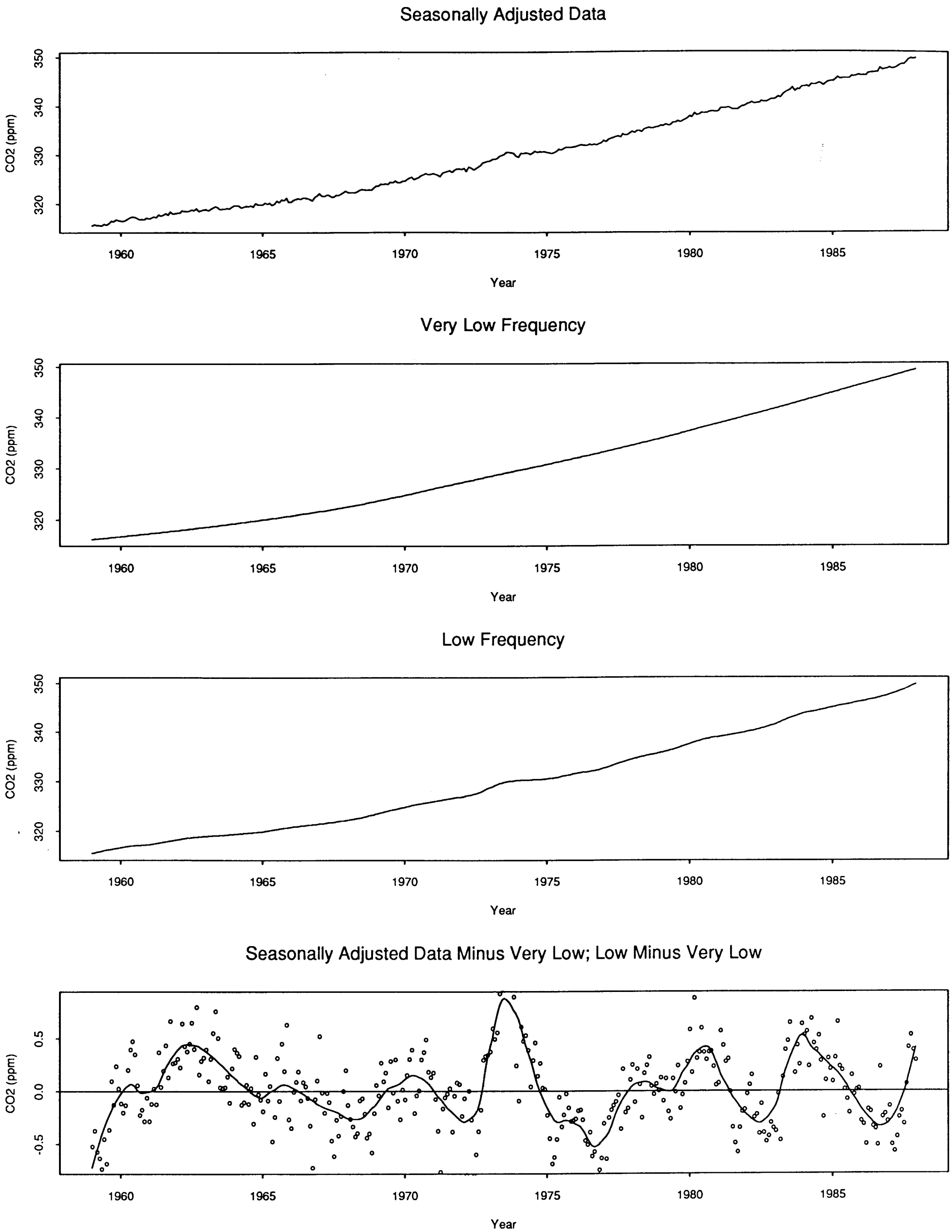

Fig. 1. Trend Post-Smoothing of Monthly Carbon Dioxide.

a very early draft and unfortunately deleted it. Here is a modified version of that discussion.

The top panel of Figure 1 shows the seasonally adjusted values of the monthly CO_2 series that we analyzed in the paper. The second panel from the top is a trend post-smoothing using loess with no robustness iterations, locally-linear fitting, and $q = 200$. The resulting trend is a very low-frequency component that describes the rise in the level of CO_2 that is being caused by the burning of fossil fuels (Kukla and Gavin 1981). The third panel of Figure 1 shows another trend post-smoothing, this time with locally-quadratic fitting and $q = 35$. This low-frequency component captures the rise in level plus cycles with periods of several years in length. To see the properties of these cycles, the low frequency component minus the very low frequency component is graphed by the solid curve in the bottom panel. These cycles are correlated with the Southern Oscillation, a measurement of the difference in atmospheric pressure between Easter Island in the South Pacific and Darwin, Australia (Bacastow 1976). The circles in the bottom panel are the data in the top panel minus the very low-frequency trend. Notice that the curve is able to climb the mountains and descend the valleys and does not distort the pattern in the data. This results from the use of locally quadratic fitting; if we use locally linear fitting, distortion occurs unless q is small, in which case the curve is too noisy.

Thus there are at least two interesting trend-cycle components for these data. If we want a description of the variation in CO_2 caused by fossil fuel burning, then an answer is the curve in the second panel; if we want a description of variation associated with the Southern Oscillation, then the curve in the bottom panel is an answer.

4. Adding Locally Constant Fitting, Locally Quadratic Fitting, and a Hybrid

Loess smooths by fitting polynomials of degree d locally. In our Fortran implementation of STL we used $d = 1$, which is locally linear fitting, for both the trend and seasonal smoothers. For the climatology data sets in which we are interested, this works quite well. But for other data sets, including some of those analyzed by the discussants, other degrees likely would perform better. For the trend smoothing it would be helpful to add $d = 2$, locally quadratic fitting; this would enable the trend smoother to be a better facilitator for the seasonal smoother in applications where there are low-frequency cycles that account for a large amount of the variation in the data. For the seasonal smoother it would be helpful to add both locally constant fitting and a hybrid smoother consisting of the average of a locally constant fit and a locally linear fit; this would widen the range of possibilities for the amount of smoothing, as we will show in Section 6.

5. The Value of Theoretical Reasoning

One of the values of having at least some ability to argue theoretically in studying seasonal adjustment methods is that some issues can be decided incisively and others, while they may need some empirical work, can at least be guided by theory. When empirical results have no theoretical guidance, they produce a morass (Kuhn 1962, p. 16). The grasp that we have on the mathematical properties of STL, which is not particularly extensive, is nevertheless enough to make headway on issues raised by several discussants, without turning from our desks to the computer terminal.

5.1. Initial trend value (Trewin)

We have an argument for why, if the default values of $n_{(l)}$ and $n_{(t)}$ are used, the nontrivial eigenvalues of the trend and seasonal operators have distinct sets of eigenvectors. This means that even though we start with a trend component identically equal to 0, we get convergence on the first step. This is not quite a proof for the actual STL because, as Gray and Thomson have pointed out, we have made a circularity assumption in the analysis of Section 5 of the paper. Still, we would expect this idealized setting to provide good guidance for the actual case. And when we turned to our computer terminals for verification, we found that convergence was always very rapid, and never found a case where more than two steps were needed. The empirical results, which rest on the theoretical foundation, lead us to believe that it is unnecessary to use any other set of starting values.

5.2. Component series (Trewin)

If the robustness feature of STL is not used – that is, $n_{(o)}$ is zero – and the values of the smoothing parameters are kept the same for the component series and the total, then the adjusted total will equal the sum of the adjusted components. Suppose, however, that robust estimation is deemed necessary. The following procedure would also produce equality: get the final robustness weights from the adjustment of the total, and then for each component series do Steps 1 to 6 a total of $n_{(i)}$ times using these weights.

5.3. Competition between trading-day and seasonal components (Trewin)

The standard way to estimate a trading-day component is to regress on independent variables that are based on the number of times each day of the week occurs in a month (Young 1965). If we add the trading-day estimation to STL as described in Section 6.3 of the paper, the projection matrix of the regression is the trading-day operator. Thus we can study competition between the trading-day and seasonal components using the theory of Buja, Hastie, and Tibshirani (1989).

5.4. Smoother design (Gray and Thomson)

The basic principle of loess is easy to grasp – to get a smoothed value at x, fit a polynomial locally and evaluate it at x. It should be appreciated that while the details of loess are new, this basic idea goes back decades (Macaulay 1931). The basic principle of loess is also general in that it allows us to fit at any value of x. This is important because it gives fitted values at the middle of the series, at the ends, beyond the ends, or anywhere else. We have not had to come up with separate procedures for trend smoothing, seasonal smoothing, special end-value procedures, another special procedure for predicting one-step beyond, and yet another procedure for filling in missing values. Having a basic principle allows us to gain certain insights into the properties of the smoother for all of the smoothing tasks simultaneously. For example, here are two immediate statements we can make about bias: (1) Local fitting of a polynomial of degree d reproduces a polynomial of degree d, a statement so obvious as to seem trivial, but in fact quite essential in considering bias properties. (2) Away from the ends, locally linear fitting to a quadratic effect results in a constant (additive) bias.

Contrast this with X-11. Since there is no basic principle of underlying smoother design, each smoothing task – trend, seasonal, end-value, and one-step prediction – has its own ad hoc solution. The consequence is considerable difficulty in reasoning about the properties of the X-11 smoothers. And no one has yet devised a solution to the problem of treating missing values with X-11, in part because a sensible ad hoc solution to this problem without an underlying principle is difficult to conceive. In other words, the ad hoc nature of X-11 contributes to its closure as a system.

5.5. *Why not use component models*

Parametric component models provide an excellent mechanism for theoretical reasoning. Should we not then give up all filtering approaches to seasonal adjustment such as X-11 and STL and use component models? We believe the answer is "no." The reason is that for general purposes, component-model approaches do not yield sufficient flexibility in specifying seasonal components. Furthermore, far too many series are not well fitted by the parametric models that are now commonly used. Using a nonparametric approach – that is, decomposing by a filtering procedure and studying the stochastic properties of the series by spectrum estimates based on the Fourier transform – is both more flexible and easier to carry out successfully in practice.

6. Empirical Studies: Matching Smoothers

One difficulty that plagues purely empirical studies of different seasonal adjustment procedures is that results are dependent on the amounts of smoothing of the trend and seasonal smoothers. Choosing the smoothing parameters of two procedures to match the amounts of smoothing is very tricky at best and impossible in some cases.

6.1. *Standard deviations*

Suppose we have a smoother

$$x_v = \sum_u d_{uv} w_u,$$

using the notation of Section 5.1 of the paper. Then one measure for the amount of smoothing at time position v is

$$\sigma_v = \sqrt{\sum_u d_{vu}^2}.$$

This is simply the standard deviation of x_v when w_v is white noise with variance 1. The smaller σ_v is, the more the amount of smoothing for position v. For two smoothers to match everywhere in terms of the amount of smoothing, the two values of σ_v for each v would have to be equal. This will often be too much to ask.

Table 1. 100 times standard deviations of three smoothers

Position	Henderson 13-term	Loess: $d = 2, q = 17$	Loess: $d = 1, q = 11$
1	62	71	64
2	52	52	52
3	45	41	42
4	43	36	36
5	43	36	35
6	45	38	38
7	45	39	38
8	45	40	38
9	45	42	38

Table 2. 100 times standard deviations of two smoothers

Position	X-11: 3 × 9	Loess: hybrid, $q = 13$
1	44	44
2	41	40
3	38	36
4	35	33
5	33	32
6	32	32
7	32	34

6.2. Matching the STL and X-11 smoothers by standard deviation

A number of discussants appeared to want trend and seasonal components of STL whose amounts of smoothing are closer to those to which they are accustomed with X-11. With the help of Brian Monsell, we studied the standard deviations of the smoothers of STL and X-11. The amount of smoothing of the Henderson k-term smoother – for $k = 5, 9, 13$, and 23 – is best matched, in our judgement, by the amount of smoothing of loess with locally quadratic fitting and $q = k + 4$. The amount for the Henderson is also reasonably well matched by the amount for loess with locally linear fitting and $q = k - 2$, but only for $k = 9$, 13, and 23. As an example, Table 1 shows the standard deviations of the three smoothers for $k = 13$. The values in the first column are the positions: the beginning position is 1, the next is 2, and so forth. The $3 \times k$ X-11 seasonal smoother is best matched by a hybrid loess smoother: the mean of the locally linear fit with $q = k + 4$ and of the locally constant fit with $q = k + 4$. As an example, Table 2 shows the standard deviations of the two smoothers for $k = 9$. Matching by standard deviation does not, of course, mean that the X-11 and STL smoothers are the same. In particular they still will have different bias properties.

6.3. Sliding-spans analysis

Sliding-spans analysis (Findley et al. 1990) is clearly a useful diagnostic tool. However, comparing two seasonal adjustment procedures by this method for a particular series is a delicate matter. The ideal procedure is the following: for each method find decompositions that are deemed adequate for representing the seasonal variation; choose in each case the one with the smallest sliding-spans measure; and compare the result for the first seasonal adjustment procedure with the result for the second. The next best procedure is simply to match the amounts of smoothing of the two seasonal smoothers and of the two trend smoothers. Note that it is not enough to match just the seasonal smoothers. The sliding-spans diagnostics depend on the amounts of smoothing of both the trend and seasonal smoothing; for example, Table 3 in the discussion of Findley and Monsell shows that the diagnostics change with the value of $n_{(t)}$. The reason is that the seasonal operator matrix, $\tilde{S}$, depends on both S and T. It would be interesting to do the sliding-spans comparisons that Findley and Monsell carried out, matching the amounts of smoothing of X-11 and STL as described in Section 6.2.

7. End Effects

Gray and Thomson have argued that loess does heavier smoothing at the ends than in the middle. We can see from the previous section that this is not so when the amount of smoothing is measured by the standard deviation: loess does not ever reduce the standard deviation appreciably more at the ends than in the middle and usually smooths much less. The difficulty with the second-difference measure used by Gray and Thomson is that it measured smoothness in a very local way that depends on artifacts of the weight function and that does not reflect what we normally would think of as smoothness when we look at a plot. However, there is one sense in which Gray and Thomson are correct – the potential for bias is greater at the ends than in the middle.

Gray and Thomson have correctly observed that our analysis in Section 5 ignores end effects. It would, however, be possible to carry out an analysis that includes end effects by looking at the eigenvalues of the actual trend and seasonal operator matrices, rather than the circularized ones as we have done, and also invoking the results of Buja, Hastie and Tibshirani (1989).

8. SABL and STL

Nicholls has asked about the computational aspects of STL and SABL, our first attempt at building a seasonal-trend decomposition procedure. SABL, like X-11, is slow. It would have been a major computational chore, perhaps on the order of hours rather than seconds, to apply SABL to the daily CO_2 data that we analyzed in the paper. Furthermore, SABL, like X-11, is much more complicated than STL. In fact, we undertook the design of STL when certain experiments indicated that SABL was more complicated than necessary.

9. Accomplishments and Additions

In our paper we have described a seasonal-trend decomposition procedure that is considerably simpler than X-11 and yet has many new capabilities. Still, the basic template of STL uses a number of ideas from X-11, and if the U.S. Bureau of the Census did not have the claim to the name, we might have called our procedure "X-12."

We have chosen the details of our first implementation so that the procedure works on the data in which we are interested – climatology measurements. We fully acknowledge that these details will not achieve desired goals for all types of data or perhaps even any other type of data. But because we have devised an open system, other researchers can attempt to modify the details of STL to make it work for their data. The following is a list of desirable additions to our first implementation that have arisen from our own observations and from those of the discussants; we are sure, however, that the list would grow considerably if we had a round-table discussion of all authors and discussants:

- Add locally-quadratic fitting to the trend smoothing; add locally-constant fitting and a hybrid of locally-constant and locally-linear fitting to the seasonal smoothers
- Add trading-day estimation
- Add a capability to the code for specifying robustness weights a priori to facilitate the adjustment of component series and their total
- Write a *C* implementation of STL
- Write an implementation of STL that provides for missing values
- Experiment further with STL-ARIMA
- Experiment with cross-validation to pick $n_{(s)}$
- Study the eigenvalue properties of the trend and seasonal operator matrices without making the circularity assumption
- Study the eigenvalue properties when trading-day estimation is included
- Carry out further sliding-spans analyses of STL.

We hope that others in seasonal adjustment will take the basic template of STL and see if the details of our first implementation can be modified to work on their data, perhaps by implementing some of the above changes. We believe one result would be a more fertile environment for pursuing new ideas and improvements because, as we have emphasized, having an open system and an increased ability to reason abstractly would be of benefit to a field now constrained by the closed system of X-11.

10. References

Bacastow, R.B. (1976). Modulation of Atmospheric Carbon Dioxide by the Southern Oscillation. Nature, 261, 116–118.

Buja, A., Hastie, T.J., and Tibshirani, R.I. (1989). Linear Smoothers and Additive Models (with discussion). The Annals of Statistics, 17, 453–555.

Findley, D.B., Monsell, D.B., Shulman, H.B., and Pugh, M.G. (1990). Sliding-Spans Diagnostics for Seasonal and Related Adjustments. Journal of the American Statistical Association (to appear).

Kuhn, T.S. (1962). The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

Kukla, G. and Gavin, J. (1981). Summer Ice and Carbon Dioxide. Science, 214, 497–503.

Macaulay, F.R. (1931). The Smoothing of Time Series. New York: National Bureau of Economic Research.

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