How the X-11 program implements a trend-seasonal decomposition

Alistair Gray
Mathematical Statistics Division
Department of Statistics
Wellington New Zealand

August 1990

The basic design of the X-11 procedure applied to monthly series is shown in Figure 1. Here

- Y_t is decomposed multiplicatively or additively into T_t, S_t, I_t;
- W_t is a robustness weight associated with each point;
- Initial estimates of T_t, S_t, I_t are obtained in a manner to be described;
- Two iterations are carried out;
- Within the iterated steps it is possible to estimate a further component, the calendar component C_t, having first initialized it as done for the other components. The details of how X-11 does this will be described later. For now, it is sufficient to say that at the end of the second seasonal smooth, on the first iteration, a modified I_t is computed and this is used as the regrediend for the regressors, $X_{1t}, \ldots, X_{7t}$, where X_{it} is essentially the number of times the ith day of the week occurs in month t. The predicted values of I_t under the model are removed from Y_t before the *Adjust W_t* step.

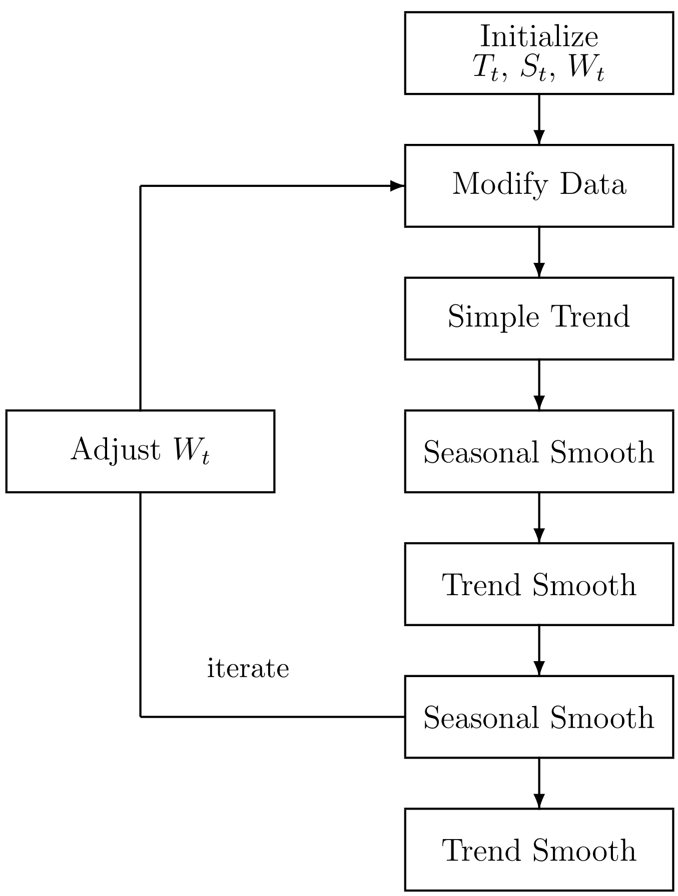

Figure 1: X-11: basic design

The *Modify Data* step is shown in Figure 2. Here

- *Compute* I_t is either dividing Y_t by T_tS_t or subtracting from Y_T, T_t and S_t;

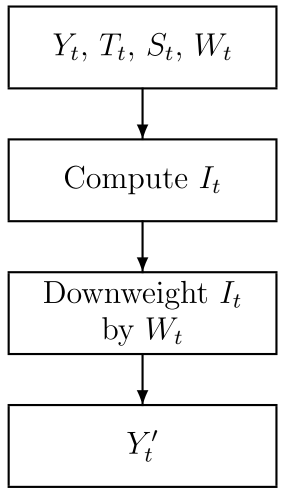

Figure 2: X-11: modify data

- Y'_T is either $T_tS_t(1 + W_t(I_t - 1))$ or $T_t + S_t + W_tI_t$

The *Simple Trend* step applies to the modified data a centred 12 term weighted moving mean (average) with weights given by

$$w_j = \begin{cases} \frac{1}{s} & \text{if } s = 2p + 1 \text{ and } j = -p, \ldots, p \\ \frac{1}{s} & \text{if } s = 2p \text{ and } j = -p + 1, \ldots, p - 1 \\ \frac{1}{2s} & \text{if } s = 2p \text{ and } j = -p, p. \end{cases}$$

It can be implemented as a 12 term simple moving mean, followed by a 2 term simple moving mean, written 2×12 ma. This loses half a cycle at each end of the series.

The *Seasonal Smooth* step is shown in Figure 3. Here

- *Remove T_t from Y_t* is dividing T_t into, or subtracting it from Y_t;
- The seasonal filters can be chosen from any of the simple moving means: 3 ma, 3×3 ma, 3×5 ma, 3×9 ma, or the overall mean;
- At the ends of the series asymmetric filters are used so that no further points are lost. The design of the asymmetric filters is not consistent across filters. The main principle as quoted in Shiskin et al. (1967) is that the

  > *future seasonal [...] values will be at approximately the same level as the most recent values*

  This seems more or less the case for the 3×3 and 3×5 moving averages. But the 3 term and 3×9 term moving averages seem to be Musgrave end filters based on a minimum revisions criterion (q.v. discussion on *Trend Smooth*).
- *Remove trend from S_t* is applying a 2×12 ma to S_t: the months in the half cycle at each end which are lost are regained by repeating the nearest available seasonal for that month (this is the simplest form of DeForest extension);

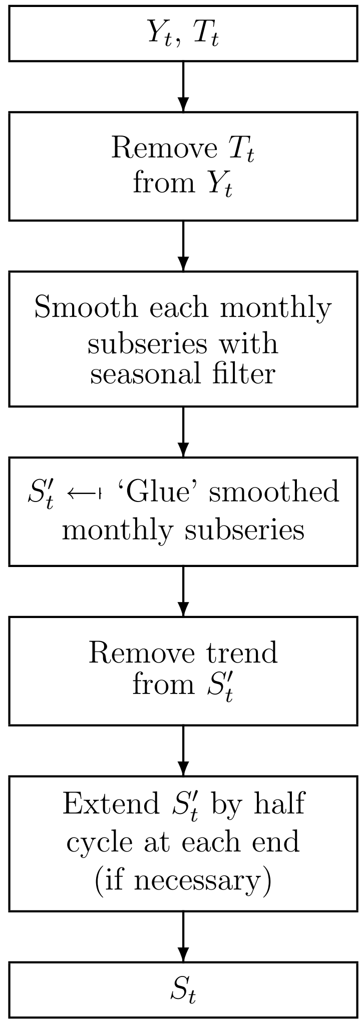

Figure 3: X-11: seasonal smooth

- The above step helps to ensure that no part of the trend has leaked into the seasonal and that

$$\frac{1}{s}\sum_{s} S_i = \begin{cases} 1 & \text{if multiplicative} \\ 0 & \text{if additive;} \end{cases}$$

- When a simple trend has been used prior to the seasonal smooth then S_t is extended by half a cycle at each end by repeating the nearest available seasonal as in removing the trend from S_t.

In the *Trend Smooth* step any of a 9, 13, or 23 term Henderson weighted moving mean is applied to the modified data with the seasonal subtracted from it. Because there is no need to remove the seasonal, as is the case with the initial trend, both the window width and the filter design can be tailored to other criteria. Such filters are designed to estimate the 'smoothest' piecewise third order polynomial passing through the data. In economic time series we often expect to see turning points in the trend, and so the requirement that the trend filter pass a low order polynomial is sensible.

At the ends of the series asymmetric weights are used so that no further points are lost. The design of these filters for the 9, 13 and 23 term Henderson filters is due to Musgrave (1964) and is based on the following idea. Suppose that the data really follows

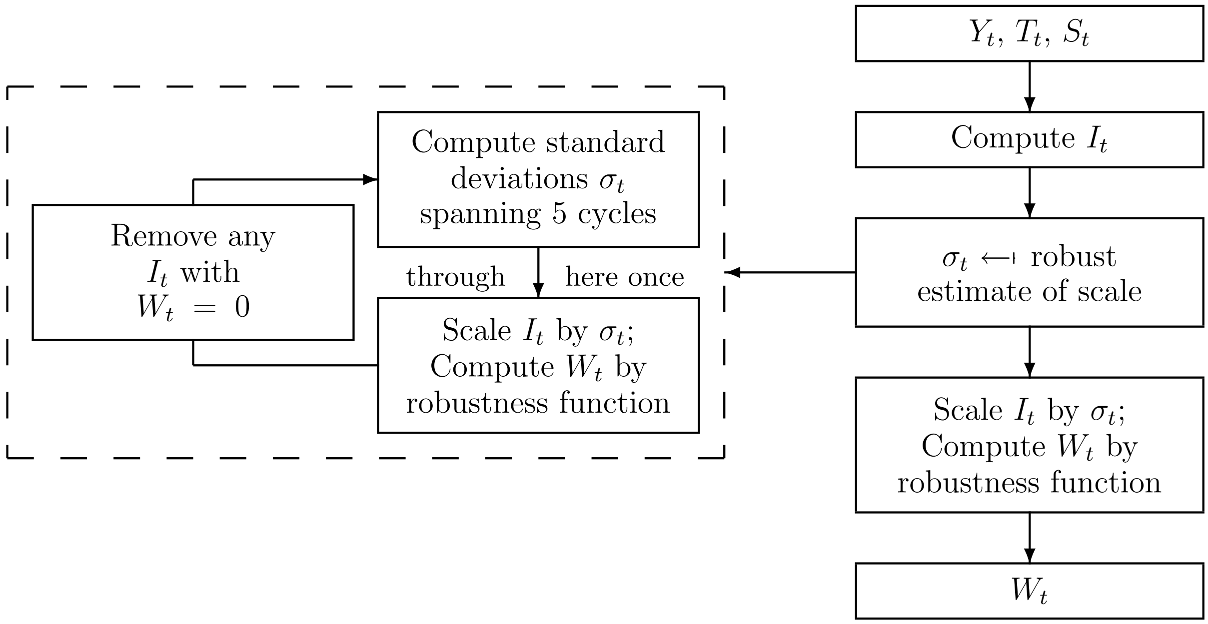

Figure 4: X-11: adjust weight

a linear trend with noise. What asymmetric filter would produce smoothed values whose values would be changed the least as enough data came to hand to use the symmetric weight. This is a miminum revisions criterion.

The *Adjust* W_t step updates the estimates of how reliable each point is based on the current values of Y'_t, T_t, and S_t, and is shown in Figure 4. Here

- In *Robust estimate of scale*, in *Compute standard deviation spanning 5 cycles*, for the first two and last two years, the value for the third and third to last year, respectively, is used;

- In both the Initialize step and the iterated loop, *Scale* I_t *by* σ_t mean forming

$$I'_t \longleftarrow \frac{|I_T - \mathbf{E}(I_t)|}{\sigma_t};$$

- In the Initialize step the robustness function is

$$W_t = \begin{cases} 1 & \text{if } I'_t \leq \lambda_{zeroweight} \\ 0 & \text{otherwise} \end{cases}$$

  where $\lambda_{zeroweight}$ can be altered and by default is 2.5;

- In the iterated loop the robustness function is

$$W_t = \begin{cases} 1 & \text{if } I'_t \leq \lambda_{fullweight} \\ \alpha & \text{if } \lambda_{fullweight} < I'_t < \lambda_{zeroweight} \\ 0 & \text{if } I'_t \geq \lambda_{zeroweight} \end{cases}$$

  where

$$\alpha = \frac{(I'_t - \lambda_{zeroweight})}{(\lambda_{fullweight} - \lambda_{zeroweight})}$$

  and $\lambda_{fullweight}$ can be altered and is by default 1.5.

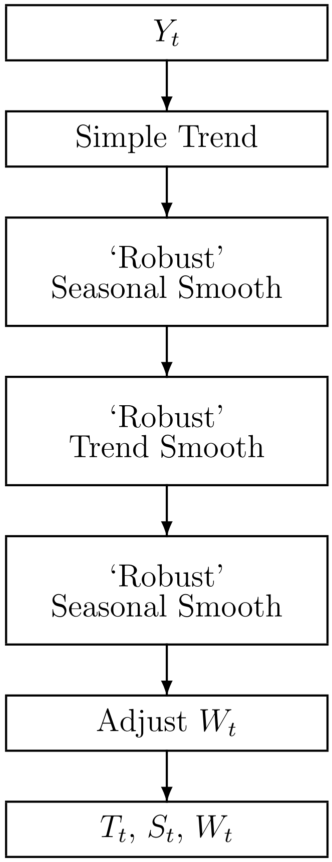

Figure 5: X-11: initialize T_t, S_t, W_t

- Shiskin et al. (1967) says that the robustness function in the iterated loop with the robust estimate of scale accounts for extremes in the trend, improves stability of the seasonal in the presence of additional data and identifies more extremes in a single month. This is based on widespread empirical testing rather than theoretical properties.

How X-11 initializes T_t, S_t, and W_t is as complicated as how it computes them in the iterated steps of the basic design described above, and in some ways are not consistent with more recent ideas. Yet it has to be said that X-11 works in practice, which can't be said of some recent seasonal adjustment procedures.

The *Initialize* T_t, S_t, W_t step is shown in Figure 5. The initial estimate of the calendar component if required is made after the second 'robust' seasonal step.

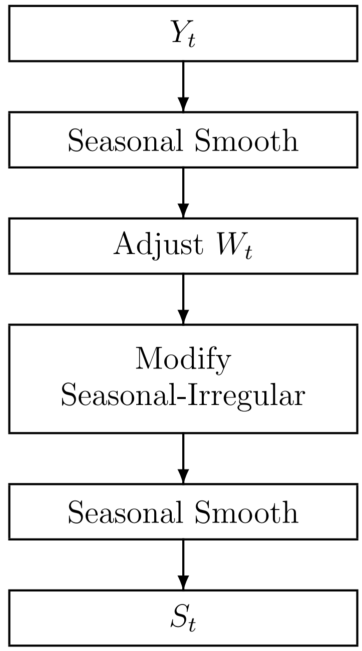

Figure 6: X-11: 'robust' seasonal smooth

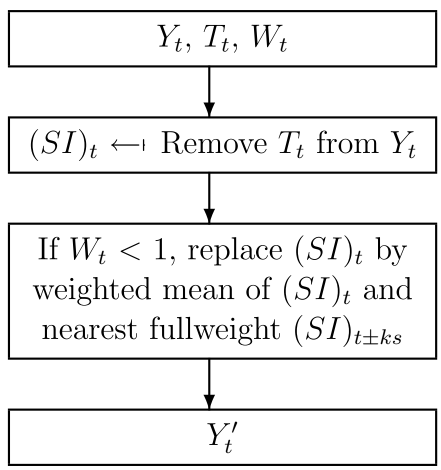

Figure 7: X-11: modify seasonal-irregular

The *'Robust' Seasonal Smooth* step is shown in Figure 6. Here

- The *Seasonal Smooth* step is the same process as in iterated loop but applied to Y_t with the first simple trend removed. However the filter is restricted to a 3×3 ma the first time and a 3×5 ma the second time.

- The *Adjust* W_t step is also the same as in the iterated body but applied to Y_t with the simple trend and the first seasonal smooth removed;

- The *Modify Seasonal-Irregular* step is shown in Figure 7. Here

  – In *Replace* $(SI)_t$, for the 3rd year to 3rd - to - last year, the nearest two fullweight $(SI)_{t \pm ks}$'s either side of the non fullweight $(SI)_t$ are used; otherwise the nearest three fullweight $(SI)_{t \pm ks}$'s are used, (s is the length of the cycle); the weighted mean uses weights $1/(n + W_t)$ for the n fullweight $(SI)_{t \pm ks}$'s and weight $W_t/(n + W_t)$ for the outlier $(SI)_t$.

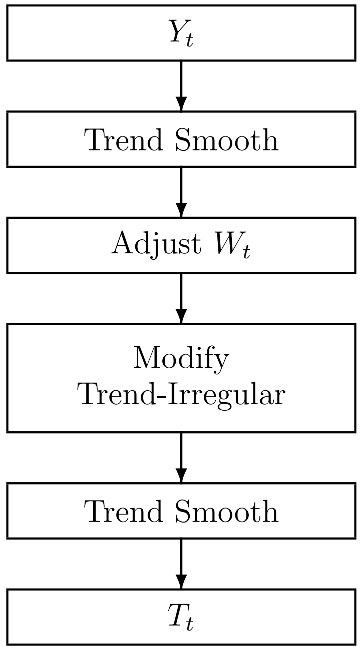

Figure 8: X-11: 'robust' trend smooth

- the modified data, Y_t', is either $T_t(SI)_t$ or $T_t + (SI)_t$;
- If W_t is almost 1 then the replaced $(SI)_t$ is almost the mean of the five values, whereas if W_t is almost 0 then the replaced $(SI)_t$ is almost the mean of the four 'good values'.

  A trimmed mean for each month might seem to be better way because it is simpler. However, Shiskin et al. (1967) suggests the a trimmed mean approach leads to substantial revisions in the seasonal factors when new data comes to hand. The above method of weighting and replacement seems to overcome this.

The *'Robust' Trend Smooth* step is shown in Figure 8. Note that this only happens if the 'strike option' is activated (Table B 7). Here

- The *Trend Smooth* is the same process as in the iterated body but applied to Y_t with the first 'robust' seasonal removed;
- The *Adjust* W_t step is also the same as in the iterated body but applied to Y_t with the first 'robust' seasonal and the first trend smooth removed;
- The *Modify Trend-Irregular* step is shown in Figure 9. Here
  - In *Replace* $(TI)_t$, for the 3rd month to 3rd - to - last month, the nearest two fullweight $(TI)_{t\pm ks}$'s either side of the non fullweight $(TI)_t$ are used; otherwise the nearest three fullweight $(TI)_{t\pm ks}$'s are used;

    the weighted mean uses weights $1/(n+W_t)$ for the n fullweight $(TI)_{t\pm ks}$'s and weight $W_t/(n+W_t)$ for the outlier $(TI)_t$.
  - the modified data, Y_t' is either $(TI)_tS_t$ or $(TI)_t + S_t$;
  - If W_t is almost 1 then the replaced $(TI)_t$ is almost the mean of the five values, whereas if W_t is almost 0 then the replaced $(TI)_t$ is almost the mean of the four 'good values'.

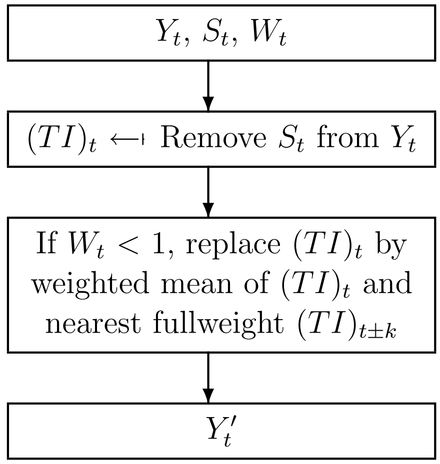

Figure 9: X-11: modify trend-irregular

- Similar comments can be made about the sensibility of the downweighting as were made in discussing the 'robust' seasonal smooth.

The *Adjust* W_t step is also the same as in the iterated loop but applied to Y_t with the 'robust' trend smooth and the second 'robust' seasonal smooth removed.

References

Musgrave, J. C. (1964). A set of end weights to end all end weights. Working paper, Bureau of the Census, US Department of Commerce, Washington, D.C.

Shiskin, J., Young, A. H., and Musgrave, J. C. (1967). The X-11 variant of the Census method II seasonal adjustment program. Technical Paper 15, Bureau of the Census, U.S. Department of Commerce, Washington, D.C.