

# Systematic Approach to Seasonal Adjustment of Time Series Data: Methods and Example

**Chuanzhong Sun**

SAS Institute Inc, 100 SAS Campus Drive, Cary, NC 27513, USA

Email: Charles.Sun@sas.com

## Abstract

There have been three primary methods in the world of seasonal adjustment. The first one is the X-12-ARIMA based filter (also referred to as x11 smoothing) developed by the United States Census Bureau. This technique has been used by statistics-based governmental agencies in US, United Kingdom, Australia, Canada, and other international agencies. The second technique is ARIMA model-based signal extraction (also referred to as SEATS), developed primarily by the Bank of Spain. This technique has been used by many organizations in Europe and elsewhere. The third method is called the structural time series-based decomposition (also referred to as UCM). It has been implemented in popular statistical software such as SAS, STAMP, and STATA, as well as in open-source statistical software, but has not been widely used by the agencies. Each of the three methods has advantages and disadvantages for specific situations. In this paper, I propose a hybrid and systematic way to do the seasonal adjustment by combining these three methods. Data of monthly retail sales of consumer goods in China from January 2000 to December 2011 is used for analysis. Results show that for the given time series, if a desired time series model can be built, then seasonal adjustment results with these three methods are very similar. Otherwise, a systematic approach that recursively applies these methods is suggested. SAS software is used to illustrate and account for seasonal factor adjustments. Finally, general recommendations for systematic approaches to seasonal adjustment are discussed.

**Keywords:** Seasonal adjustment, X-12-ARIMA, SEATS, UCM.

## 1. Introduction

Economic data and most retail data are collected periodically. They are typically called time series data and usually contain seasonal patterns. The seasonal patterns in a time series need to be adjusted for forecasts or other analysis purposes. For monthly or quarterly data, several methods have been developed for seasonal adjustment of a time series. The first one, called the “ad\_hoc” filter approach, includes methods that use moving average smoothing filters to estimate the signal. The X-12-ARIMA tool developed by the Census Bureau of United States is a well-known example of applying this method. The second method is the ARIMA model-based (AMB, also called SEATS) approach, which takes into account of model structure of a time series. The ARIMA model-based procedure identifies a parsimonious ARIMA model for the prior-adjusted time series. Then ARIMA models for each component are derived with the restriction that the aggregation of those models yields to the overall ARIMA model. Because there is no unique admissible decomposition, the method applies the ‘canonical decomposition’. Within this approach, the most popular algorithm is the SEATS/TRAMO software (Gómez and Maravall (1997a, 1997b)). The third method is called the structural time series based approach (STS, also referred to as UCM) (Harvey

(1989)). With this method, instead of specifying a model for the observed series, one directly assumes different stochastic linear model for each unobserved component. These models are formulated within a stochastic state space setting, and the Kalman Filter is used to estimate the parameters. As to our knowledge, this method has not been used as widely as other two methods in seasonal adjustment.

Catherine presented some comparisons between SEATS/TRAMO and X-12-ARIMA using U.S. Import/Export series data (Catherine C. Hood, 2002). In fact, each of the three methods has advantages and disadvantages under specific situations. In this paper, I will suggest a systematic way to do the seasonal adjustment by considering all of these three methods. Series of monthly retail sales of consumer goods in China from January 2000 to December 2011 is used for demonstration. SAS software is used for seasonal adjustment example.

## **2. Systematic Method for Seasonal Adjustment**

Given the availability of the three methods in software such as SAS, I am suggesting a systematic approach for seasonal adjustment by considering all of the methods. The basic idea is to start with the three methods in their simple and default options and compare the seasonal adjustment results, then pick the “best” one. For a given time series, if the picked results are desirable, the process stop. Otherwise, repeat the process by refining the model options in the three methods. General steps for this approach are given below. Figure 2.1 show the process in a flowchart.

Step 1: Plot the time series to visually check any abnormal observations, missing values, or other outstanding issues.

Step 2: Test the need for proper transformation and do the data transformation if needed. Most commonly it will be the log transformation.

Step 3: Use TRAMO type automatic modeling process to test and estimate related calendar effects such as trading day, moving holiday effects such as Easter Day, outliers, and user specified effects, and find the ‘best’ regARIMA model for the time series. If a ‘desirable’ model can be found, the series is prio-adjusted and go to Step 4. Otherwise go to Step 5.

Step 4: Apply X-11 smoothing filter on the prio-adjusted series and use the obtained ARIMA model for SEATS decomposition on the same prio-adjusted series. Meantime, test the significance of seasonality in the prio-adjusted time series.

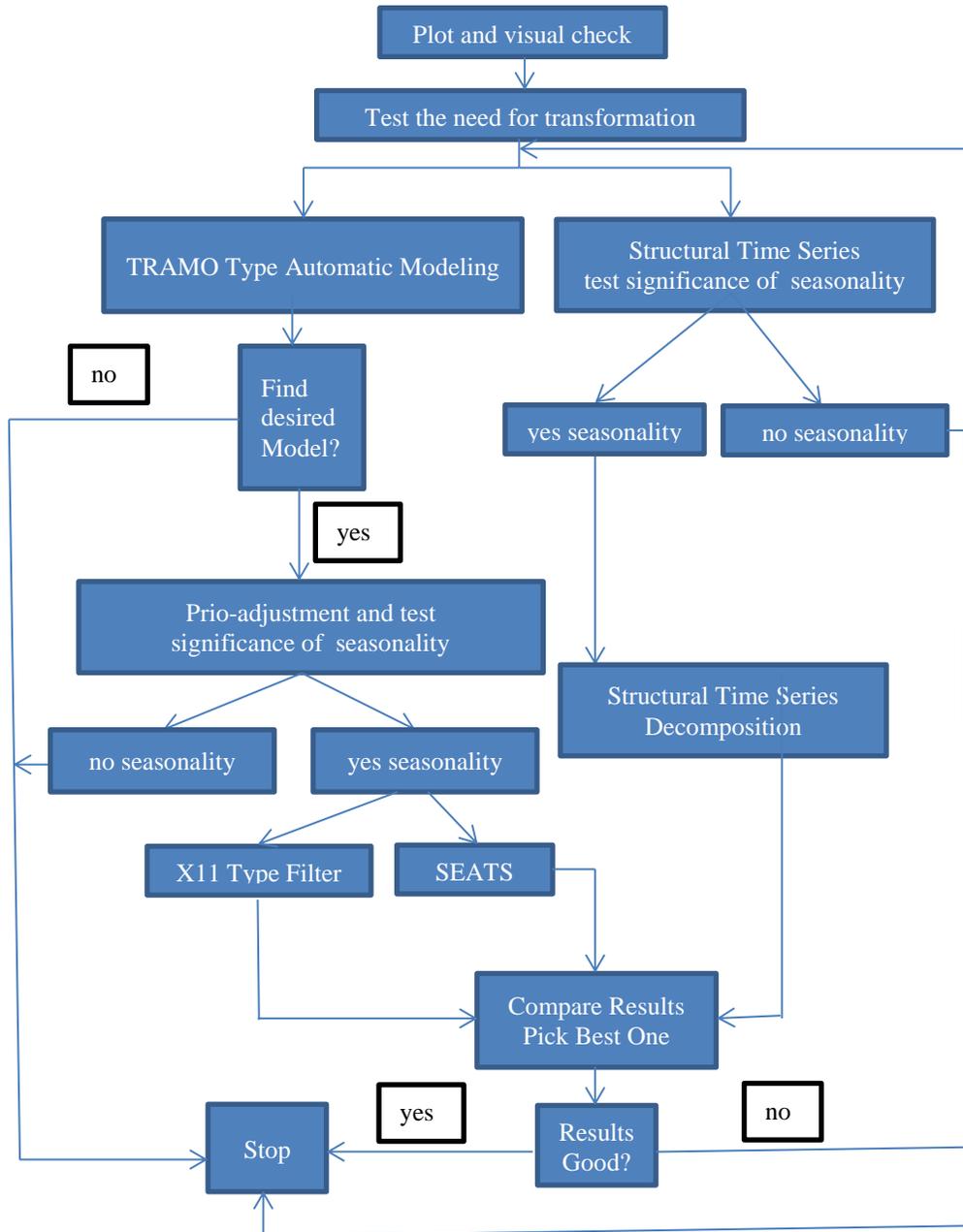
Step 5: User the structural time series method to specify each part of the components. Start with Basic Structural Model (BSM) and gradually increase the complexity of model specification if needed. Test the significance of seasonality. If the seasonal components are significant then use the smoothed seasonal components as seasonal factor estimates.

Step 6: If the “desirable’ regARIMA model is found, compare the estimated seasonal factors from the three methods and pick the “best” one for seasonal adjustment. Otherwise, use the results from BSM for seasonal adjustment.

Step 7: Test seasonally adjusted series. If the results are desirable then stop, otherwise repeat Step3 through Step 5 with different model specifications.

In SAS Analytical Products 12.1, both the X-12-ARIMA and SEATS algorithms are implemented in PROC X12. PROC UCM has the algorithm for STS method. So, it is convenient to conduct above systematic approach with SAS software.

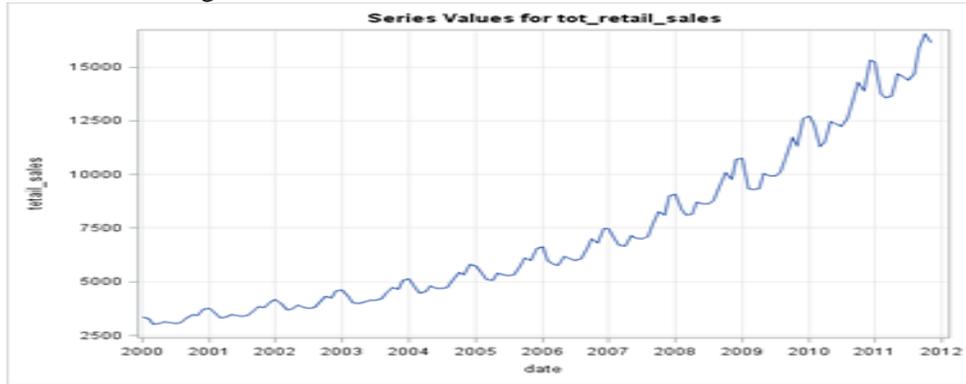
Fig. 2.1 Flowchart of Systematic Method for Seasonal Adjustment



### 3. Example With Monthly Retail Sales of Consumer Goods in China

The time series analyzed in this example is the monthly series of total retail sales of consumer goods in China from January 2000 to December 2011. Figure 3.1 shows the original time series. The graphic suggests a strong seasonality and a need for log transformation, so steps for testing the seasonality and transformation are omitted here.

Fig. 3.1 Total Retail Sales of Consumer Goods in China



To use the TRAMO type method for automatic modeling, we need to consider other potential effects on the time series. For sales data one of the important factors is the Chinese New Year. Three dummy variables are created for these effects:  $x_{wb}$ ,  $x_w$ , and  $x_{wa}$ . They represent the effects of before festival, during festival, and after festival respectively. They are measured in the similar way as the Easter Day dummy variables in X-12-ARIMA. Assumed that the time series is stored in a dataset called `cc.china_dat1_x` and the variable name is `tot_retail_sales`. To find the “best” regARIMA model, following SAS program is submitted:

```
proc x12 data=cc.china_dat1_x date=date
interval=12 span=(jan00,dec11)
mdlinfoout=mymdl_5;
var tot_retail_sales;
transform function=log;
regression uservar=x_wb x_w x_wa/usertype=holiday holiday holiday;
automdl difforder=(1,1) balanced;
estimate;
check maxlag=12 print=all;
outlier type=(AO TC LS) ;
run;
```

Output from submitting this program is omitted. The primary results are:

- 1) Chinese New Year effects  $x_{wb}$  and  $x_w$  are significant. The After Chinese New Year effect  $x_{wa}$  is not significant.
- 2) Automatically selected “best” ARIMA model is (1 1 0)(0 1 0).
- 3) Following outliers are automatically identified: AO jan2002, AO may2003, AO jan2005, AO feb2005, AO dec2007, AO jan2008, LS jan2010, AO feb2010.

Based on the results, the original SAS program is modified as:

```
proc x12 data=cc.china_dat1_x date=date
interval=12 span=(jan00,dec11)
mdlinfoout=mymdl_6;
var tot_retail_sales;
transform function=log;
regression uservar=x_wb x_w /usertype=holiday holiday;
automdl difforder=(1,1) balanced;
estimate;
check maxlag=12 print=all;
outlier type=(AO TC LS);
event ao01jan2002d ao01may2003d ao01jan2005d ao01feb2005d ao01dec2007d
ao01jan2008d ls01jan2010d ao01feb2010d;
run;
```

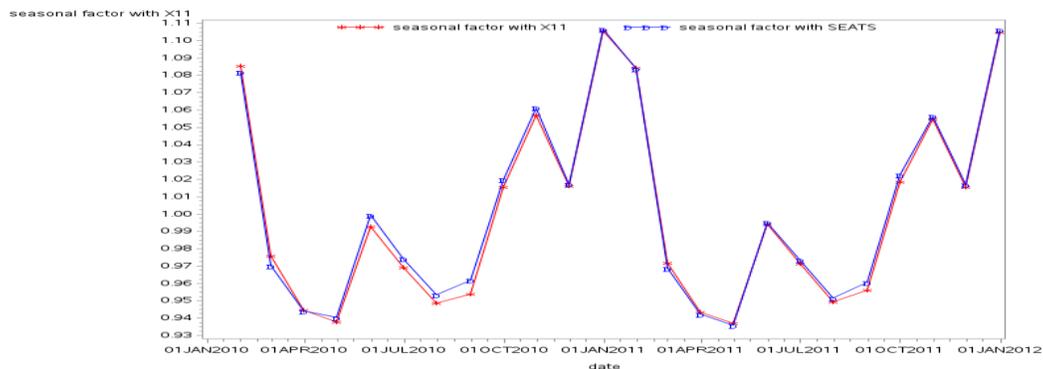
Running this program, the results show that:

- 1) All the regression parameters are significant.
- 2) No significant autocorrelation in model residual.
- 3) Model residual passed the normality test.
- 4) Same model ARIMA:(1 1 0)(0 1 0) is selected

So, the obtained model seems desirable. Next, we will use the obtained model for seasonal adjustment with both SEATS and X-12-ARIM methods. Following SAS program are submitted for this purpose. The seasonal factor estimates from these two methods are actually very close. Figure 3.2 shows the estimated seasonal factors for last two years.

```
proc x12 data=cc.china_dat1_x date=date interval=12 span=(jan00,dec11)
mdlinfoin=mymdl_6;
var tot_retail_sales;
regression uservar=x_wb x_w /usertype=holiday holiday;
event ao01jan2002d ao01may2003d ao01jan2005d ao01feb2005d ao01dec2007d
ao01jan2008d ls01jan2010d ao01feb2010d;
forecast lead=24 nback=24;
output out=sales_x11 B1 d11 D12 D10;
X11 final=all;
SEATSDECOMP OUT=sales_seats LEAD=24 NBACK=24;
run;
```

Fig. 3.2 Estimated Seasonal Factors for Last Two Years

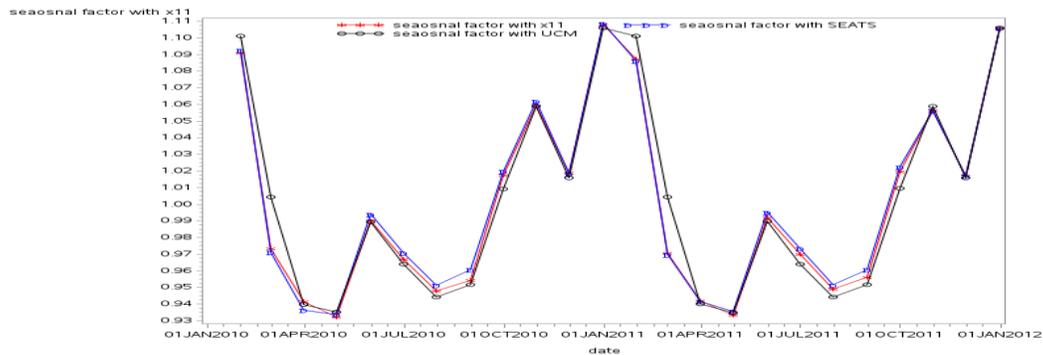


Diagnostics on the seasonally adjusted series show that there is no significant remaining seasonality in seasonally adjusted series. Although the results seem desirable, for demonstrating purpose, we are giving a SAS example to use structural time series method:

```
proc ucm data=chdat1(where=(date between '31Jan2000'd and '31Dec2011'd));
id date interval = month;
model logsales=x_wb x_w ;
irregular ;
level checkbreak print=smooth;
slope print=smooth;
season length = 12 type=trig ;
outlier print=short;
estimate outest=est1 plot=panel;
forecast back=12 lead=12 print=fdecomp outfor=my_ucm;
run;
```

In above SAS program, the dependent variable logsales is the logarithm transformation of original variable tot\_retail\_sales. The SAS data set chdat1 is created based on previous data set cc.china\_dat1\_x. The seasonality is specified in trigonometric form with slowly evolution. This is a basic structure model with regression. The UCM model specification is relatively simple compared with regARIMA model. Again, the model residual diagnostics is acceptable and the estimated seasonal factors are quite close to those from SEATS and X-12-ARIMA. In addition, only two outliers are detected here. One of reasons could be the fact that UCM is using a random walk as trend. Figure 3.3 shows the plots of the last two years from the three methods.

Fig. 3.3 Estimated Seasonal Factors for Last Two Years



#### 4. Conclusions and Recommendations

By considering the three methods simultaneously, we will more likely to get a “better” seasonal adjustment results. In our example, a decent regARIMA model is found, so the estimated seasonal factors from the three methods are very close. It is possible that a decent regARIMA model may not be found and the results from X11 filter and SEATS may not be desirable. In such a case, with the choice of STS, one may still get a good seasonal adjustment. Also, with STS method, one may need a much simpler model to get the decent seasonal factor estimates. More researches are needed to investigate the favorable situations for each method.

In SAS software, procedure X12 has the functionality of automatic regARIMA modeling, X11 filters, and SEATS decomposition. The procedure UCM can handle the task of structural time series modeling. Therefore, it is convenient to practice the systematic methods within SAS system. In addition, the graphics feature in SAS system will be very helpful in monitoring the process.

#### References

- Bell, W.R. and S. C. Hillmer. (1984) “Issues involved with the seasonal adjustment of economic time series,” *Journal of Business and Economic Statistics* 2, 291–320
- Burman, J. P. (1980) “Seasonal adjustment by signal extraction,” *Journal of the Royal Statistical Society series A* 143, 321–336
- Catherine C. Hood. (2002) “Comparison of Time Series Characteristics for Seasonal Adjustments from SEATS and X-12-ARIMA,” *JSM Proceedings 2002*, URL <http://www.amstat.org/sections/srms/Proceedings/y2002/Files/JSM2002-000767.pdf>
- Gómez, V. and Maravall, A. (1997a), *Guide for Using the Programs TRAMO and SEATS, Beta Version*, Madrid: Banco de España.
- Gómez, V. and Maravall, A. (1997b), *Programs TRAMO and SEATS: Instructions for the User, Beta Version*, Madrid: Banco de España.
- Harvey, A. C. (1989), *Forecasting, Structural Time Series Models, and the Kalman Filter*, Cambridge: Cambridge University Press.
- U.S. Bureau of the Census (2010) “X-12-ARIMA Seasonal Adjustment Program,” URL <http://www.census.gov/srd/www/x12a/>