

Paper 261-2010

Financial Analysis Using SAS® PROCs

Somaye Gharibvand, Multimedia University, Cyberjaya, Malaysia

Lida Gharibvand, University of California, Riverside, USA

ABSTRACT

Financial services industry is interested in analyzing vast financial data including price trends from stock exchanges around the world. SAS analytical and graphical tools are extremely useful to enable statistical analysis of various financial data including stock price information. Time Series statistical method is a valuable approach for analysis of financial data. Time Series modeling involves trends, seasonality, cyclical behavior, and forecasting future trends using historical data collected at regular time intervals. Many advanced Time Series analysis procedures are available in SAS/ETS module. In this paper, three PROCs are shown: PROC TIMESERIES, PROC FORECASTING, and PROC UCM. The advanced features of SAS procedures were used to analyze Texas Instruments stock with TXN symbol. It will be demonstrated that PROC UCM (Unobserved Components Model) is better suited to analyze the stock price movements.

INTRODUCTION TO TIME SERIES

The term Time Series refers to a sequence of data points which are spaced at uniform intervals in time and are measured repeatedly at successive times. The Time Series data points usually arise when monitoring industrial operation processes or measuring corporate business metrics. The term Time Series Analysis refers to methods that are used to understand the behavior of the time series events. Time Series Analysis takes into account the fact that data points collected over time often have an internal structure such as autocorrelation, trend or seasonal variation. The Time Series Analysis is performed to determine where the underlying data points came from and make a prediction about the future behavior of the data points. Time Series Analysis is used for many applications such as Economic Forecasting, Currency Exchange Rate Index, Sales Forecasting, Budgetary Analysis, Stock Market Analysis, Yield Projections, Process and Quality Control, Inventory Studies, Workload Projections, Utility Studies, Census Analysis, Cigarette Smoking Activity Patterns, Wireless Fading Channels, etc. The term Time Series Forecasting refers to the use of a statistical model to forecast future events based on known past events. In a time series model, the observations that are close together in time will be more closely related than the observations that are farther apart in time. Time Series statistical model can be used to try to predict the future stock price of a company based on the stock price data points from the past. This paper is an attempt to introduce important SAS tools that can be applied to analyze Time Series data related to the market price of a company's stock.

SAS ODS STATISTICAL GRAPHICS

ODS graphics was experimental in SAS 9.1 and is now embedded in SAS 9.2 (SAS Institute, Inc. (2008)). With ODS Graphics, a procedure creates the graphs that are most commonly needed for a particular analysis. The capabilities of ODS Graphics has been expanded to include new graph types, an interactive editor, etc. The use of ODS Graphics allows us to make customized statistical graphics with ease while maintaining a professional appearance.

SAS PROCs:

In this paper, three SAS PROCs are used. The TIMESERIES procedure is descriptive and well suited for pre processing data, and to perform explanatory graphical analysis using SAS ODS Graphics. The FORECASTING procedure provides a one-step method to automatically generate forecasts for hundreds of time series at a time. However, no ODS Graphics are available in the FORECASTING procedure. The UCM procedure uses the Unobserved Components Model (UCM) to analyze and forecast the equally spaced univariate time series data. It breaks down the response series into components that are useful in explaining and predicting its behavior such as trends, seasonal factors, cycles, and regression effects due to predictor series. The UCM procedure combines the versatility of the ARIMA model with interpretability of the smoothing model. Advanced Time Series analysis plots can be generated using the ODS Graphics option in UCM PROC.

TIME SERIES DATA SOURCE

The data used for this presentation was downloaded from Yahoo Finance web site for TXN stock. From the figure called Series Plot, we can see some discontinuities in the graph which correspond to the five times when the stock was split by Texas Instruments.

Texas Instruments Stock Split Dates:

- June 15, 1987: 3 to 1 split (stockholders received 3 shares for 1)
- August 21, 1995: 2 to 1 split (stockholders received 2 shares for 1)
- November 24, 1997: 2 to 1 split (stockholders received 2 shares for 1)
- August 17, 1999: 2 to 1 split (stockholders received 2 shares for 1)
- May 23, 2000: 2 to 1 split (stockholders received 2 shares for 1)

Conclusion: 1 share on June 15, 1987 grew 1x3x2x2x2x2 to 48 shares in 13 years

PROC TIMESERIES

The TIMESERIES procedure computes various statistics to analyze the seasonal factors and the time period trends, and transforms the data into different Time Series format. Further analysis can be performed on the working Time Series by using techniques from PROC TIMESERIES: descriptive (global) statistics, seasonal decomposition/adjustment analysis, correlation analysis, and cross-correlation analysis. The analysis result is stored in output data sets or printed using the Output Delivery System (ODS) which can be used to create graphics. The Time Series format such as a working Time Series is useful for preparing the data for subsequent analysis using other SAS/ETS procedures.

```
ods graphics on;
proc timeseries data=txi out=monthtxi plot=(series corr decomp);
  Id date interval=month accumulate=median;
  var close;
run;
ods graphics off;
```

Code Box 1: PROC TIMESERIES

The monthly close from this data set was used. The daily series were converted to monthly series.

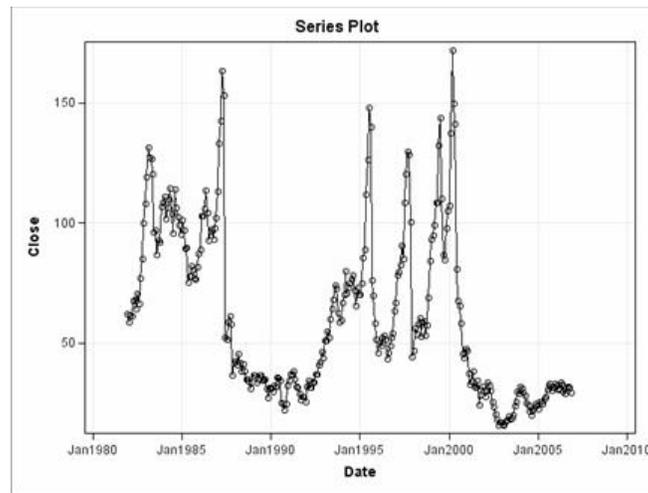
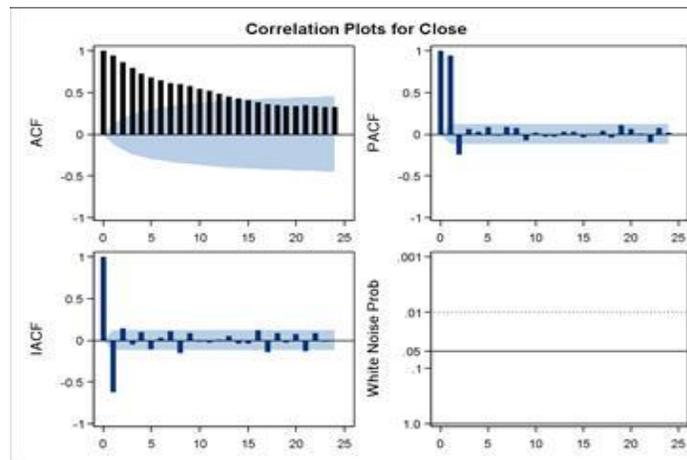
**Figure1:** The output from PROC TIMESERIES

Figure 1 is the plot of actual data which is the output from PROC TIMESERIES procedure.

**Figure 2:** Correlation Panel (Experimental)

The graphical displays are requested by specifying the ODS GRAPHICS statement and the PLOTS= options in the PROC TIMESERIES statement.

PROC FORECAST

The FORECAST procedure is based on extrapolation of data and the forecast for a series is solely a function of the time and the past values of that series. It is a one-step method to automatically generate forecasts for hundreds of series at a time. The Stepwise Autoregressive Method combines time trend regression with an autoregressive model and uses a stepwise method to select the lags to use for the autoregressive process. The Exponential Smoothing Method produces a time trend forecast but the parameters are allowed to change gradually over time, and earlier observations are given exponentially declining weights. The Holt-Winters Method combines a time trend with multiplicative seasonal factors to account for regular seasonal fluctuations. The additive version uses additive seasonal factors. The FORECAST procedure generates the forecasts and confidence limits for an output data set. PROC FORECAST uses extrapolation to generate practical results quickly which lead to approximate statistical results such as confidence limits. To use PROC FORECAST, the input and output data sets and the number of periods to forecast must be specified in the PROC FORECAST statement. Then, the variables to forecast must be listed in a VAR statement.

```
proc forecast data=monthtxi interval=month
              method=expo trend=2 lead=6
              out=out outfull outest=est;
  id date;
  var close;

run;
```

Code Box 2: PROC FORECAST

There is no ODS Graphics output in the Forecasting Procedure. Therefore, we need to use PROC GPLOT to generate the graphics.

```
symbol1 i=none c=black v=X f='Arial'
symbol2 i=spline c=red v=dot cv=blue ;
symbol3 i=spline c=blue l=3 v=L w=2 f='Arial' ; /* for _type_=L95 */
symbol4 i=spline l=3 c=orange v=U w=2 f='Arial' ; /* for _type_=U95 */

proc gplot data=out;
  plot close * date = _type_ / VAXIS=axis1 haxis=axis2 FRAME vminor=0
      hminor=0 cframe = white skipmiss;
axis1 label=( a=90 r=0 "close $")
      value=( ) width=2 ;
axis2 label=( "year")
      value=( ) width=2
      offset=(4 pct);

run;
quit;
```

Code Box 3: PROC GPLOT

Plot of Forecasts from Expo Method Texas Instrument stock

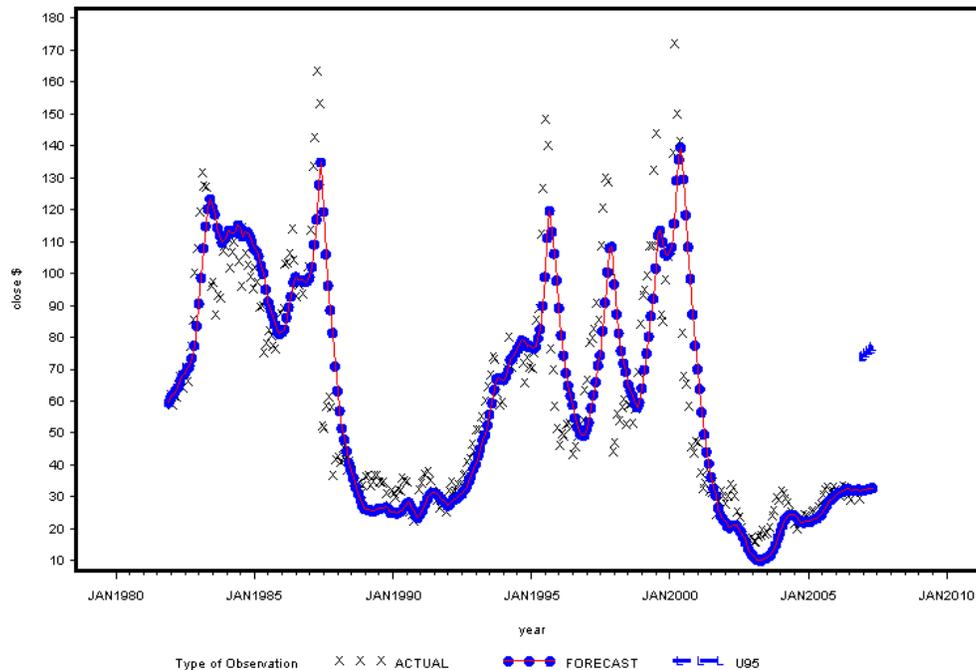


Figure 3: The output from PROC FORECASTING procedure using PROC GPLOT

Figure 3 shows the forecast plot with confidence limits.

PROC UCM

The UCM procedure uses the Unobserved Components Models (UCM) to analyze and forecast the equally spaced univariate Time Series data. UCM model breaks down the response series into components that are useful in explaining and predicting its behavior. It is used to fit a wide range of data with complex patterns such as trends, seasonal factors, and cyclical behavior which might include multiple predictors. It provides a variety of diagnostic tools to assess the fitted model and suggest modifications. UCM model is useful for forecasting the values of response series and component series in the model, obtaining a model-based seasonal decomposition, and obtaining the full sample or "smoothed" estimates of the component series in the model. The Basic Structural Model assumes data has a trend and seasonal component: $y_t = m_t + g_t + e_t$ where m_t is the trend, g_t is the seasonal and e_t is the irregular term. Trend and seasonal components can be modeled in a few different ways. A common model is the *locally linear time trend*:

$$\begin{aligned} m_t &= m_{t-1} + b_{t-1} + h_t, & h_t &\sim i.i.d. & N(0, s_h^2) \\ b_t &= b_{t-1} + e_t, & e_t &\sim i.i.d. & N(0, s_e^2) \end{aligned}$$

In this model, the level m_t and the slope b_t are assumed to be stochastic.

Special cases: $s_e^2=0$ implies a fixed slope equal to b_0 . $s_h^2=0$ and $s_e^2=0$ implies a deterministic trend equal to $m_0 + b_0 t$.

```
proc ucm data=monthtxi;
  id date interval=month;
  model close=june87;
  irregular;
  level;
  slope ;
  season length=12 type=trig ;
  deplag lags=2;
run;
```

Code Box 4: PROC UCM

Final Estimates of the Free Parameters					
Component	Parameter	Estimate	Approx Std Error	t Value	Approx Pr > t
Irregular	Error Variance	0.00000867	0.0068881	0.00	0.9990
Level	Error Variance	92.89802	8.60685	10.79	<.0001
Slope	Error Variance	3.970427E-8	0.00003430	0.00	0.9991
Season	Error Variance	3.343656E-9	2.67797E-6	0.00	0.9990
june87	Coefficient	-92.45851	9.83711	-9.40	<.0001
DepLag	Phi_1	0.23502	0.05536	4.25	<.0001
DepLag	Phi_2	-0.08908	0.05538	-1.61	0.1077

Figure 4: The parameter estimates for this model

The estimates suggest that the Slope can be treated as constant, i.e., has zero variance. Since slope and season components are not significant, it may then be useful to check if they can be dropped from the model. This can be checked by examining the significance analysis table of the components given in table below.

Significance Analysis of Components (Based on the Final State)			
Component	DF	Chi-Square	Pr > ChiSq
Irregular	1	0.00	0.9999
Level	1	146.08	<.0001
Slope	1	0.19	0.6635
Season	11	17.17	0.1029

Figure 5: Component Significance Analysis

From tables above we can conclude that Slope and Season components are not significant and should be dropped in the model. The Slope component can be made deterministic by holding the value of its error variance fixed at zero. This is done by modifying the SLOPE statement as follows: slope variance=0 noest.

The Irregular component's contribution appears insignificant towards the end of the estimation span; however, since it is a stochastic component it cannot be dropped from the model on the basis of this analysis alone. Although Slope and Season components are not significant for this data set, they might be significant for a different data set.

```
ods graphics on;
proc ucm data=monthtxi;
  id date interval=month;
  model close=june87;
  irregular;
  level;
  slope variance=0 noest;
  deplag lags=2;
  estimate back=24 plot=(residual normal acf);
  forecast back=6 lead=6 plot=(forecasts decomp);
run;
ods graphics off;
```

Code Box 5: Revised PROC UCM

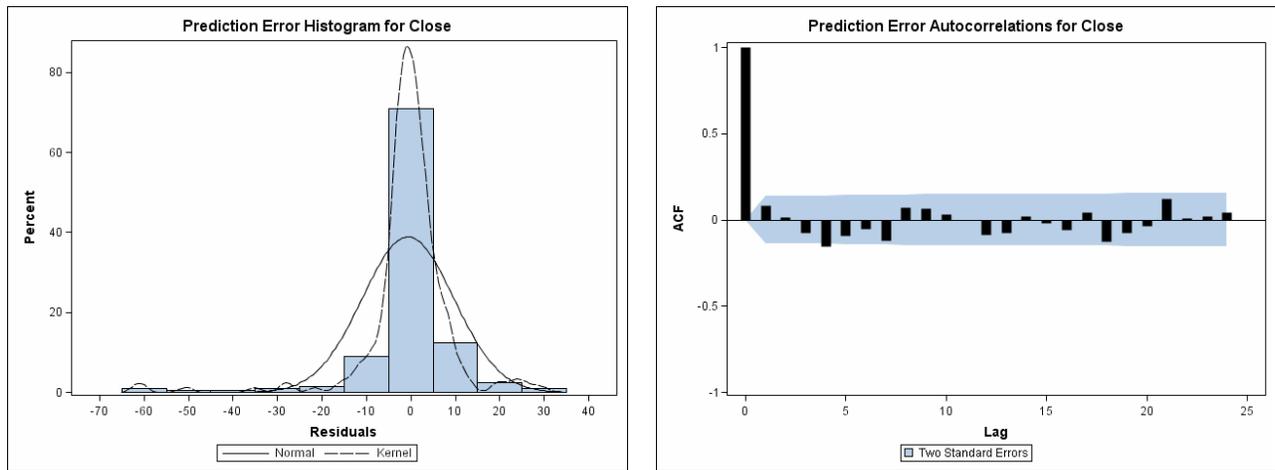


Figure 6: Residual Diagnostics

Figure 6 shows the diagnostic plots based on the one-step-ahead residuals. The autocorrelation function (ACF) doesn't show any significant violations of the whiteness of the residuals. Therefore, on the whole, the model seems to fit the data well.

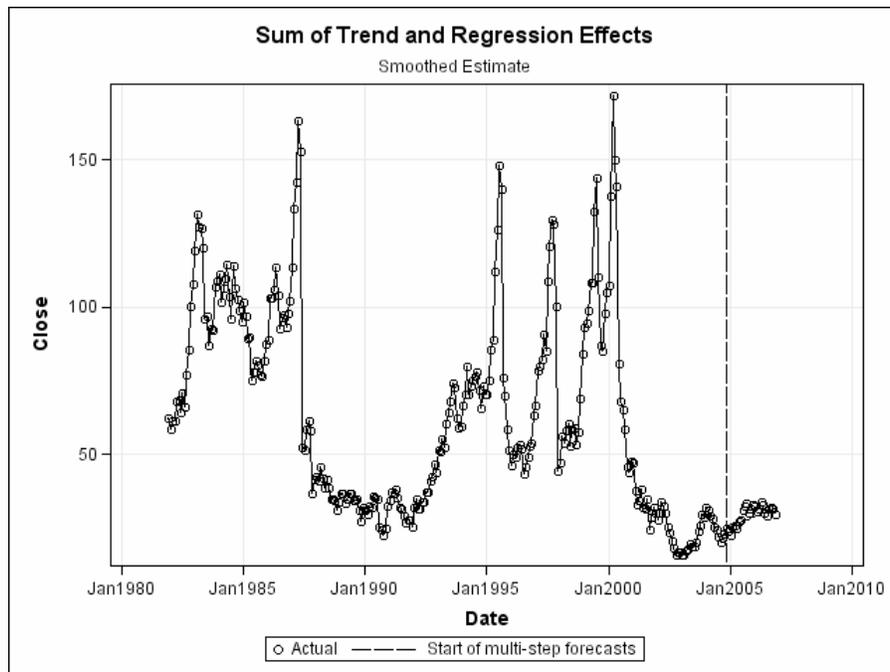


Figure 7: Plot of Smoothed Estimate

Figure 7 shows the forecast plot and demonstrates the model predictions were quite good.

CONCLUSIONS

By using the NYSE stock market data for Texas Instruments, the advanced features of three important SAS/ETS PROCs (TIMESERIES, FORECASTING, and UCM) were demonstrated. The graphical representation of the PROC TIMESERIES procedure is a significant tool which facilitates clear understanding of the underlying time series data and also helps in reliable forecasting. The PROC FORECASTING procedure provides a one-step method to automatically generate forecasts for hundreds of time series at a time. The PROC UCM procedure is a new PROC which is very useful to find a suitable model for the

series of interest, to obtain extensive model diagnostics, to generate series forecasts, and to forecast the constituent components. The comparison of these three procedures indicates that PROC UCM is better suited to analyze the stock price fluctuations.

REFERENCES

SAS Institute, Inc. (2008), TEMPLATE Procedure: ODS Graphics Overview

ODS: <http://support.sas.com/documentation/cdl/en/grstatproc/60786/HTML/default/a003159568.htm>

SAS Institute, Inc. (2008), SAS/ETS(R) 9.2 User's Guide TIMESERIES Procedure Functional Summary, SAS OnlineDoc® 9.2, Cary, NC: SAS Institute, Inc.

TIMESERIES: http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/timeseries_toc.htm

SAS Institute, Inc. (2008), SAS/ETS(R) 9.2 User's Guide FORECAST Procedure Functional Summary, SAS OnlineDoc® 9.2, Cary, NC: SAS Institute, Inc.

FORECAST: http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/forecast_toc.htm

SAS Institute, Inc. (2008), SAS/ETS(R) 9.2 User's Guide UCM Procedure Functional Summary, SAS OnlineDoc® 9.2, Cary, NC: SAS Institute, Inc.

UCM: http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/ucm_toc.htm

CONTACT INFORMATION

Your comments are greatly appreciated and encouraged. Contact the authors at:

Somaye Gharibvand
Multimedia University (MMU)
Jalan Multimedia 63100
Cyberjaya Selangor Malaysia
Work Phone: 60176703817
Email: s_gharibvand@yahoo.com

Lida Gharibvand
University of California, Riverside
900 University Ave.
Riverside, CA 92521 USA
Work Phone: (949) 230-5439
Email: lida.gharibvand@email.ucr.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.