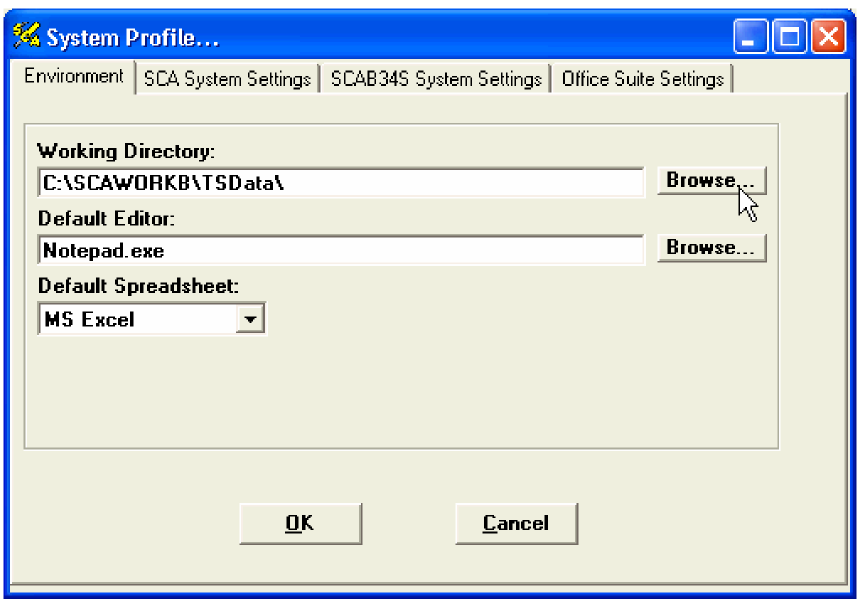
Introduction to SCA WorkBench 5.0
Interactive Use of the SCA Statistical System

SCA WorkBench provides a graphical user interface (GUI) for the SCA Statistical System. Under this GUI, users can

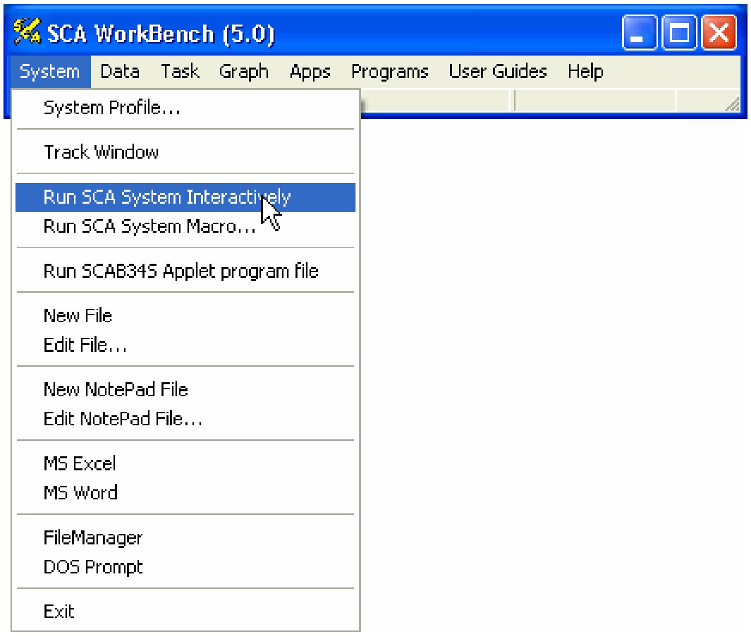
- Employ easy-to-use menus and dialogs supporting all SCA commands
- Type SCA commands directly at the SCA Input Console
- Organize and execute SCA macro procedures
- Import data from Excel and other spreadsheet programs

This document provides an example of using the menus and dialogs to perform interactive analysis in the SCA Statistical System. After launching SCA WorkBench, it is important to select the working directory. The working directory will contain the output and saved files from an SCA session. It also typically contains the data files associated with the analysis project. The working directory is set by selecting **System Profile** under the **System** menu.

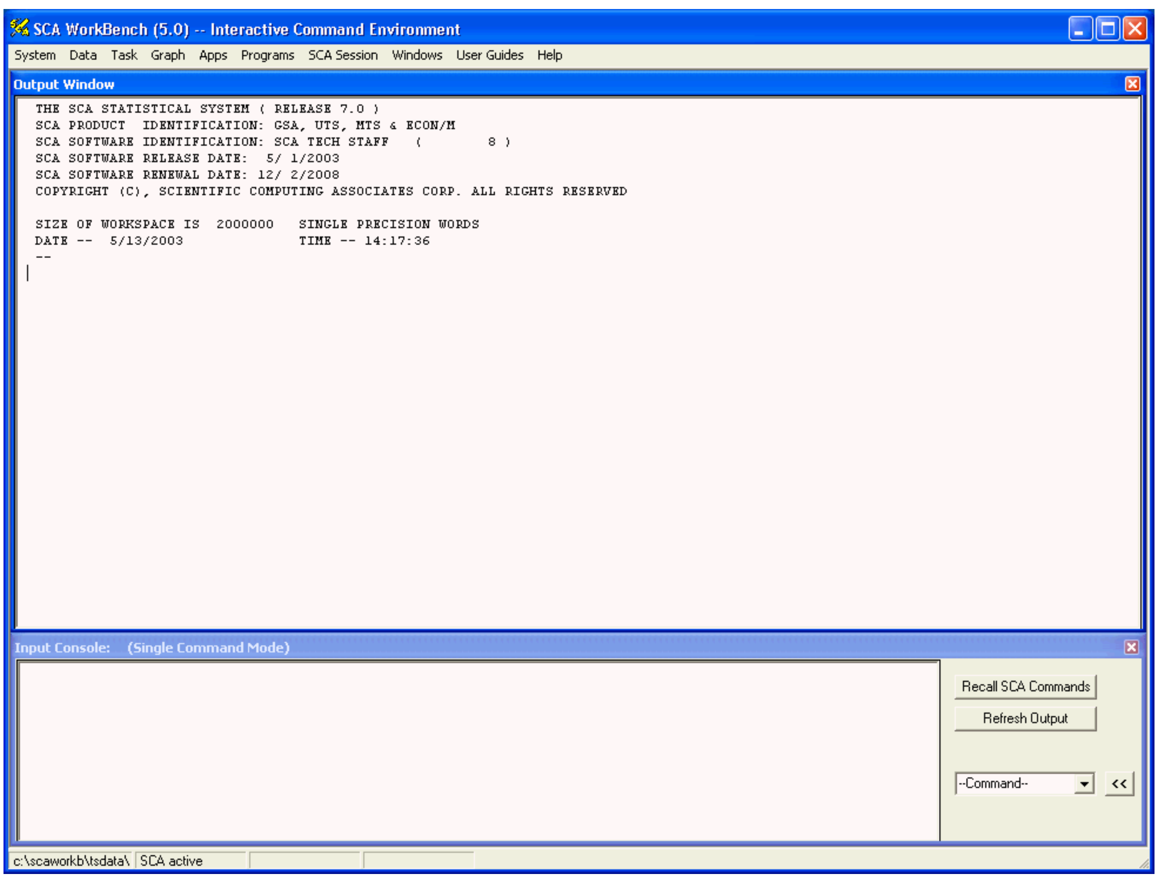


The working directory can be typed directly in the text box or it can be selected by clicking on the *Browse* button. For this illustrative example, the TSDATA directory is selected as the working directory. This directory contains a variety of time series data and is automatically installed by SCA WorkBench under the SCAWORKB directory.

Once the working directory is specified, an interactive SCA System session is started by selecting **Run SCA System Interactively** from the **System** menu.

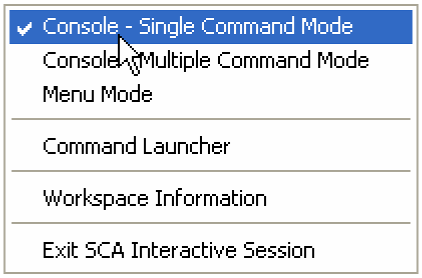


The initial environment of an interactive SCA session is displayed below.

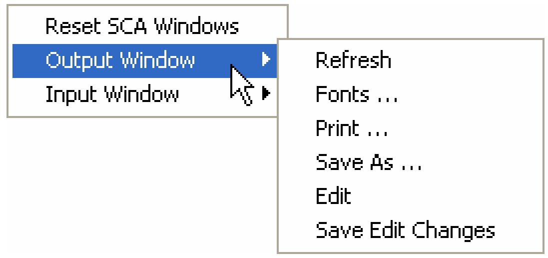


The items under the **SCA Session** and **Windows** topics are displayed below.

SCA Session



Windows



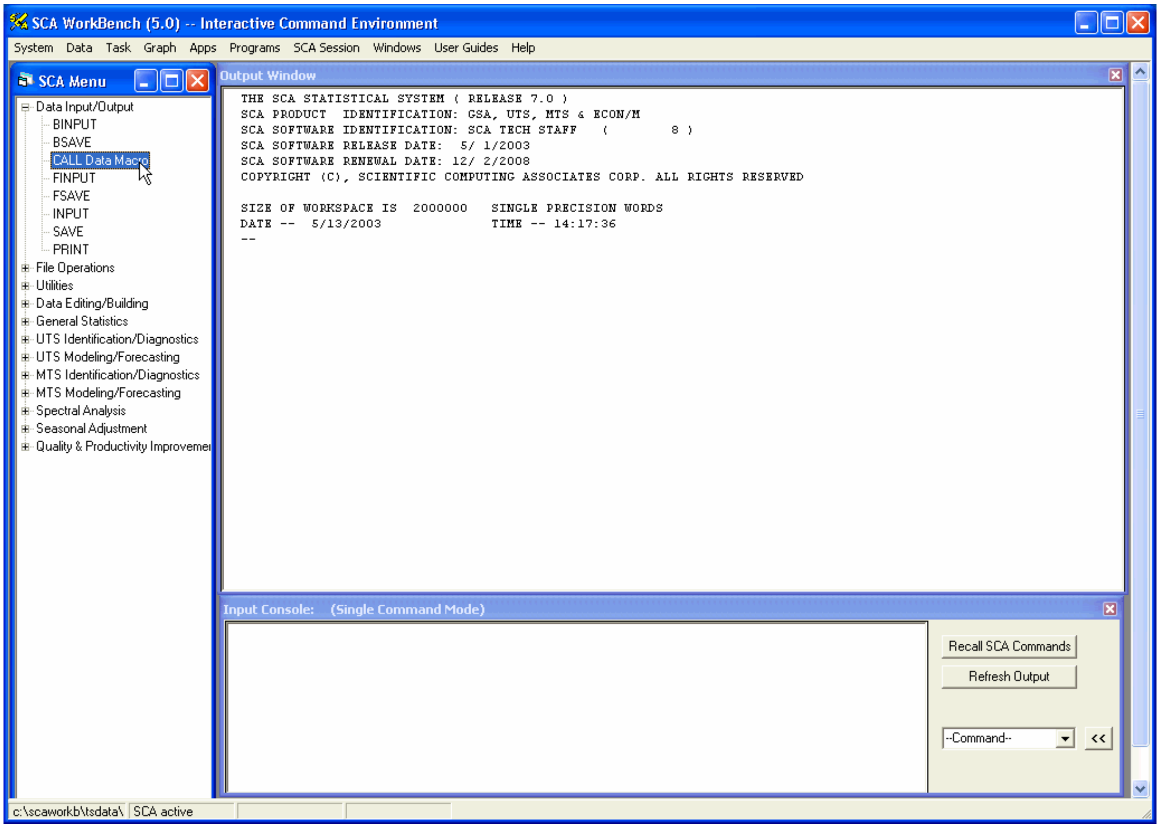
From the **SCA Session** menu, users can select a preferred interface mode, view a sorted list of SCA commands, and display information about variables currently stored in the SCA System workspace. The **Windows** menu provides control over the input console window and output window. From this menu, users can change fonts, add comments to the output, or print/save the output and SCA command history.

If a user is familiar with the SCA command syntax, SCA commands can be typed directly at the input console. This is often an efficient mode of operation when executing simple SCA commands. For example, the SCA command to compute the sample autocorrelation function of a time series (assuming all default options) is

```
ACF variable-name
```

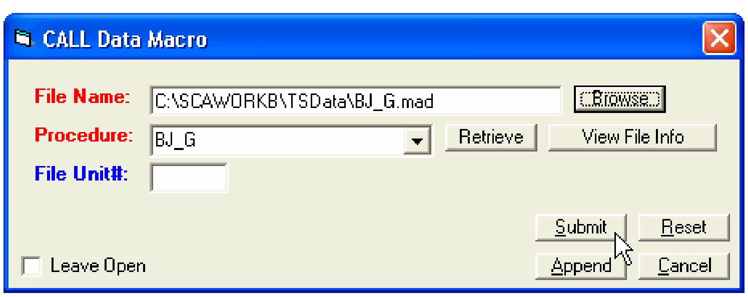
which consists of a two word command. By typing the command, it eliminates the need to navigate through menus. However, if a user is not familiar with the SCA command syntax or if the command has several options, typing the commands at the input console may become less desirable. In this circumstance, it may be easier for a user to select options from a dialog.

SCA WorkBench provides a GUI interface that consists of menus and dialogs for all SCA commands. Currently, SCA provides more than 100 individual commands that cover a wide range of topics from data and workspace management to statistical modeling and analysis. To view the dialogs by topic, select **Menu Mode** from the **SCA Session** menu. This is the recommended mode of operation for new users.



The items under each topic correspond to the SCA command name with few exceptions. A longer description of the command is displayed by clicking once on the item. The dialog box is launched by double-clicking the item.

When an SCA session is started, the first action is typically reading data into the workspace. The SCA System can read data into its workspace using the INPUT, FINPUT or BINPUT commands. Alternatively, the data can be read from an SCA data macro. In this example, we select the **CALL Data Macro** dialog which builds the command to read data into the workspace via an SCA data macro.

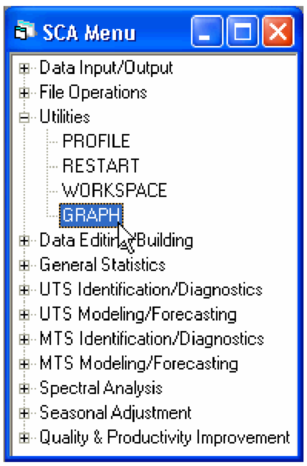


The items in red signify required elements of the command. The items in blue or black indicate that the element is optional.

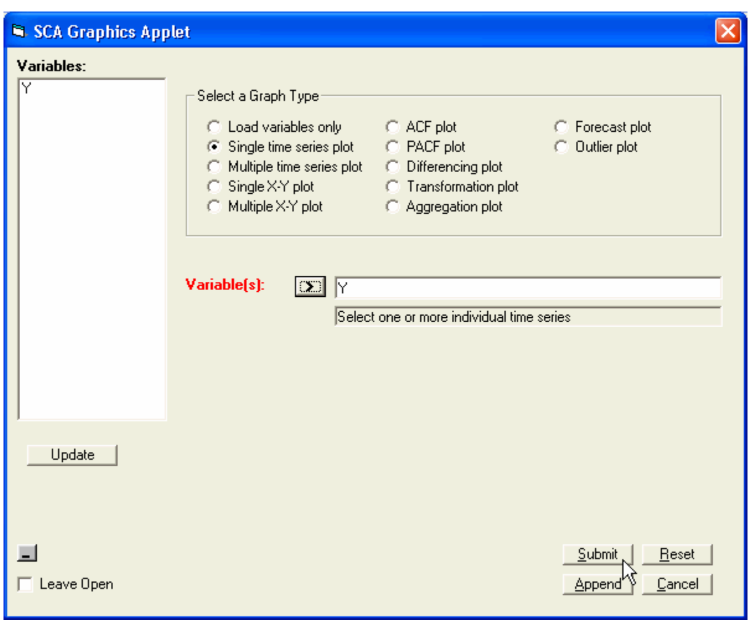
The *Browse* button is used to select the SCA data macro file to read. Here, we select the "BJ_G.mad" file located under the TSData sub-directory of SCAWORKB. After the file is selected, the procedures contained in the data macro are automatically loaded into the *Procedure* drop-down box. If the SCA data macro file is typed directly into the File Name text box, the *Retrieve* button must be executed to retrieve the individual data macro procedures into the *Procedure* drop-down box. Click on *Submit* to execute the command immediately in the SCA System or *Append* to copy the SCA command syntax to the input console window for execution later. In this example, we click on *Submit* which in turn builds the SCA command

```
CALL BJ_G.  FILE 'BJ_G.mad'.
```

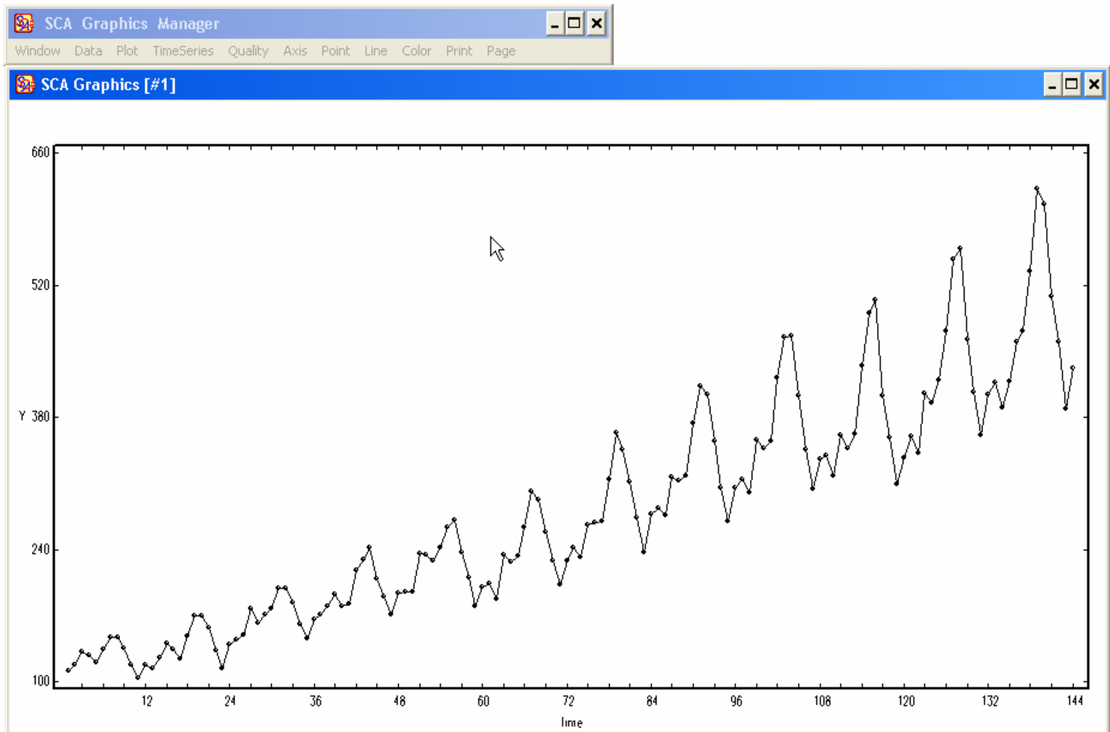
and sends it to the SCA System for execution. The SCA command is echoed in the output window and is followed by the message that the variable Y is stored in the workspace and is 144 observations. The BJ_G series (referred to as Y) is the monthly airline passenger data from Box-Jenkins (1970). The SCAGRAF applet is used to plot the Y variable and investigate characteristics of the data. Under the **Utility** menu, double-click on **GRAPH**.



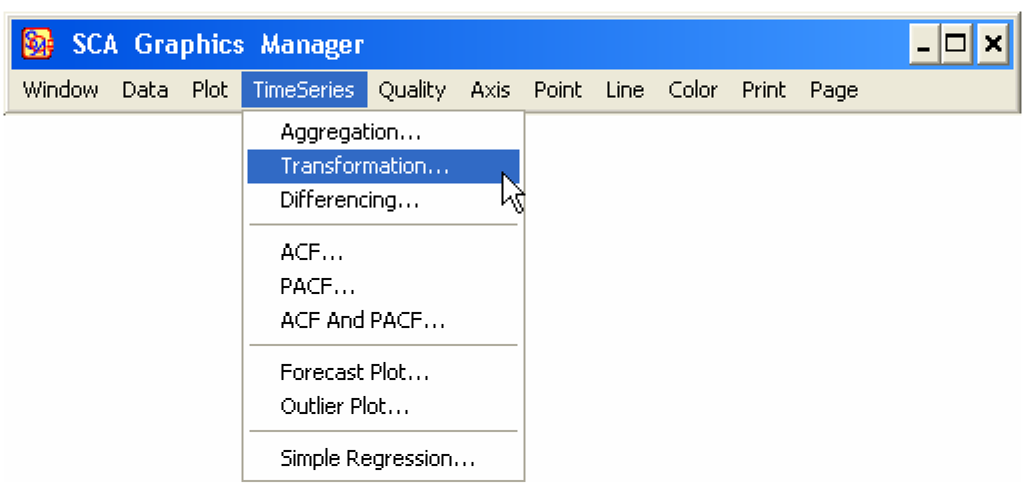
The following dialog is displayed after the **GRAPH** item is selected.



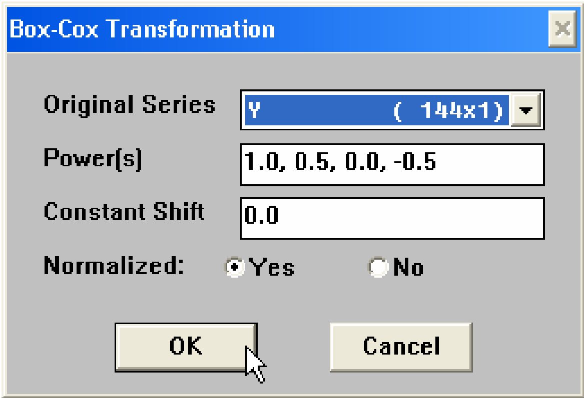
Here, a high resolution time series plot of Y is selected. The variable is exported from the SCA workspace into the SCA Graphics Manager where the time series plot is displayed.



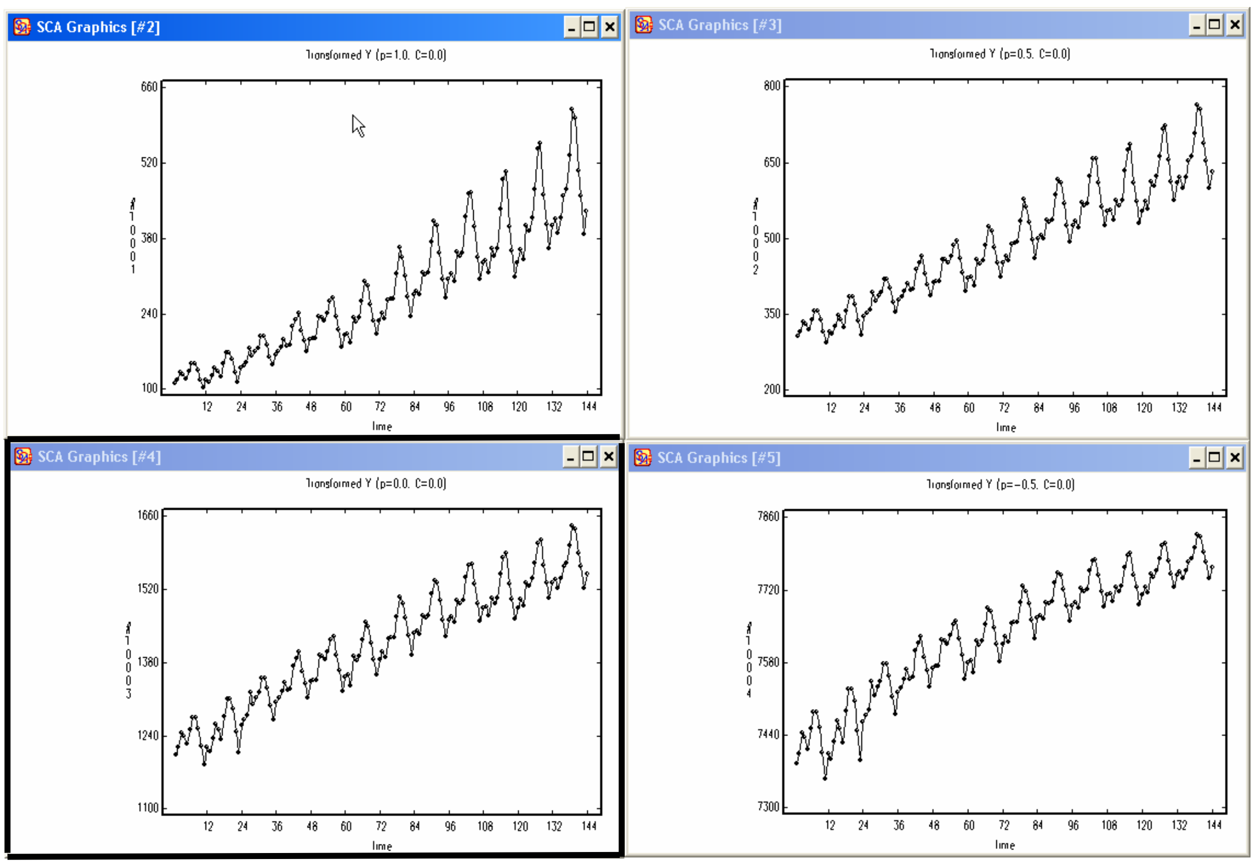
The time series plot reveals a strong monthly seasonal pattern. It is also evident that the series is not stationary and the variance of the series increases over time. We investigate if a transformation will stabilize the variance of the series. From the **SCA Graphics Manager**, select **Transformation** from the **TimeSeries** topic.



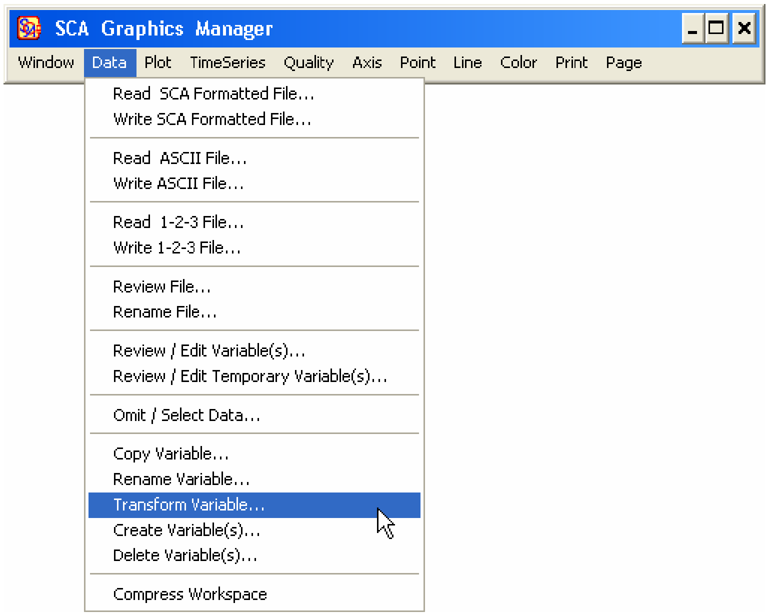
The Y variable is selected from the Box-Cox Transformation dialog and we investigate how the series would look using no transformation (1.0), a square root transformation (0.5), logarithmic transformation (0.0), and an inverse square root transformation (-0.5).



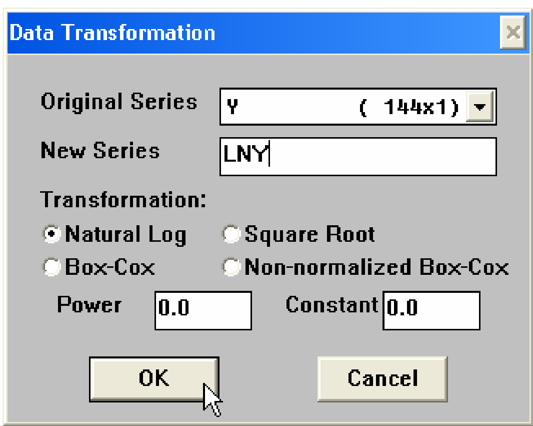
After selecting the appropriate choices, click on OK to generate the four transformation plots shown below.



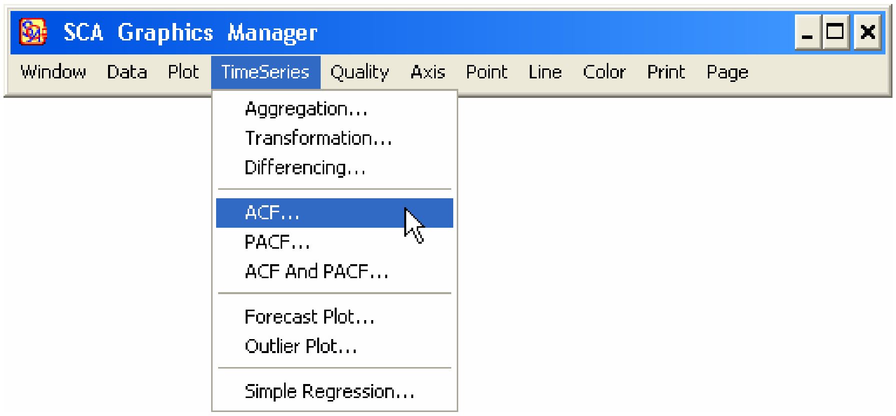
Upon visual inspection, it seems that a logarithmic transformation does an adequate job of stabilizing variance over time. The transformed series displayed in the graphs are currently temporary variables in SCAGRAF. Since we are interested in furthering our analysis of the logarithmic transformed series in SCAGRAF, we select **Transform Variable** under the **Data** topic.



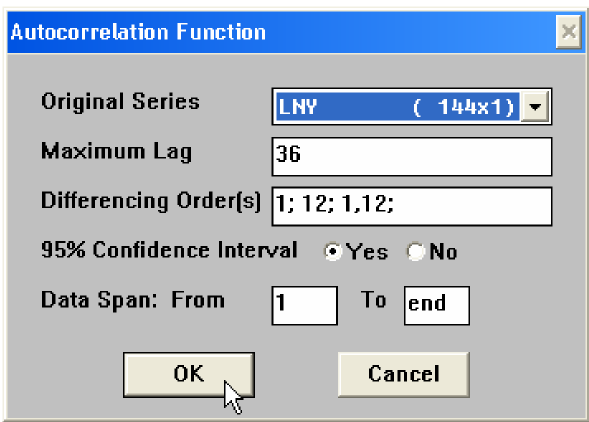
This action will display the **Data Transformation** dialog where the following options are selected and where LNY is specified as the new series created.



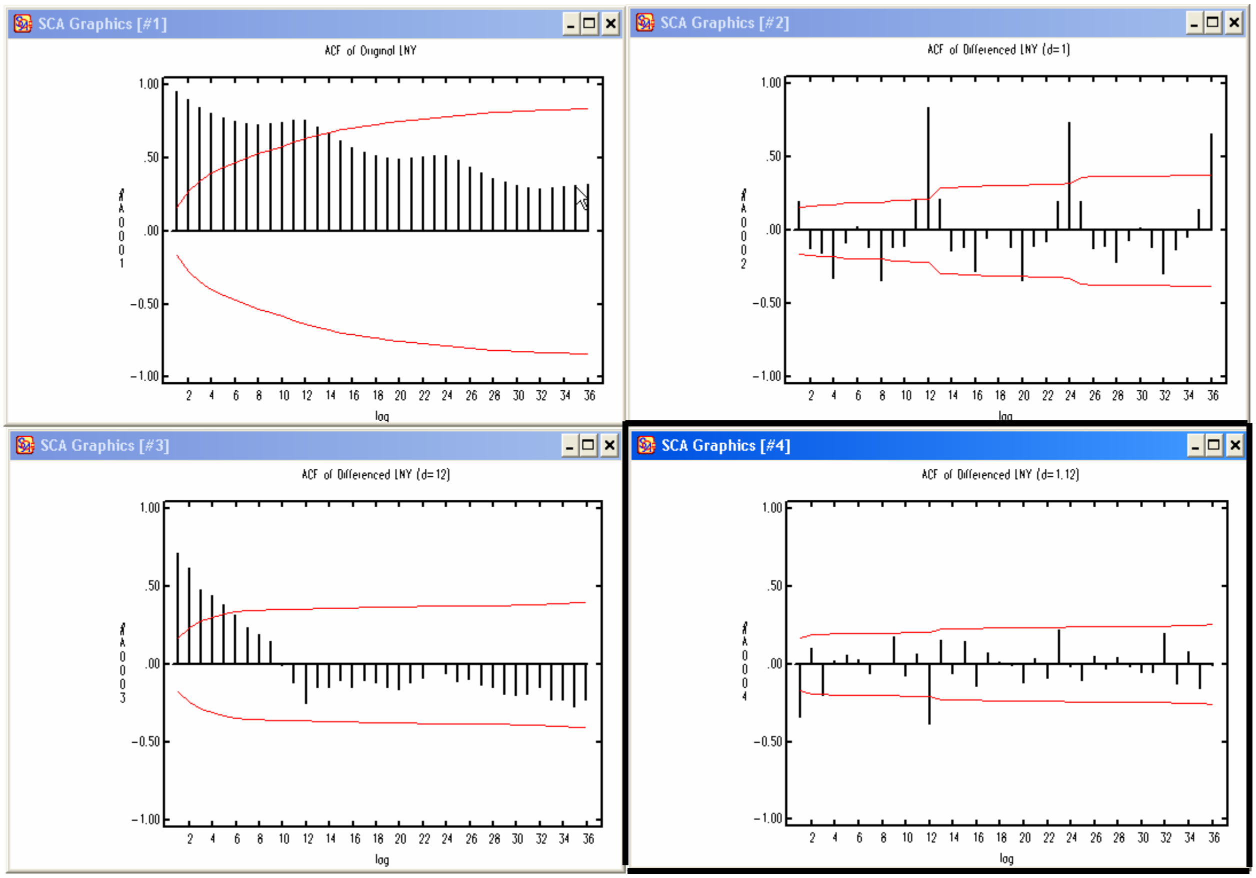
Since the transformed series LNY remains a non-stationary process, differencing should be considered to induce stationary. To explore the differencing orders that adequately induce stationarity, the autocorrelation function of the LNY series is examined.



After selecting **ACF** from the **TimeSeries** topic, the following dialog box will appear allowing multiple differencing orders to be explored in a convenient manner.



Here, the LNY series is selected as the original series and we enter the differencing orders that are to be considered and separate them by a semi-colon. The following ACF plots are produced.



Based on the ACF graphs shown above, double differencing (d=1,12) is required to induce stationarity. We can also determine from the ACF of the double-differenced series (d=1,12) that a seasonal ARIMA model of the form

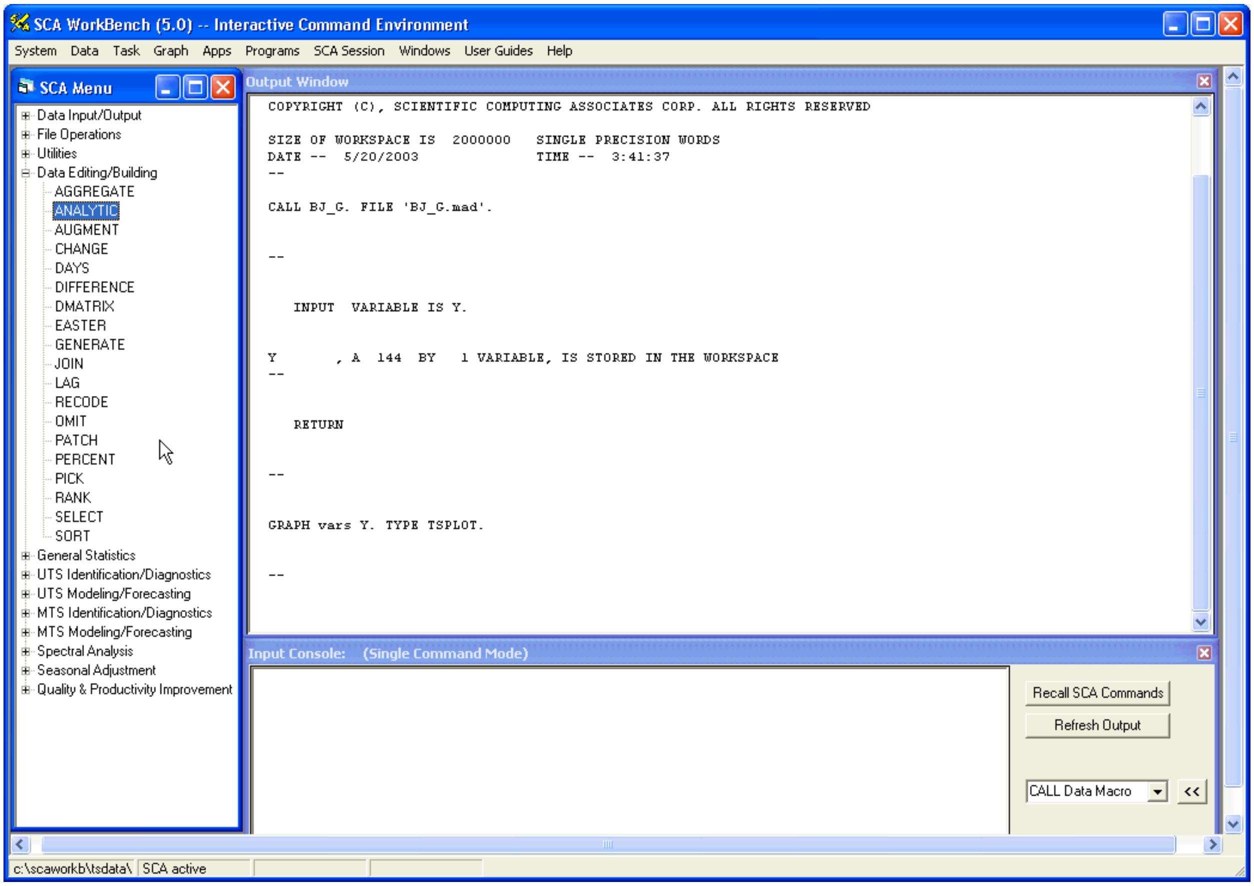
$$(1-B)(1-B^{12})LNY = (1-\theta_1 B)(1-\Theta_{12}B^{12})a_t.$$

may be a reasonable model to consider. We now have enough information to go back to the interactive SCA System session to specify our proposed model and examine its adequacy.

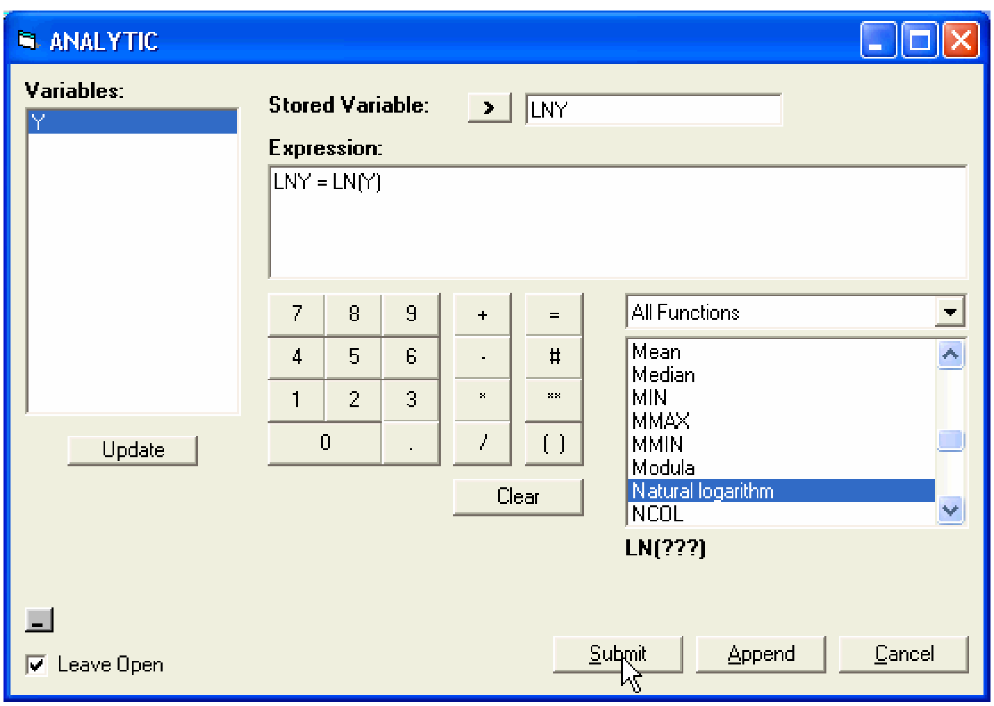
It is recommended that the SCA Graphics Manager session be terminated before returning to the SCA session if the user intends to employ the SCAGRAF dialog again.

The SCA System allows only one SCAGRAF session to be opened. If an SCAGRAF session is already open, the variables can not be passed from the SCA System to the SCA Graphics Manager. This is applicable to Windows NT, ME and XP operating systems. Under Windows 95/98 operating systems, this restriction does not apply.

The first task once we are back in the interactive SCA session, is to take a natural log of the Y variable. Select **Analytic** under the **Data Editing/Building** topic.

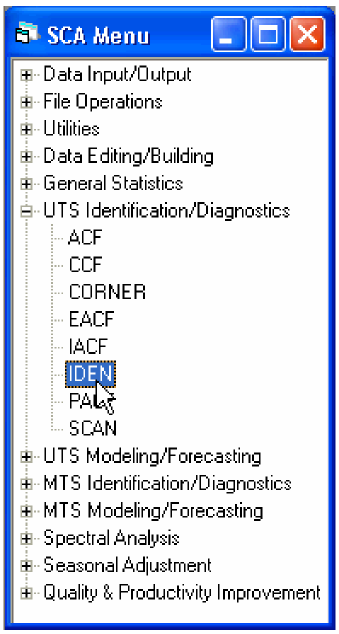


The **Analytic** dialog provides a calculator-like interface to specify SCA analytic functions or other assignment operations.

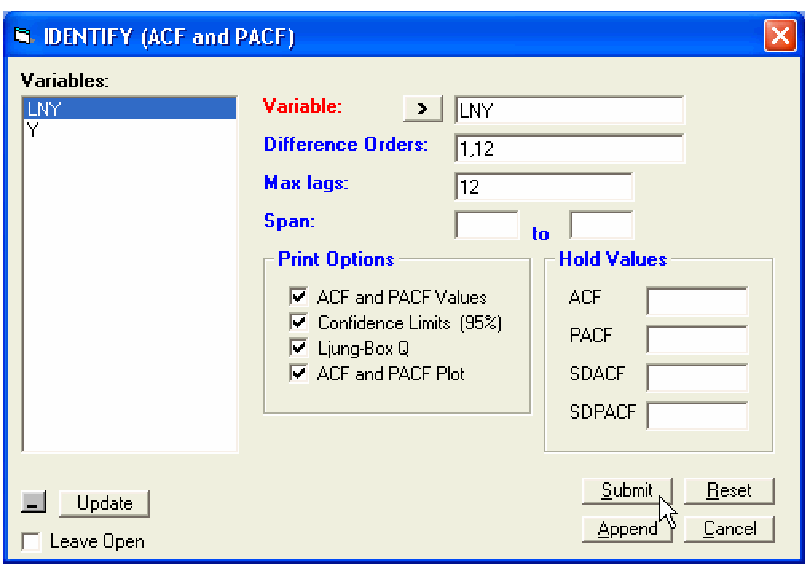


Type the name of the series that will contain the logarithmic transformed series as LNY in the *Stored Variable* text box. Next, double-click on natural logarithm from the functions list. Three question marks will be enclosed in parentheses. This indicates that a variable needs to be selected. Double-click on Y in the variables list to complete the expression as shown above. Click on *Submit* to execute the operation in the SCA System. Note that we could have simply typed the expression LNY=LN(Y) directly at the input console to achieve the same result.

We now select **IDEN** from the **UTS Identification/Diagnostics** topic to compute both the autocorrelation and partial autocorrelation function for the LNY series.



That action will display the IDEN dialog box as shown below.



Select the LNY variable from the *variables list* and set the differencing orders to “1, 12” to examine the ACF and PACF of the double differenced series. Note that the maximum number of lags is specified as 12 for brevity. Typically, a user should set the maximum number of lags so at least three or four seasonal cycles can be viewed. For monthly data as we have here, we should specify 36-48 lags to view as we have done in SCAGRAF.

Click on *Submit* to execute the command.

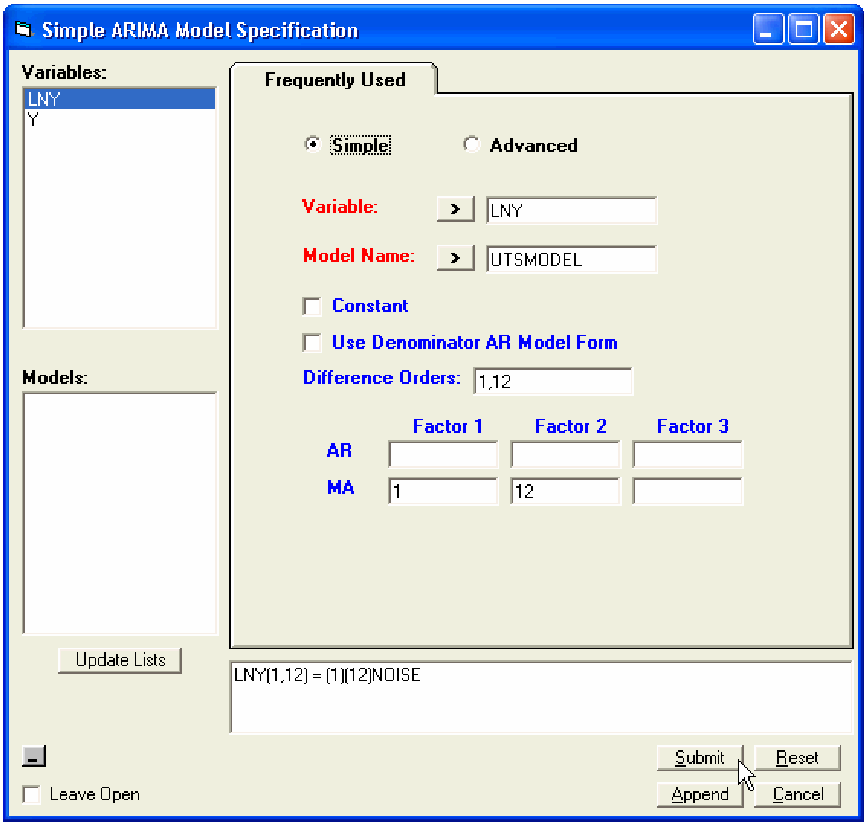
Output Window

```
                                                  1      12
DIFFERENCE ORDERS. . . . . . . . . . . (1-B  ) (1-B  )
NAME OF THE SERIES . . . . . . . . . .             LNY
TIME PERIOD ANALYZED . . . . . . . . .   1  TO    144
MEAN OF THE (DIFFERENCED) SERIES . . .         0.0003
STANDARD DEVIATION OF THE SERIES . . .         0.0457
T-VALUE OF MEAN (AGAINST ZERO) . . . .         0.0729


AUTOCORRELATIONS

  1- 12     -.34  .11 -.20  .02  .06  .03 -.06 -.00  .18 -.08  .06 -.39
  ST.E.      .09  .10  .10  .10  .10  .10  .10  .10  .10  .10  .10  .10
   Q        15.6 17.1 22.6 22.7 23.1 23.3 23.7 23.7 28.1 29.0 29.6 51.5


          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
                                     I
  1  -0.34                  XXXXX+XXXI   +
  2   0.11                      +    IXXX +
  3  -0.20                      XXXXXI    +
  4   0.02                      +    IX   +
  5   0.06                      +    IX   +
  6   0.03                      +    IX   +
  7  -0.06                      +   XI    +
  8   0.00                      +    I    +
  9   0.18                      +    IXXXX+
 10  -0.08                      +  XXI    +
 11   0.06                      +    IXX  +
 12  -0.39                 XXXXX+XXXXI    +
```

The ACF plot is shown in the output window. The PACF plot is not visible here but can be viewed by using the scroll bar. Based on the t-value of the mean (against zero), we can determine that a constant or trend term is not required in our model. Furthermore, the ACF plot supports our identification of the seasonal ARIMA model

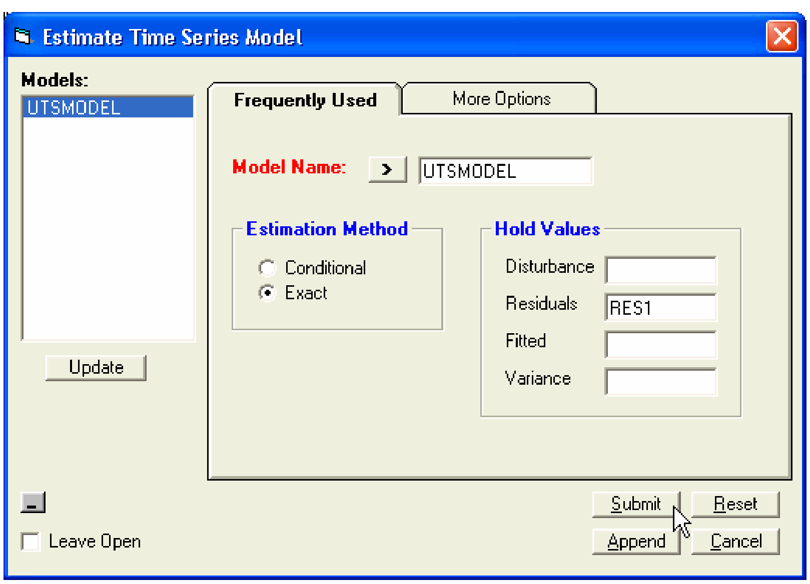
$$(1-B)(1-B^{12})LNY = (1-\theta_1 B)(1-\Theta_{12}B^{12})a_t.$$

To specify the univariate time series model, select **ARIMA** under the **UTS Modeling/Forecasting** topic.

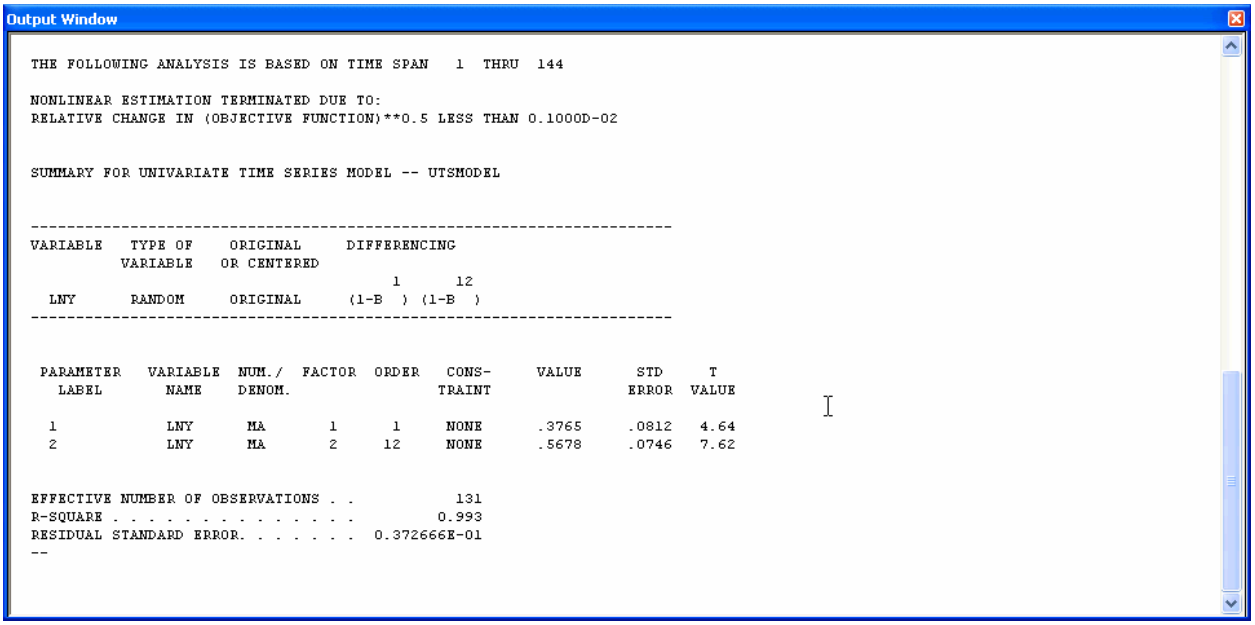


Select LNY from the variables list, and enter the model components as shown. The syntax for the specified model is automatically constructed and displayed in the model text box. Click on Submit to save the model in the SCA System. The specified model will then be stored under the name UTSMODEL in the SCA workspace.

It is now necessary to estimate the UTSMODEL. Select **ESTIM** from the **UTS Modeling/Forecasting** topic.

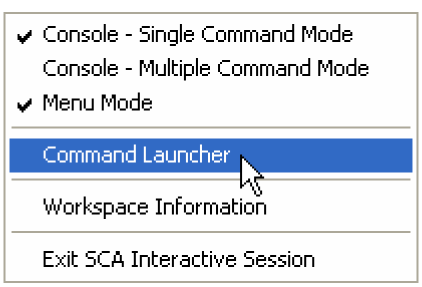


Select UTSMODEL from the models list. If more than one model has been specified, the other model names would also appear in the models list. Since the UTSMODEL contains moving average terms and the LNY series is not long, select the EXACT estimation method. Hold the residuals of the estimated model in a new variable named RES1.

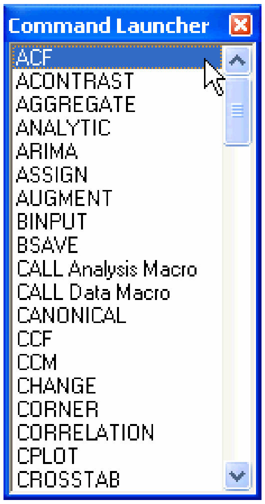


The command is echoed in the output window along with the estimation results as shown above.

Once a user becomes familiar with SCA commands, the command dialogs can also be launched from a sorted quick list. There are two places where a user can access the dialog quick list. The first place is a dropdown list on the **Input Console**. Simply click on the down arrow to expose the list or type the first few letters of the command in the dropdown list until it matches the command you need. Alternatively, if the user prefers to have a floating list of SCA commands available, select **Command Launcher** from the **SCA Session** menu.

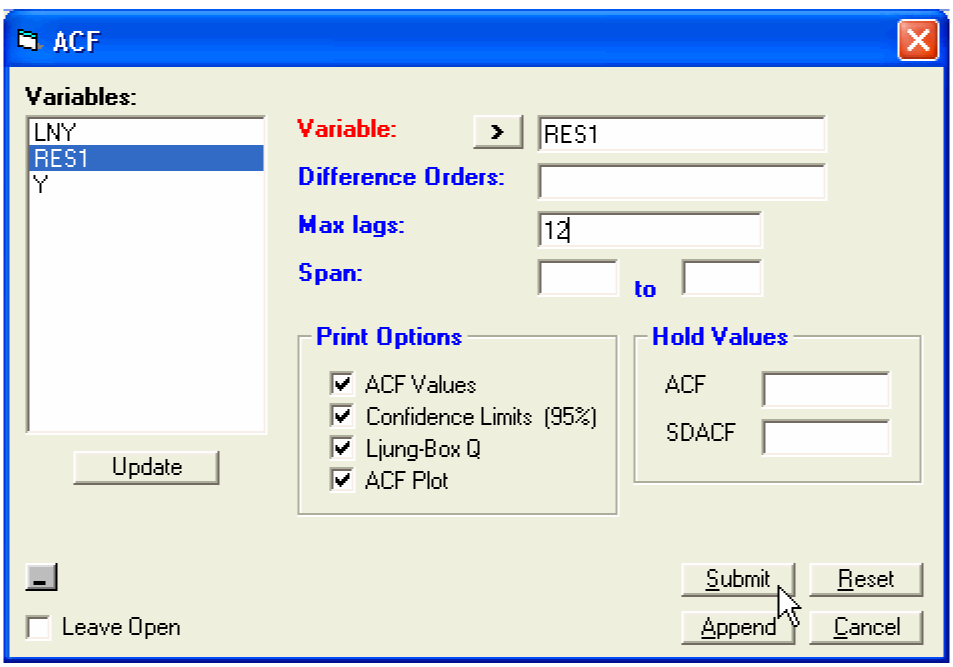


The **Command Launcher** displays the SCA command list on your desktop in a floating menu as shown below.



The same dialogs are accessible from the tree list, the drop-down command list, and the command launcher menu. It is simply a matter of preference as to how the user wishes to select a command dialog.

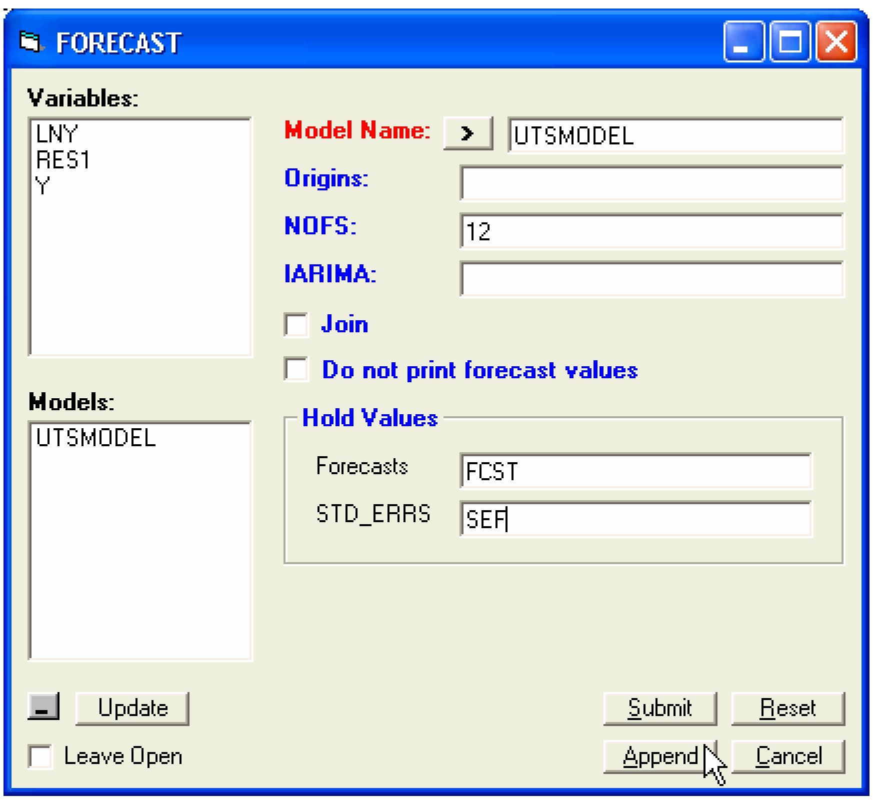
To complete our sample analysis, it is prudent to check the adequacy of the UTSMODEL specified. Select ACF from the Command Launcher to access the dialog for computing the autocorrelation function.



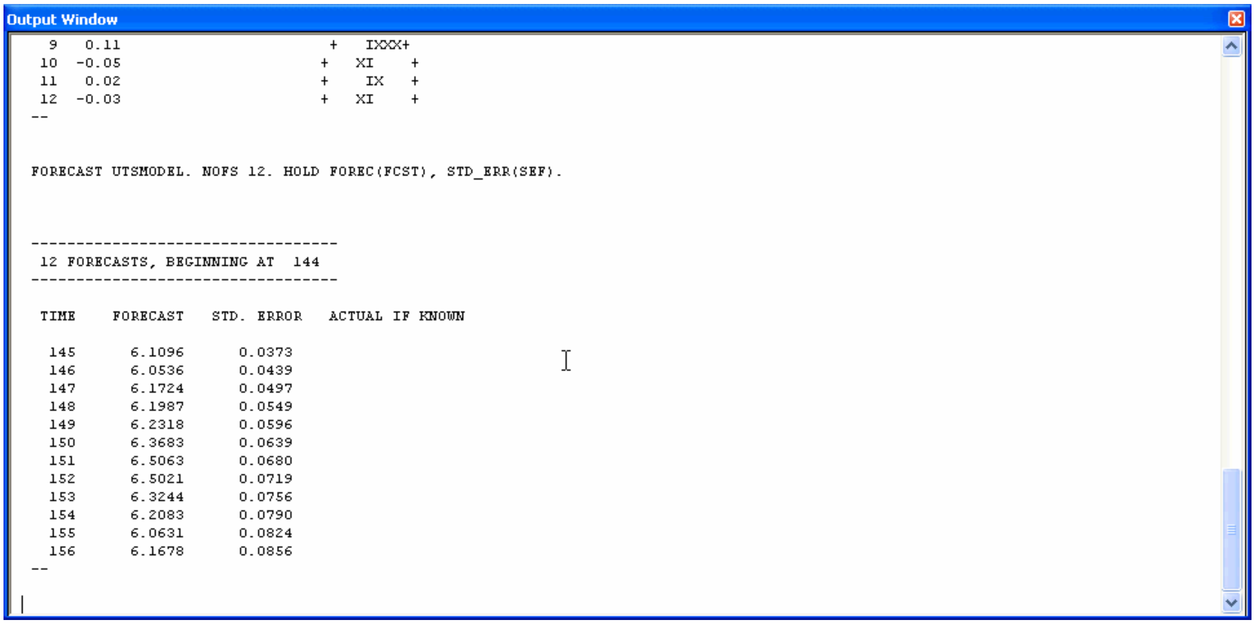
The residuals from the estimated UTSMODEL were stored in RES1. Select RES1 from the variables list and click on Submit to execute the command. The results are shown below.

```
Output Window
ACF RES1. MAXLAG 12.


NAME OF THE SERIES . . . . . . . . . .           RES1
TIME PERIOD ANALYZED . . . . . . . . . 14   TO    144
MEAN OF THE (DIFFERENCED) SERIES . . .         0.0020
STANDARD DEVIATION OF THE SERIES . . .         0.0372
T-VALUE OF MEAN (AGAINST ZERO) . . . .         0.6072


AUTOCORRELATIONS

  1- 12      .01  .03 -.12 -.10  .08  .08 -.05 -.02  .11 -.05  .02 -.03
  ST.E.      .09  .09  .09  .09  .09  .09  .09  .09  .09  .09  .09  .09
   Q          .0   .1  2.1  3.6  4.4  5.2  5.6  5.7  7.5  7.9  7.9  8.1


          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
                                     I
  1   0.01                       +   I   +
  2   0.03                       +   IX  +
  3  -0.12                       +XXXI   +
  4  -0.10                       +XXXI   +
  5   0.08                       +   IXX +
  6   0.08                       +   IXX +
  7  -0.05                       +  XI   +
  8  -0.02                       +  XI   +
  9   0.11                       +   IXXX+
 10  -0.05                      +   XI    +
 11   0.02                      +    IX   +
 12  -0.03                      +   XI    +
```

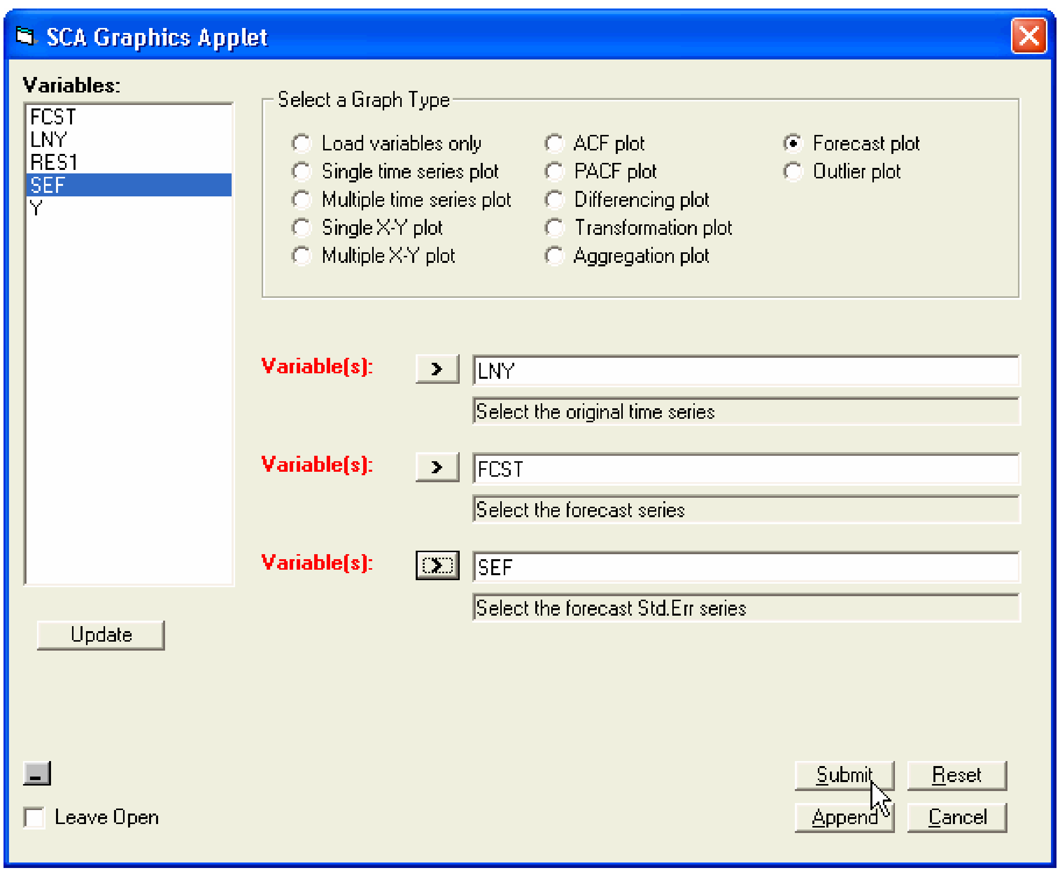
Examining the autocorrelations, we conclude that the UTSMODEL adequately represents the LNY series and that serial correlation in the residuals has been removed. Finally, the UTSMODEL is used for forecasting. Select **FORECAST** from the **Command Launcher** menu.



Select the UTSMODEL and hold the forecasts and standard errors of the forecasts in the new variables FCST and SEF. This will produce the forecasts shown below.

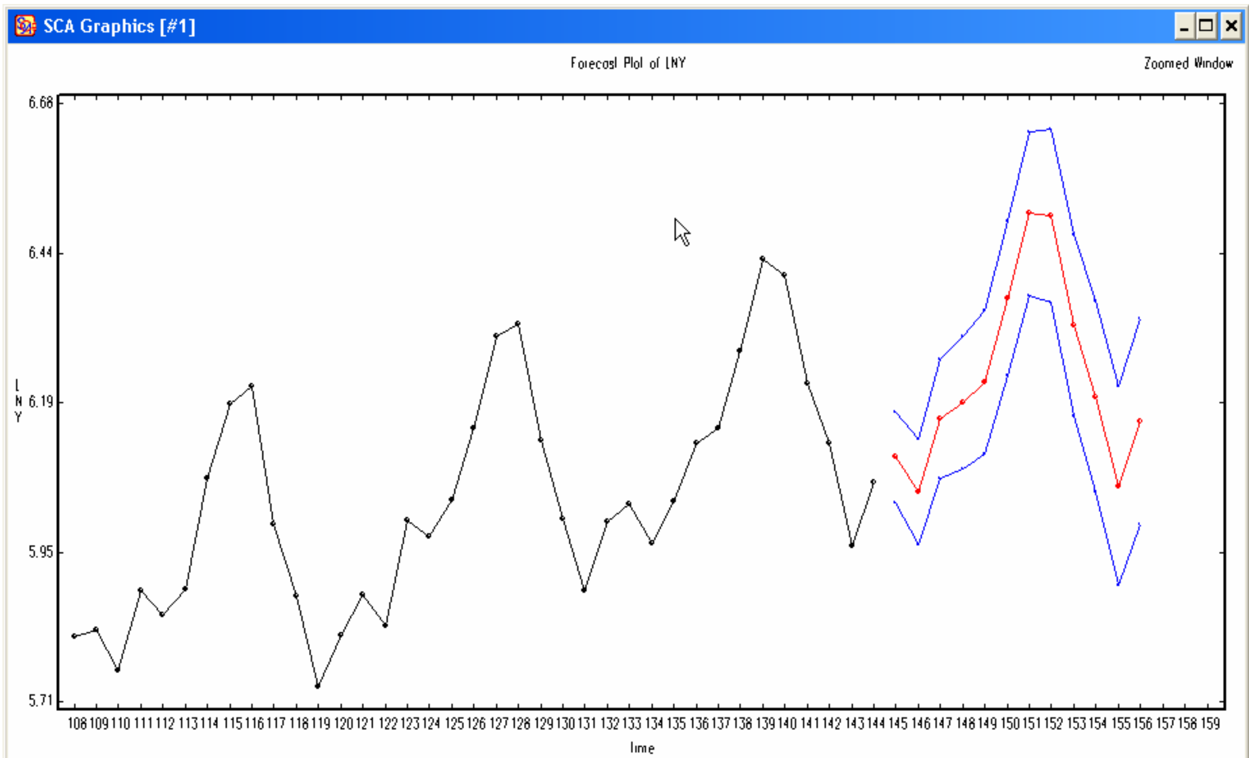


A graph of the forecasts can be view by going to the GRAPH dialog and entering the options are shown below.

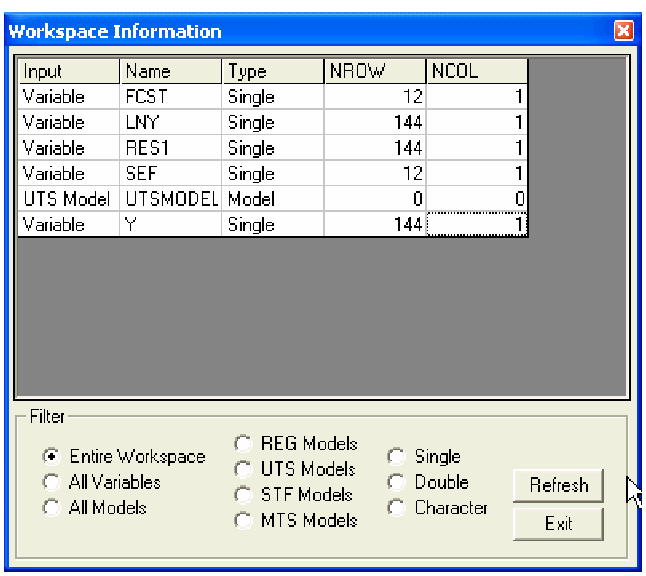


Upon submitting the GRAPH dialog, the forecast plot is generated.

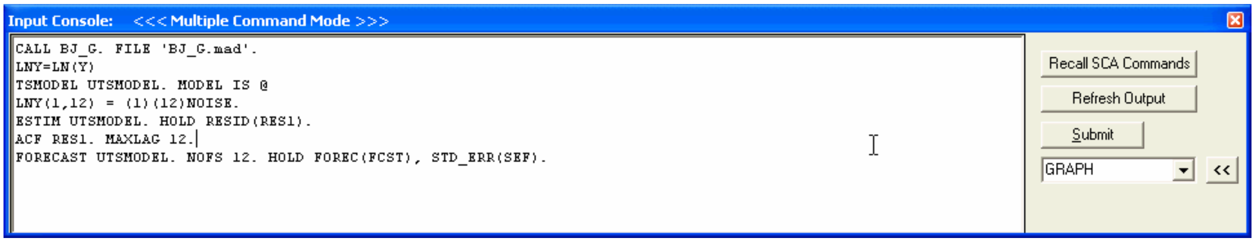
Please be sure the SCA Graphics Manager is not active on your desktop before submitting the GRAPH dialog or the graph will not be generated.



At any time during an interactive SCA session, the information contained within the SCA workspace can be viewed by selecting **Workspace Information** under the **SCA Session** menu.



The command dialogs also have memory of the last command executed making it convenient to submit a command then go back and modify the options without too much fuss. Users may also find it convenient to know that the command syntax, generated from a dialog or typed in directly at the Input Console can be recalled or saved to an external file. The individual commands may be recalled by pressing the "*CTRL* and +" or "*CTRL* and –" keys at the same time while your cursor is in the input console window. The entire history of commands can be recalled by clicking on the **Recall SCA Commands** button located on the **Input Console**.



After the **Recall SCA Commands** button is executed, the **Input Console** will be placed into **Multiple Command Mode**. This mode allows the user to edit more than one command at the console before submitting the group of commands to the SCA System for execution. The output window will be refreshed after all SCA commands have been processed by the SCA System.

A user may also copy the command history into a text editor (e.g., notepad) and convert the SCA command history into an SCA macro procedure which can be executed later to replicate an analysis.

Depending on the number of commands submitted as a group or the computational intensity of the commands executed, a message may be displayed asking whether you wish to continue waiting for the commands to finish. If you select No, the output window may need to be manually refreshed to display the output from the pending execution. Click on the *Refresh Output* button to manually refresh the output window. If you responded Yes (i.e., wait) then the output window will automatically refresh when the commands have been completely processed in the SCA System.