

Forecasting



Revised: 10/10/2017



Summary	2
Data Input.....	3
Analysis Options	5
Forecasting Models	9
Analysis Summary	21
Time Sequence Plot	23
Forecast Table	24
Forecast Plot.....	25
Model Comparisons	26
Residual Plots.....	27
Residual Autocorrelations.....	29
Residual Autocorrelation Function	30
Residual Partial Autocorrelations	31
Residual Partial Autocorrelation Function.....	32
Residual Periodogram Table.....	32
Residual Periodogram Plot	34
Residual Integrated Periodogram.....	35
Tests For Randomness	35
Residual Crosscorrelations.....	36
Residual Crosscorrelation Plot.....	38
Save Results	39
Calculations.....	40

Summary

The **Forecasting** procedure is designed to forecast future values of time series data. A *time series* consists of a set of sequential numeric data taken at equally spaced intervals, usually over a period of time or space. The models provided to forecast future values include a moving average, a random walk, various types of exponential smoothers, trend models, and parametric ARIMA models. Statistics are calculated to compare the fit of up to 5 models at any one time.

This procedure is designed for users who wish to select their own model. The *Automatic Forecasting* procedure tries various models and automatically selects the model that is best according to a specified goodness-of-fit criteria.

Sample StatFolio: *tsforecast.sgp*

Sample Data:

The file *golden gate.sgd* contains monthly traffic volumes on the Golden Gate Bridge in San Francisco for a period of $n = 168$ months from January, 1968 through December, 1981. The table below shows a partial list of the data from that file:

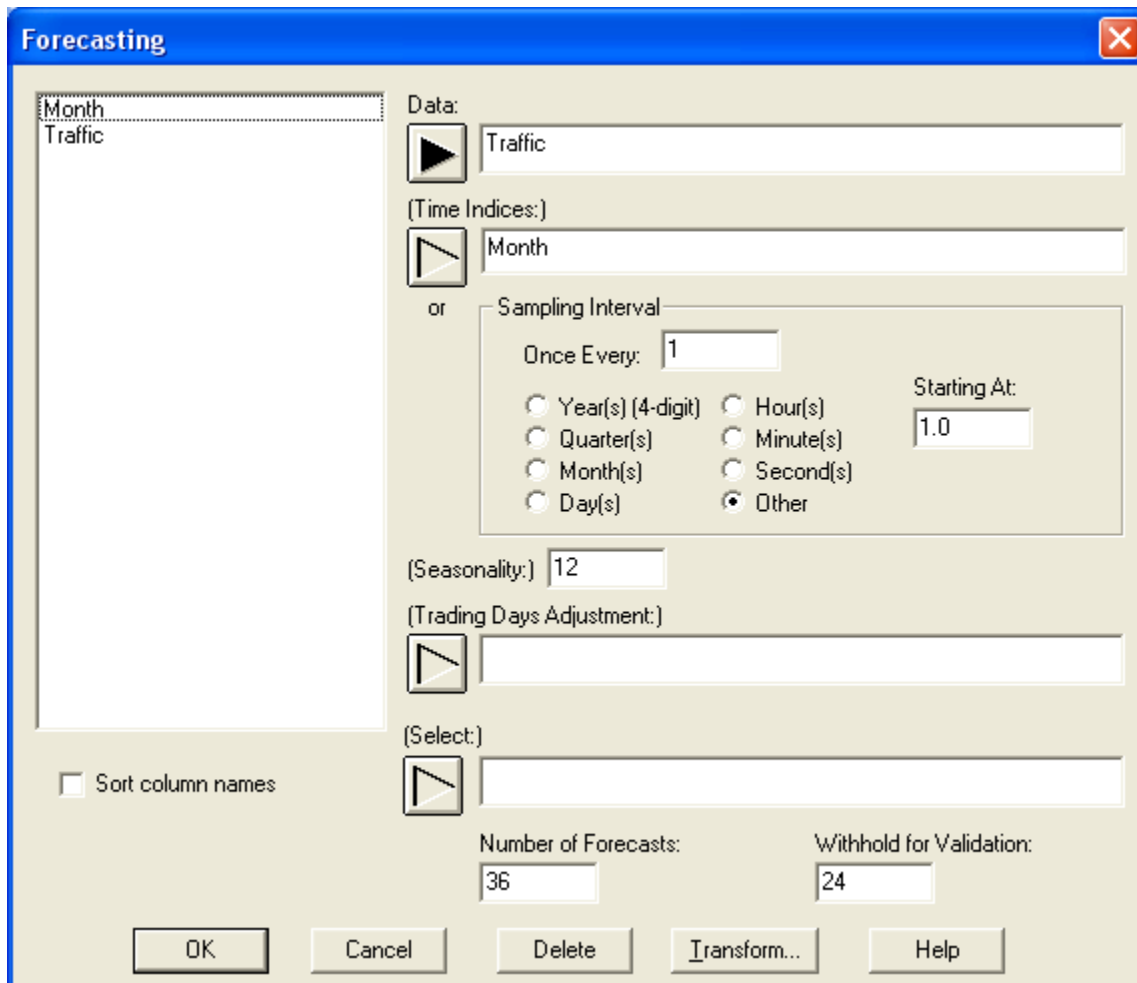
<i>Month</i>	<i>Traffic</i>
1/68	73.637
2/68	77.136
3/68	81.481
4/68	84.127
5/68	84.562
6/68	91.959
7/68	94.174
8/68	96.087
9/68	88.952
10/68	83.479
11/68	80.814
12/68	77.466
1/69	75.225
...	...

The data were obtained from a publication of the Golden Gate Bridge.

As an exercise, the data for the last two years (1980 and 1981) will not be used to estimate the forecasting model, but will be used instead to calculate validation statistics.

Data Input

The data input dialog box requests the name of the column containing the time series data:



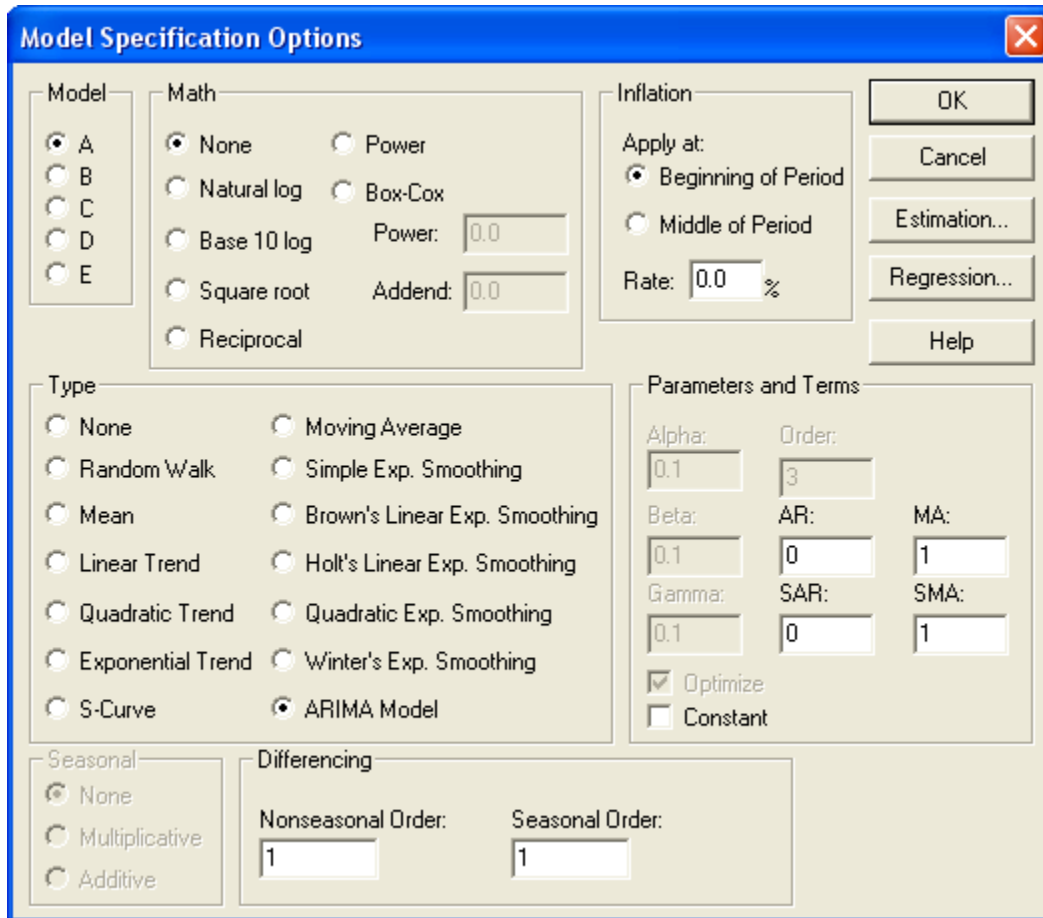
- **Data:** numeric column containing n equally spaced numeric observations.
- **Time indices:** time, date or other index associated with each observation. Each value in this column must be unique and arranged in ascending order.
- **Sampling Interval:** If time indices are not provided, this defines the interval between successive observations. For example, the data from the Golden Gate Bridge were collected once every *month*, beginning in January, 1968.
- **Seasonality:** the length of seasonality s , if any. The data is seasonal if there is a pattern that repeats at a fixed period. For example, monthly data such as traffic on the Golden Gate Bridge have a seasonality of $s = 12$. Hourly data that repeat every day have a seasonality of $s = 24$. If no entry is made, the data is assumed to be nonseasonal ($s = 1$).

- **Trading Days Adjustment:** a numeric variable with n observations used to normalize the original observations, such as the number of working days in a month. The observations in the *Data* column will be divided by these values before being plotted or analyzed. There must be enough entries in this column to cover both the observed data and the number of periods for which forecasts are requested.
- **Select:** subset selection.
- **Number of Forecasts:** number of periods following the end of the data for which forecasts are desired.
- **Withhold for Validation:** number of periods m at the end of the series to withhold for validation purposes. The data in those periods will not be used to estimate the forecasting model. However, statistics will be calculated describing how well the estimated model is able to forecast those observations.

In the current example, the traffic data is monthly beginning in January, 1968, and has a seasonality of $s = 12$. $m = 24$ observations at the end of the series will be withheld for validation purpose, while forecasts will be generated for the next 36 months.

Analysis Options

The *Forecasting* procedure is controlled by the *Analysis Options* dialog box:



- **Model:** the model to which the other settings on the dialog box apply. Up to five forecasting models may be considered at the same time, labeled A, B, C, D, and E.
- **Math:** Before fitting a model, the data may be transformed using any of the indicated operations. With the exception of the Box-Cox transformation, the selections are self-explanatory. The Box-Cox transformation is used when necessary to make the data more Gaussian. For a detailed discussion, see the documentation for the *Box-Cox Transformations* procedure.
- **Seasonal:** seasonally adjust the data using the indicated method before fitting the model. Seasonal adjustments are designed to remove any seasonal component from the data. The methods used are discussed in the documentation for the *Seasonal Decomposition* procedure.
- **Inflation:** adjusts the data for inflation using the specified inflation rate λ before fitting the model. If applied at the beginning of the period, the adjustment is

$$y'_t = \frac{y_t}{(1 + \lambda)^{(t-t_0+1)}} \quad (1)$$

where t_0 is the index of the first observation. If applied at the middle of the period, the adjustment is

$$y'_t = \frac{y_t}{(1 + \lambda)^{(t-t_0+0.5)}} \quad (2)$$

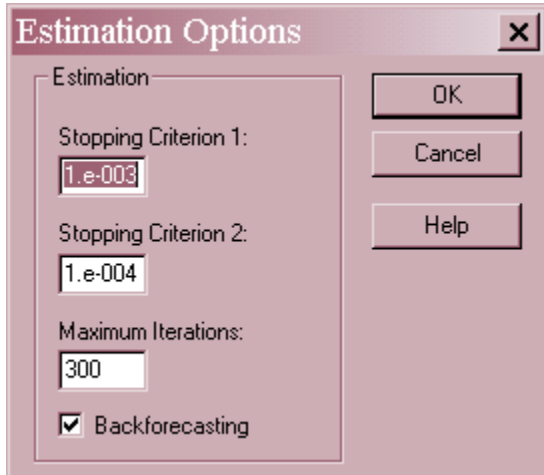
Note: Transformations are applied to the data before the forecasting model is fit. If more than one transformation is requested, they are applied in the following order:

1. trading day adjustment
2. inflation adjustment
3. math adjustment
4. seasonal adjustment

After the forecasts are generated, inverse transformations are applied to the forecasts in reverse order.

- **Type:** the type of forecasting model to be fit. For an explanation of the different types of models, see the discussion below.

- **Parameters and Terms:** options for different forecasting models.
 - *Alpha, beta, and gamma:* parameters for the *Exponential Smoothing* models. Each parameter must be greater than 0 and less than 1. The lower the value of a parameter, the greater the amount of smoothing that is performed.
 - *Order:* the number of terms in the *Moving Average* model.
 - *AR, MA, SAR, and SMA:* the order of the various components of the *ARIMA* models, referred to as p , q , P , and Q respectively in the discussion below.
 - *Optimize:* whether optimal values of the parameters should be found. If checked, the parameter values specified are used as starting values for the search procedures. If not checked, the values entered will be used in the model.
 - *Constant:* whether a constant term should be included when fitting a *Random Walk* or *ARIMA* model.
- **Differencing:** the order of seasonal and nonseasonal differencing to be applied when fitting the *ARIMA* models, referred to as d and D in the discussion below.
- **Estimation Button:** displays a dialog box that controls the nonlinear estimation procedure used when optimizing the exponential smoothing and *ARIMA* models.



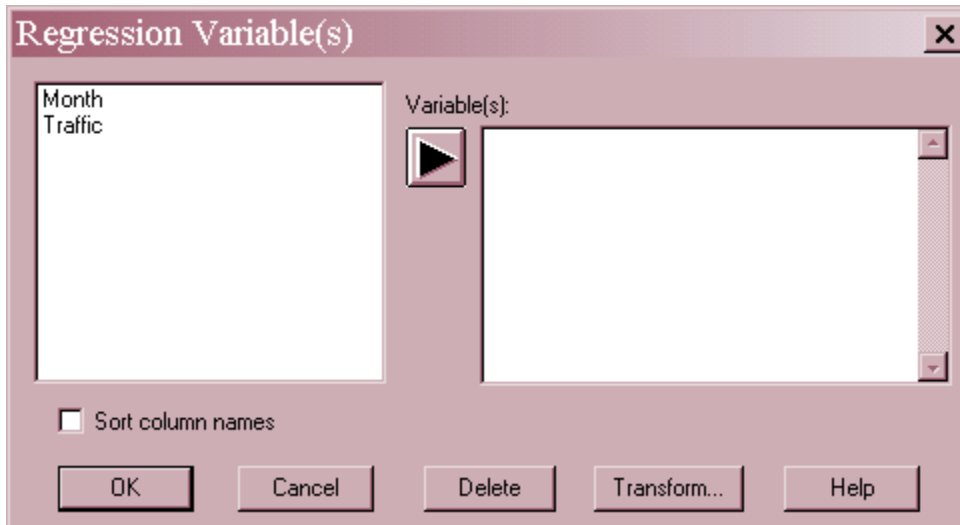
Stopping Criterion 1: The algorithm is assumed to have converged when the relative change in the residuals sums of squares from one iteration to the next is less than this value.

Stopping Criterion 2: The algorithm is assumed to have converged when the relative change in all parameter estimates from one iteration to the next is less than this value.

Maximum Iterations: Estimation stops if convergence is not achieved within this many iterations.

Backforecasting: Uses a method called *backforecasting* to forecast values prior to time $t = 1$. These values are used to generate the initial values which are needed to generate forecasts for small values of t . For details, see Box, Jenkins and Reinsel (1994).

- **Regression Button:** adds additional independent variables to the forecasting model when estimating a trend or ARIMA model. Typically, such variables are lagged values of leading indicators.



Variables: values of X variables to include in the model. If you wish to include a column named X but lag the data by 3 rows so that the model includes a term involving X_{t-3} , enter $LAG(X,3)$ instead of just X .

Note: Whichever letter is selected in the *Model* field when the dialog box is closed is taken to be the primary model. This is the model used when generating all of the tables and plots (except for the *Model Comparisons* pane, which compares them all).

Forecasting Models

Each of the forecasting models takes a different approach for forecasting future values. In the discussions below, the following notation will be used:

Y_t = observed value at time t , $t = 1, 2, \dots, n$

n = sample size (number of observations used to fit the model)

$F_t(k)$ = forecast for time $t+k$ made at time t

e_t = one period ahead forecasting errors calculated from

$$e_t = Y_t - F_{t-1}(1) \quad (3)$$

Given that m observations at the end of the time series have been withheld for validation purposes, two important validation statistics are:

RMSE = root mean squared error over the validation period, given by

$$RMSE = \sqrt{\frac{\sum_{i=1}^m e_{n+i}^2}{m}} \quad (4)$$

MAPE = mean absolute percentage error over the validation period, given by

$$MAPE = 100 \frac{\sum_{i=1}^m |e_{n+i} / Y_{n+i}|}{m} \% \quad (5)$$

The RMSE estimates the standard deviation of the one-ahead forecast errors. The MAPE estimates the average percentage one-ahead forecasting error. Small values of RMSE and MAPE are desirable.

Random Walk Model

The random walk model is very simple. Without a constant, it uses the current value of the time series to forecast all future values, i.e.,

$$F_t(k) = Y_t \text{ for all } k \geq 1 \quad (6)$$

This model is often used for data that does not have a fixed mean and for which the history of the process is irrelevant given its current position. The time series is thus equally likely to go up or down at any point in time.

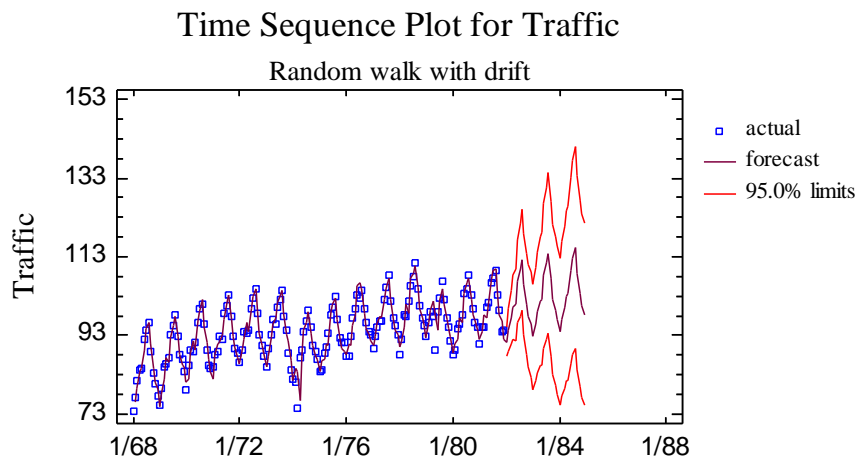
If a constant is included, then the forecast is given by

$$F_t(k) = Y_t + k\hat{\Delta} \quad (7)$$

where $\hat{\Delta}$ estimates the average change from one period to the next. The forecast function for such a model is a straight line with slope equal to $\hat{\Delta}$.

For the sample data, the random walk model could be used if a constant is included and the time series is first seasonally adjusted. The results are shown below:

Model	Constant	Seasonal adj.	Validation RMSE	Validation MAPE
Random walk	Yes	Multiplicative	2.72	1.46%



The plot shows:

1. *Observed data*: shown using point symbols.
2. *One-ahead forecasts*: shown as a solid line passing through the data.
3. *Forecasts for future values*: extension of the forecasts past the end of the data.
4. *95% prediction limits*: the red bounds around the forecasts.

Note the wide prediction limits, typical of random walk models.

Trend Models

The *Mean*, *Linear Trend*, *Quadratic Trend*, *Exponential Trend*, and *S-Curve* models all fit various types of regression models to the data, using time as the independent variable. The models are fit by least squares, resulting in estimates of up to 3 coefficients: a , b , and c . Forecasts from the models are as follows:

$$\text{Mean model: } F_t(k) = \bar{Y} \quad \text{where } \bar{Y} \text{ is the average of the data up to and including time } t. \quad (8)$$

$$\text{Linear trend: } F_t(k) = \hat{a} + \hat{b}(t+k) \quad (9)$$

$$\text{Quadratic trend: } F_t(k) = \hat{a} + \hat{b}(t+k) + \hat{c}(t+k)^2 \quad (10)$$

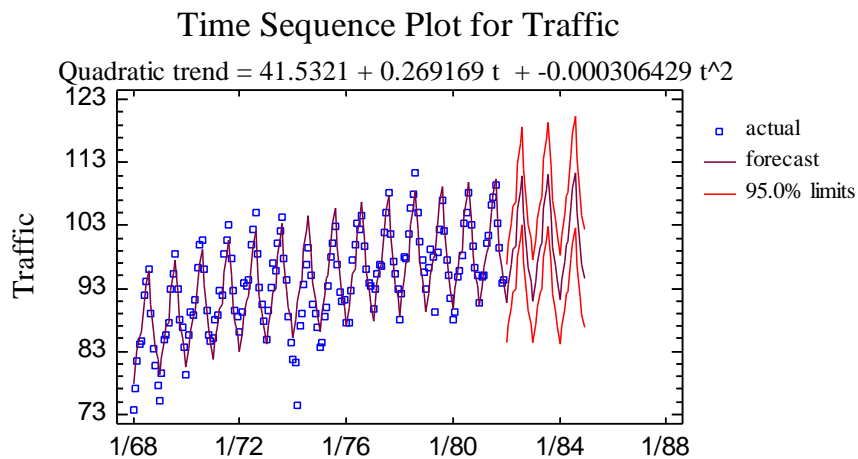
$$\text{Exponential trend: } F_t(k) = \exp(\hat{a} + \hat{b}(t+k)) \quad (11)$$

$$\text{S-Curve: } F_t(k) = \exp(\hat{a} + \hat{b}/(t+k)) \quad (12)$$

Since they weight all data equally, regression models are often not the best methods for forecasting time series data.

For the sample data, the best-fitting trend model is the *Quadratic Trend*, fit after seasonally adjusting the data.

<i>Model</i>	<i>Seasonal adj.</i>	<i>Validation RMSE</i>	<i>Validation MAPE</i>
Quadratic trend	Multiplicative	2.50	1.42%



Moving Average

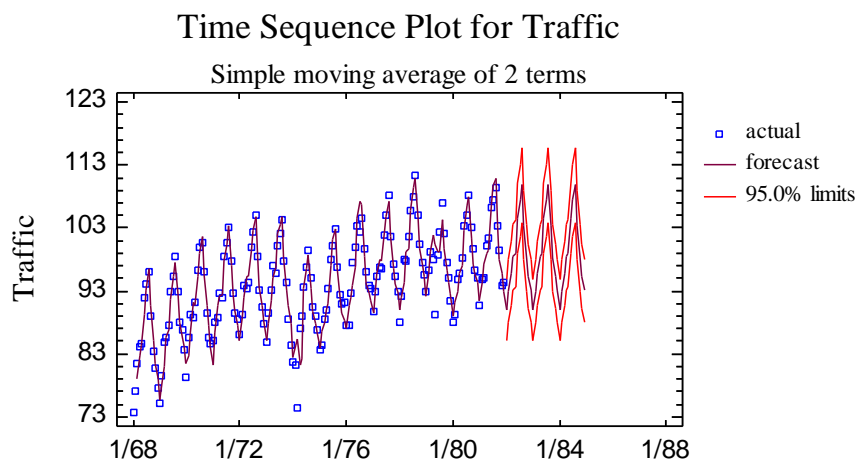
The *Moving Average* model uses the average of the most recent c observations to forecast future values. The forecasts are given by:

$$F_t(k) = \frac{\sum_{i=0}^{c-1} Y_{t-i}}{c} \quad \text{for all } k \geq 1 \quad (13)$$

Such a model can track a series that moves up and down, but tends to lag behind the actual series.

Experimenting with various orders of moving averages, it was found that an average of $c = 2$ observations gave the best fit for the traffic data during the validation period.

<i>Model</i>	<i>Order</i>	<i>Seasonal adj.</i>	<i>Validation RMSE</i>	<i>Validation MAPE</i>
Moving average	2	Multiplicative	2.08	1.27%



Note that the forecast function has no trend, which is counter-intuitive given the observed behavior. However, the one-month-ahead forecasts appear to be very good.

Exponential Smoothing

The *Simple Exp. Smoothing*, *Brown's Linear Exp. Smoothing*, and *Quadratic Exp. Smoothing* models estimate trends similar to the *Mean*, *Linear Trend*, and *Quadratic Trend* models, respectively. However, they do so by weighting recent observations more heavily than observations that are further in the past.

To generate the forecasts, up to three passes of an exponential smoother are made:

$$S'_t = \alpha Y_t + (1 - \alpha)S'_{t-1} \quad (14)$$

$$S''_t = \alpha S'_t + (1 - \alpha)S''_{t-1} \quad (15)$$

$$S'''_t = \alpha S''_t + (1 - \alpha)S'''_{t-1} \quad (16)$$

The initial values at time $t = 0$ are determined by backforecasting (unless suppressed using the *Estimation* button on the *Analysis Options* dialog box), which first smoothes the time series backwards and then uses the backforecasts to initialize the forward smoothing. The forecasts are then generated from:

$$\text{Simple smoothing: } F_t(k) = S'_t \quad (17)$$

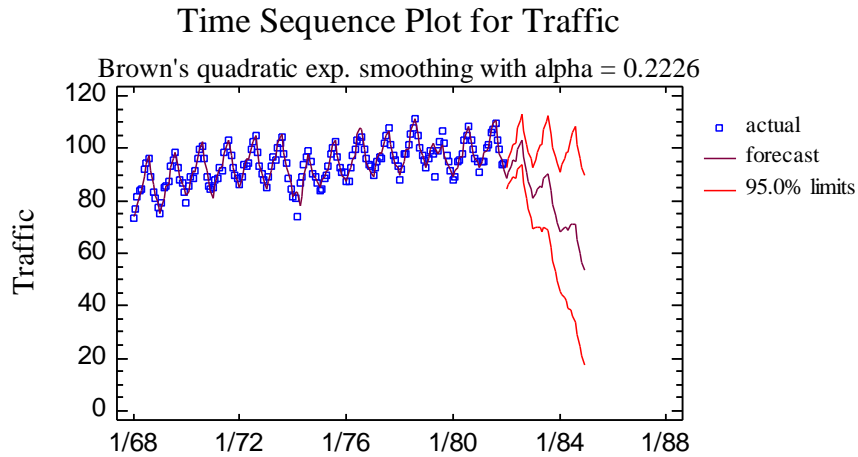
$$\text{Linear smoothing: } F_t(k) = 2S'_t - S''_t + k \frac{\alpha}{1 - \alpha} (S'_t - S''_t) \quad (18)$$

Quadratic smoothing:

$$F_t(k) = \frac{3S'_t - 3S''_t + S'''_t + k \frac{\alpha}{2(1 - \alpha)^2} ((6 - 5\alpha)S'_t - (10 - 8\alpha)S''_t + (4 - 3\alpha)S'''_t)}{+ k^2 \frac{\alpha^2}{2(1 - \alpha)^2} (S'_t - 2S''_t + S'''_t)} \quad (19)$$

The *Quadratic Exp. Smoother* gives the best results of the three procedures during the validation period when forecasting one period ahead. However, extrapolation of a quadratic trend into the future is always problematic, as can be seen in the plot below.

<i>Model</i>	<i>Alpha (optimized)</i>	<i>Seasonal adj.</i>	<i>Validation RMSE</i>	<i>Validation MAPE</i>
Quadratic exp. smoothing	0.2226	Multiplicative	2.42	1.31%



This example illustrates several important facts:

1. It is important to look at the results whenever a forecasting model is fit to be sure that the results make sense.
2. Models that are good for short-term forecasting may not behave well at forecasting values far into the future.
3. Models involving polynomials of order 2 or higher can behave erratically.

Holt's Linear Exponential Smoothing

Holt's Linear Exp. Smoothing is similar to Brown's Linear Exp. Smoothing in that it generates forecasts that follow a linear trend. However, Holt's procedure uses two smoothing constants, α and β , one to estimate the level of the series at time t and a second to estimate the slope. The procedure is as follows:

1. Smooth the data to estimate the level using

$$S_t = \alpha Y_t + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (20)$$

2. Smooth the first smooth to estimate the slope using

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1} \quad (21)$$

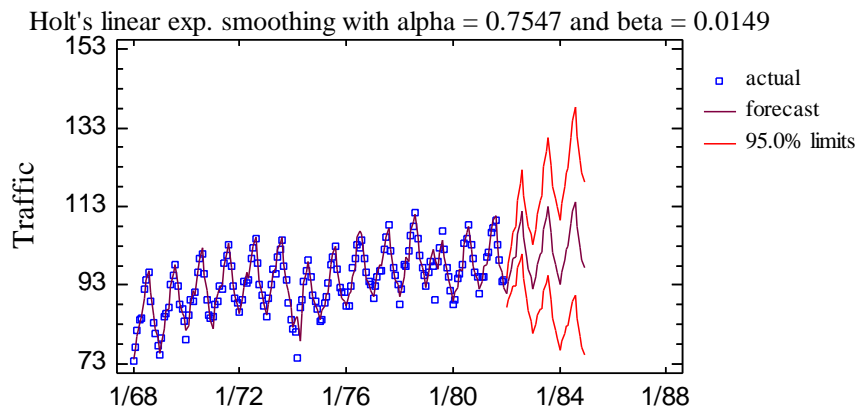
3. Calculate the forecasts using

$$F_t(k) = S_t + kT_t \quad (22)$$

The following shows the results of optimizing Holt's smoother after seasonally adjusting the data:

<i>Model</i>	<i>Alpha (optimized)</i>	<i>Beta (optimized)</i>	<i>Seasonal adj.</i>	<i>Validation RMSE</i>	<i>Validation MAPE</i>
Holt's linear exponential smoothing	0.7547	0.0149	Multiplicative	2.29	1.34%

Time Sequence Plot for Traffic



The results look quite reasonable.

Winter's Exponential Smoothing

All of the forecasting methods described above handle the seasonality by first seasonally adjusting the data, then applying the forecasting model, and then putting back the seasonality. *Winter's Exp. Smoothing* procedure handles the seasonality directly by estimating seasonality at the same time that it estimates the level and trend. It extends Holt's procedure by adding an additional parameter γ to use in a third smoother. The procedure is as follows:

1. Estimate the seasonality by smoothing the ratio of the data to the estimated level at time t using:

$$I_t = \gamma \frac{Y_t}{S_t} + (1 - \gamma)I_{t-s} \quad (23)$$

where s is the length of seasonality.

2. Estimate the level of the series by smoothing the data divided by the estimated seasonality using

$$S_t = \alpha \frac{Y_t}{I_{t-s}} + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (24)$$

3. Estimate the slope of the series using

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1} \quad (25)$$

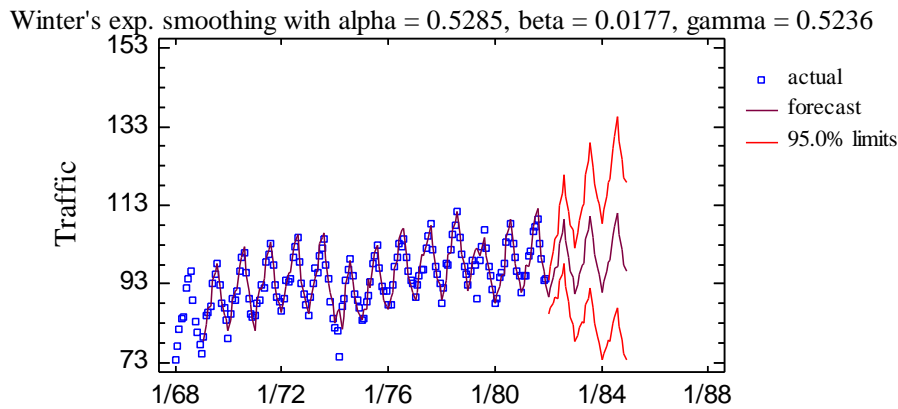
4. Calculate the forecasts using

$$F_t(k) = (S_t + kT_t)I_{t-s+m} \quad (26)$$

The following shows the results of optimizing Winter's smoother:

<i>Model</i>	<i>Alpha (optimized)</i>	<i>Beta (optimized)</i>	<i>Gamma (optimized)</i>	<i>Validation RMSE</i>	<i>Validation MAPE</i>
Winter's seasonal exp. smoothing	0.5285	0.0177	0.5236	3.34	1.58%

Time Sequence Plot for Traffic



Its performance on the traffic data is not as good as with some of the other methods. Also, the parameters are hard to estimate numerically and may vary quite a bit depending on the starting values of the search procedure.

ARIMA Models

The final choice of forecasting models, the ARIMA models, are the most general and include many of the other models as special cases. *ARIMA* models (short for “*AutoRegressive, Integrated, Moving Average*”), express the observation at time t as a linear function of previous observations, a current error term, and a linear combination of previous error terms.

The general form of the model is most easily expressed in terms of the backwards operator B , which operates on the time index of a data value such that $B^j Y_t = Y_{t-j}$. Using this operator, the model takes the form

$$\begin{aligned} & (1 - B - B^2 - \dots - B^p) (1 - B^s - B^{2s} - \dots - B^{Ps}) (1 - B)^d (1 - B^s)^D Z_t \\ & = (1 - B - B^2 - \dots - B^q) (1 - B^s - B^{2s} - \dots - B^{Qs}) a_t \end{aligned} \quad (27)$$

where

$$Z_t = Y_t - \mu \quad (28)$$

and a_t is a random error or shock to the system at time t , usually assumed to be random observations from a normal distribution with mean 0 and standard deviation σ_a . For a stationary series, μ represents the process mean. Otherwise, it is related to the slope of the forecast function. μ is sometimes assumed to equal 0.

The above model is often referred to as an $ARIMA(p,d,q) \times (P,D,Q)_s$ model. It consists of several terms:

1. A nonseasonal autoregressive term of order p .

2. Nonseasonal differencing of order d .
3. A nonseasonal moving average term of order q .
4. A seasonal autoregressive term of order P
5. Seasonal differencing of order D .
6. A seasonal moving average term of order Q .

While the general model looks formidable, the most commonly used models are relatively simple special cases. These include:

AR(1) – autoregressive of order 1

The observation at time t is expressed as a mean plus a multiple of the deviation from the mean at the previous time period plus a random shock:

$$Y_t = \mu + \phi_1(Y_{t-1} - \mu) + a_t \quad (29)$$

AR(2) – autoregressive of order 2

The observation at time t is expressed as a mean plus multiples of the deviations from the mean at the 2 previous time periods plus a random shock:

$$Y_t = \mu + \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + a_t \quad (30)$$

MA(1) – moving average of order 1

The observation at time t is expressed as a mean plus a random shock at the current time period plus a multiple of the random shock at the previous time period:

$$Y_t = \mu + a_t - \theta_1 a_{t-1} \quad (31)$$

MA(2) – moving average of order 2

The observation at time t is expressed as a mean plus a random shock at the current time period plus multiples of the random shocks at the 2 previous time periods:

$$Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad (32)$$

ARMA(1,1) – mixed model with 2 first order terms

The observation at time t is expressed as a mean plus a multiple of the deviation from the mean at the previous time period plus a random shock at the current time period plus a multiple of the random shock at the previous time period:

$$Y_t = \mu + \phi_1(Y_{t-1} - \mu) + a_t - \theta_1 a_{t-1} \quad (33)$$

ARIMA(0,1,1) – moving average of order 1 applied to the first differences

The difference between the current period and the previous period is expressed as a random shock at the current time period plus a multiple of the random shock at the previous time period:

$$Y_t - Y_{t-1} = a_t - \theta_1 a_{t-1} \quad (34)$$

It can be shown that this model is equivalent to the *Simple Exponential Smoothing* model.

ARIMA(0,2,2) – moving average of order 2 applied to the second differences

The difference of the differences is expressed as a random shock at the current time period plus multiples of the random shocks at the 2 previous time periods:

$$(Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad (35)$$

This model is equivalent to the *Holt's Linear Exponential Smoothing* model.

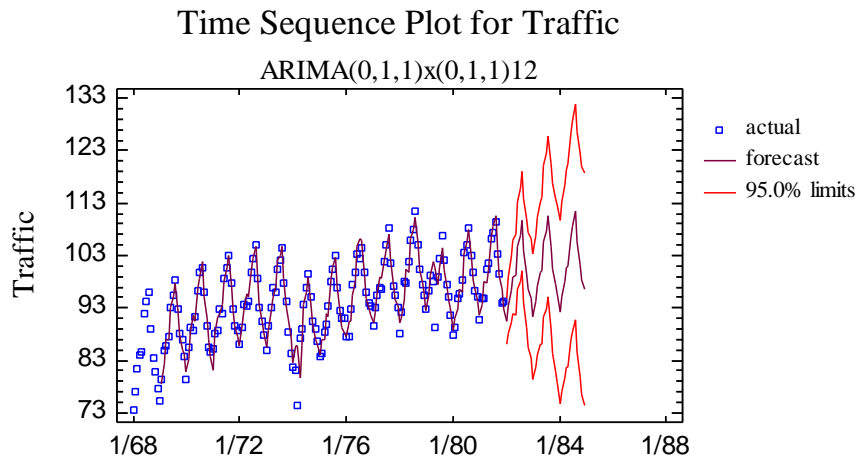
ARIMA(0,1,1)x(0,1,1)_s – seasonal and nonseasonal MA terms of order 1

The observation at time t is expressed as a combination of the observation one season ago plus the difference between the observation last period and its counterpart one season ago plus multiple of the shocks to hit the system this period, last period, and two periods one season ago:

$$Y_t = Y_{t-s} + Y_{t-1} - Y_{t-s-1} + a_t - \theta_1 a_{t-1} - \Theta_1 a_{t-s} + \theta_1 \Theta_1 a_{t-s-1} \quad (36)$$

Many economic time series with a seasonal component can be well represented by this model. It also does very well on the Golden Gate Bridge traffic data:

<i>Model</i>	<i>MA(1)</i>	<i>SMA(1)</i>	<i>Validation RMSE</i>	<i>Validation MAPE</i>
ARIMA	0.2273	0.8568	2.15	1.27%



Intuitively, the model expresses the difference in the traffic this month compared to the same month last year as being equal to the difference observed last month, plus a combination of the noise observed last month, last year, and 13 months ago.

The classic reference for constructing ARIMA models is Box, Jenkins and Reinsel (1994).

Analysis Summary

The results of fitting a forecasting model are displayed in the *Analysis Summary*. As an example, the table below shows the results of fitting the ARIMA(0,1,1)x(0,1,1)₁₂ model to the Golden Gate Bridge traffic data:

Forecasting - Traffic				
Data variable: Traffic (Golden Gate Bridge Traffic Volume)				
Number of observations = 168				
Start index = 1/68				
Sampling interval = 1.0 month(s)				
Length of seasonality = 12				
Forecast Summary				
Nonseasonal differencing of order: 1				
Seasonal differencing of order: 1				
Forecast model selected: ARIMA(0,1,1)x(0,1,1) ₁₂				
Number of forecasts generated: 36				
Number of periods withheld for validation: 24				
	<i>Estimation</i>	<i>Validation</i>		
<i>Statistic</i>	<i>Period</i>	<i>Period</i>		
RMSE	2.1868	2.14839		
MAE	1.38616	1.23025		
MAPE	1.52679	1.26567		
ME	-0.0410165	-0.000900154		
MPE	-0.080606	-0.0111287		
ARIMA Model Summary				
<i>Parameter</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>t</i>	<i>P-value</i>
MA(1)	0.277336	0.0843672	3.28724	0.001255
SMA(1)	0.85681	0.0283923	30.1776	0.000000
Backforecasting: yes				
Estimated white noise variance = 4.93026 with 153 degrees of freedom				
Estimated white noise standard deviation = 2.22042				
Number of iterations: 6				

There are several important sections in the output:

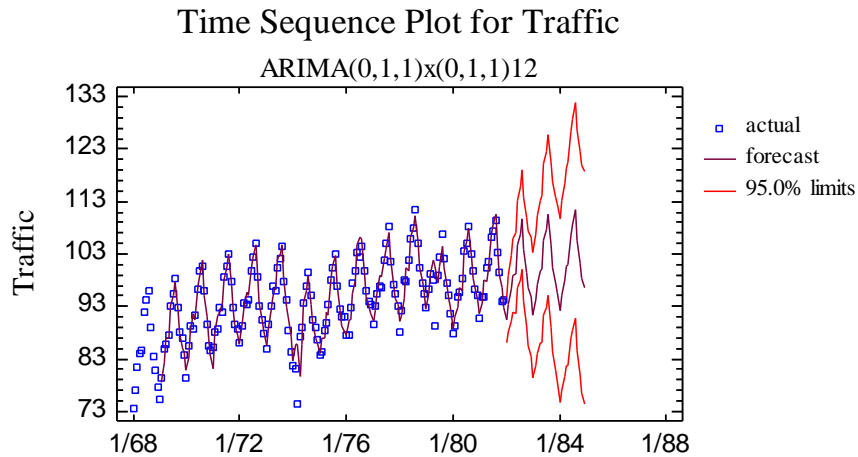
- **Data Summary:** the top section summarizes the input data and the length of seasonality s , if any.
- **Forecast Summary:** indicates any transformations that were made to the data, as well as the type of model that was fit. The number of periods m that were withheld for validation purposes is also shown.
- **Table of Statistics:** shows statistics calculated from the one-ahead forecast errors during both the estimation and validation periods. In addition to the root mean squared error (RMSE) and mean absolute percentage error (MAPE) described earlier, the program also displays the mean absolute error (MAE), the mean error (ME), and the mean percentage error (MPE). Ideally, RMSE, MAE, and MAPE will be small, since they measure the variability of the forecast errors. ME and MPE should be close to zero if the forecasts are not biased.

- **ARIMA Model Summary** – displays statistics for the coefficients of the fitted ARIMA model. A similar table is displayed when trend models are fit using least squares regression. Of interest are:
 - *Estimate*: the estimated coefficient.
 - *Std. error*: the standard error of the coefficient.
 - *t*: the value of a t statistic calculated by dividing the estimated coefficient by its standard error.
 - *P-value*: two-sided P-value calculated from Student's t distribution with the degrees of freedom indicated below the table. Small P-values (less than 0.05 if operating at the 5% significance level) correspond to statistically significant coefficients. If any P-values are greater than 0.05, consideration should be given to reducing the complexity of the model.
 - *Estimated white noise standard deviation*: estimate of the standard deviation of the noise $\hat{\sigma}_a$ that is unaccounted for by the model.
 - *Number of iterations*: the number of iterations used by the nonlinear estimation procedure.

In the example, the fitted ARIMA model has two parameters, both of which are statistically significant. The ME and MPE are close to zero in both the estimation and the validation periods, indicating little bias in the one-month ahead forecasts. Examining the RMSE, MAE, and MAPE, the model appears to do no worse (and possibly better) during the validation period than in the estimation period.

Time Sequence Plot

The *Time Sequence Plot* displays the data, the forecasts, and the forecast limits:



The plot shows:

1. The observed data Y_t , including any replacements for missing values, shown as point symbols.
2. The one-step ahead forecasts $F_t(1)$, displayed as a solid line through the points. These are created using the fitted model, forecasting each time period $t+1$ using only the information available at time t . The one-ahead forecast errors e_t are observable as the vertical distance between the observations and the solid line.
3. Forecasts for future values $F_{n+m}(k)$ made at time $t = n+m$, the last time at which observed data is available. These are shown by the extension of the solid forecast line beyond the last observation.
4. Probability limits for the forecasts at the $100(1-\alpha)\%$ confidence level, calculated assuming that the noise in the system follows a normal distribution. The limits are given by

$$F_{n+m}(k) \pm z_{\alpha/2} \sqrt{\hat{V}(k)} \quad (37)$$

where $\hat{V}(k)$ equals the estimated variance of the forecast k periods past the end of the data. The formula for the variance depends on the model used, as outlined in the *Calculations* section. It should be noted that the limits are only valid if several assumptions hold, including:

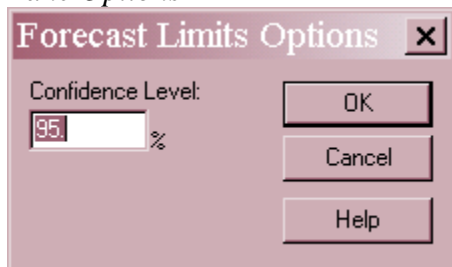
- a. The proper model has been selected.
- b. The selected model was valid for all of the historical data.
- c. The selected model continues to be valid in the future.

- d. The shocks to the system follow a normal distribution.
- e. The model has been estimated from a long enough time series that the model estimation error is small compared to the variability of the error term (except for models estimated by linear regression which include the model estimation error).

In practice, the limits should be regarded as no more than an approximation of how far the time series may stray from the forecasted values in the future.

The forecasted pattern for the Golden Gate Bridge traffic shows a continued upward trend with a strong seasonal oscillation. Although the forecast limits may seem quite wide, they must allow for the possibility of dramatic events such as were observed twice in the past.

Pane Options



- **Confidence Level:** the percentage to use for the probability limits.

Forecast Table

The *Forecast Table* displays the forecasts for both the historical and future periods. A portion of the output is shown below:

Forecast Table for Traffic				
Model: ARIMA(0,1,1)x(0,1,1) ₁₂				
V = withheld for validation				
Period	Data	Forecast	Residual	
7/81	107.415	108.624	-1.20914	V
8/81	109.385	110.652	-1.26658	V
9/81	103.266	104.117	-0.850957	V
10/81	99.432	99.1015	0.330487	V
11/81	93.965	96.5951	-2.63007	V
12/81	94.385	92.6992	1.68583	V
		Lower 95.0%	Upper 95.0%	
Period	Forecast	Limit	Limit	
1/82	90.5668	86.1802	94.9535	
2/82	93.3945	87.9823	98.8067	
3/82	96.3379	90.0656	102.61	
4/82	98.8411	91.8132	105.869	
5/82	99.3509	91.6412	107.061	
6/82	104.843	96.5074	113.179	

The top section of the output shows:

- **Period:** the time period t corresponding to each historical observation.

- **Data:** the observed data value Y_t , including any replacements for missing values.
- **Forecast:** the forecast for time t using all of the information available at time $t-1$.
- **Residual:** the one-ahead forecast error e_t , calculated by subtracting the forecast from the observed data value.
- **V:** indicates that the corresponding observation was not used to fit the model but instead was included in the validation set.

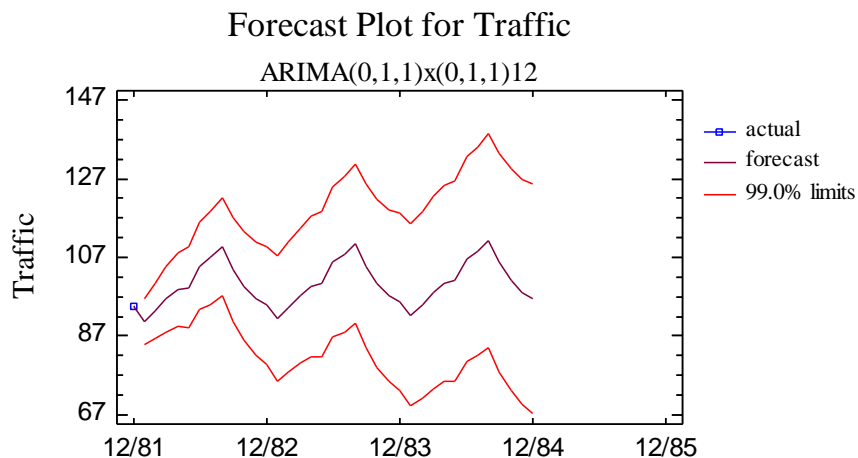
The bottom section of the output shows:

- **Period:** the time t corresponding to periods beyond the end of the observed data.
- **Forecast:** the forecast $F_{n+m}(k)$ for time period t using all of the available data.
- **Limits:** the probability limits for the forecasts.

For example, the forecasted traffic on the Golden Gate Bridge in June of 1982, made at the end of 1981, was 104.8. The 95% limits ranged from 96.5 to 113.2.

Forecast Plot

The *Forecast Plot* shows the last several observations, the forecasts, and the forecast limits:



It is similar to the *Time Sequence Plot*, except that it gives a closer view of the forecasts.

Model Comparisons

The *Model Comparisons* pane displays statistics that compare each of the models selected on the *Analysis Options* dialog box.

Model Comparison
 Data variable: Traffic
 Number of observations = 168
 Start index = 1/68
 Sampling interval = 1.0 month(s)
 Length of seasonality = 12
 Number of periods withheld for validation: 24

Models
 (A) ARIMA(0,1,1)x(0,1,1)¹²
 (B) Winter's exp. smoothing with alpha = 0.5153, beta = 0.0207, gamma = 0.4973

Estimation Period

Model	RMSE	MAE	MAPE	ME	MPE
(A)	2.1868	1.38616	1.52679	-0.0410165	-0.080606
(B)	2.40364	1.55402	1.70498	-0.275112	-0.315288

Model	RMSE	RUNS	RUNM	AUTO	MEAN	VAR
(A)	2.1868	OK	OK	OK	OK	***
(B)	2.40364	OK	OK	OK	OK	***

Validation Period

Model	RMSE	MAE	MAPE	ME	MPE
(A)	2.14839	1.23025	1.26567	-0.000900154	-0.0111287
(B)	3.27348	1.56492	1.5713	-0.0905451	-0.0585017

Key:
 RMSE = Root Mean Squared Error
 RUNS = Test for excessive runs up and down
 RUNM = Test for excessive runs above and below median
 AUTO = Box-Pierce test for excessive autocorrelation
 MEAN = Test for difference in mean 1st half to 2nd half
 VAR = Test for difference in variance 1st half to 2nd half
 OK = not significant ($p \geq 0.05$)
 * = marginally significant ($0.01 < p \leq 0.05$)
 ** = significant ($0.001 < p \leq 0.01$)
 *** = highly significant ($p \leq 0.001$)

The tables labeled *Estimation Period* and *Validation Period* display statistics calculated from the one-ahead forecast errors e_t in their respective periods:

- RMSE*: the root mean squared error.
- MAE*: the mean absolute error.
- MAPE*: the mean absolute percentage error.
- ME*: the mean error.
- MPE*: the mean percentage error.

Better models have smaller RMSE, MAE, and MAPE values, which measure the variance of the forecasting errors. ME and MPE are measures of bias and should be close to 0.

For the estimation period only, several tests are applied to the forecast errors to determine whether the model has accounted for all of the structure in the data. These tests are designed to

determine whether the residuals form a random time series (“white noise”) and are described in the *Time Series – Descriptive Methods* documentation. Included are:

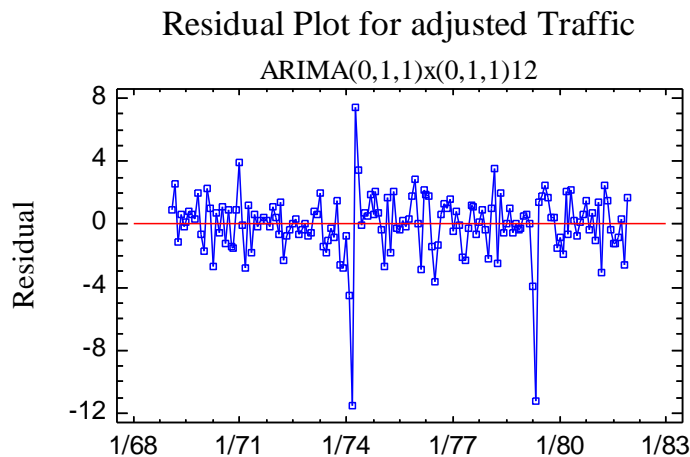
- RUNS*: a test based on the number of runs up and down.
- RUNM*: a test based on the number of runs above and below the median.
- AUTO*: a chi-squared test based on the first k residual autocorrelations, where k is set by *Pane Options* in the table displaying the residual autocorrelations.
- MEAN*: a t-test comparing the mean residuals in the first and second halves of the data.
- VAR*: an F-test comparing the variance of the residuals in the two halves.

If the entry for a particular test is *OK*, then the test is not statistically significant at the 5% significance level and the assumption of random residuals is not rejected. Otherwise, the number of stars (*) indicates the significance level at which the assumption of random residuals would be rejected.

Both of the models fit to the traffic data pass all of the tests except that comparing the two variances. The latter test is highly significant. As will be seen when the residual plots are examined, this failure is due to the presence of three large residuals during the second half of the estimation period.

Residual Plots

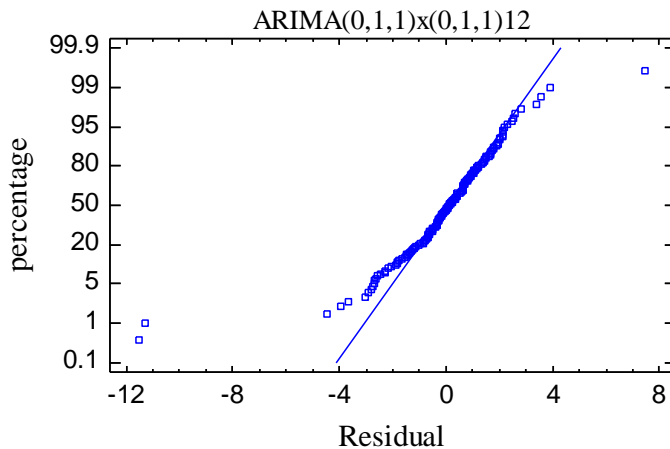
The *Residual Plots* display the one-ahead forecast errors e_t in several ways. The default plot displays the residuals in sequential order:



Notice the three large spikes occurring in March and April of 1974 and May of 1979. Traffic in those months changed much more than normal.

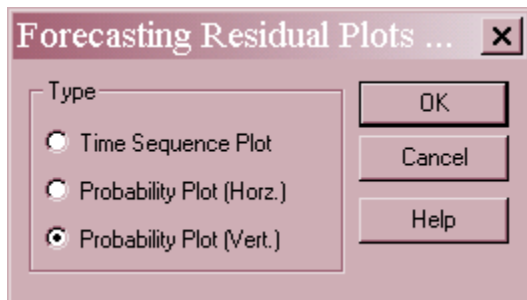
Using *Pane Options*, a normal probability plot of the residuals can be displayed instead:

Residual Plot for adjusted Traffic



If the residuals come from a normal distribution, they should fall close to the line. The plot above shows some curvature away from the line in the tails, plus 3 outliers.

Pane Options



Three different plots made be displayed:

1. *Time Sequence Plot* – a plot of the residuals versus time.
2. *Probability Plot (Horz.)* – a probability plot with the percentages displayed on the horizontal axis.
3. *Probability Plot (Vert.)* – a probability plot with the percentages displayed on the vertical axis (as shown above).

Residual Autocorrelations

It is also useful to examine the autocorrelations of the residuals. The residual autocorrelation at lag k measures the strength of the correlation between residuals k time periods apart. The residual lag k autocorrelation is calculated from

$$r_k = \frac{\sum_{t=1}^{n-k} (e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2} \quad (38)$$

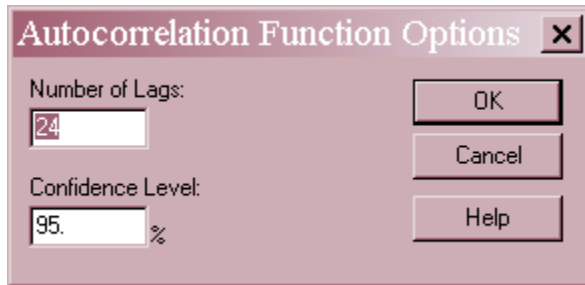
If a model describes all of the dynamic structure in a time series, then the residuals should be random and all of their autocorrelations should be insignificant.

The *Residual Autocorrelations* pane displays the residual autocorrelations together with large lag standard errors and probability limits:

Estimated Autocorrelations for residuals				
Data variable: Traffic				
Model: ARIMA(0,1,1)x(0,1,1)12				
Lag	Autocorrelation	Std. Error	Lower 95.0% Prob. Limit	Upper 95.0% Prob. Limit
1	0.020701	0.0803219	-0.157428	0.157428
2	-0.14092	0.0803563	-0.157496	0.157496
3	-0.0782807	0.0819352	-0.16059	0.16059
4	-0.0613474	0.0824163	-0.161533	0.161533
5	-0.0687687	0.0827104	-0.16211	0.16211
6	0.0266528	0.0830785	-0.162831	0.162831
7	0.0390244	0.0831336	-0.162939	0.162939
8	0.0182106	0.0832517	-0.163171	0.163171
9	-0.0183478	0.0832774	-0.163221	0.163221
10	-0.0917506	0.0833035	-0.163272	0.163272
11	0.040521	0.0839529	-0.164545	0.164545
12	-0.0970784	0.084079	-0.164792	0.164792
13	0.0550271	0.0847991	-0.166203	0.166203
14	-0.0223925	0.0850291	-0.166654	0.166654
15	0.0119691	0.0850672	-0.166729	0.166729
16	0.0101773	0.085078	-0.16675	0.16675
17	-0.00617232	0.0850859	-0.166766	0.166766
18	0.00450753	0.0850888	-0.166771	0.166771
19	0.0562645	0.0850903	-0.166774	0.166774
20	-0.00489781	0.08533	-0.167244	0.167244
21	-0.125667	0.0853318	-0.167248	0.167248
22	0.0167626	0.0865176	-0.169572	0.169572
23	0.0627018	0.0865385	-0.169613	0.169613
24	-0.059002	0.0868311	-0.170186	0.170186

Any autocorrelations that fall outside the probability limits are statistically significant at the indicated level. The StatAdvisor highlights any such autocorrelations in red.

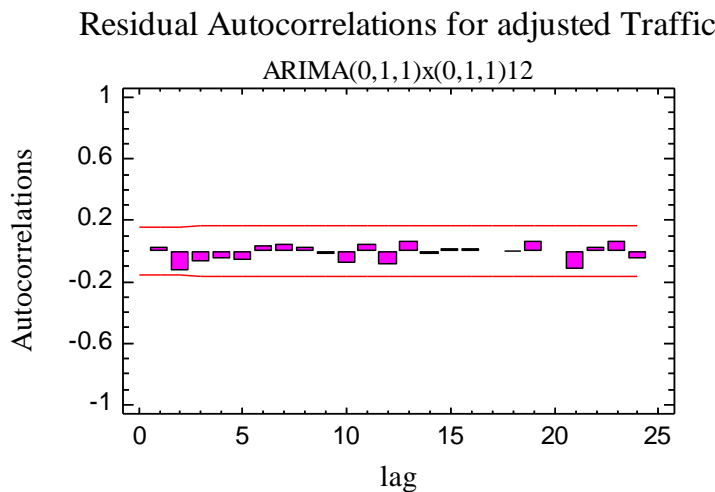
Pane Options



- **Number of lags:** maximum lag k at which to calculate the autocorrelation.
- **Confidence level:** value of $100(1-\alpha)\%$ used to calculate the probability limits.

Residual Autocorrelation Function

The *Residual Autocorrelation Function* plot displays the residual autocorrelations and probability limits:



Bars extending beyond the upper or lower limit correspond to statistically significant autocorrelations.

For the traffic data, the only estimate that is close to a probability limit is the estimate at $k = 2$. In fact, a slight reduction in the RMSE during the estimation period can be achieved by increasing the order of the nonseasonal MA term from 1 to 2. However, the performance of the model during the validation period is worse than with the current model, so the simpler model has been selected.

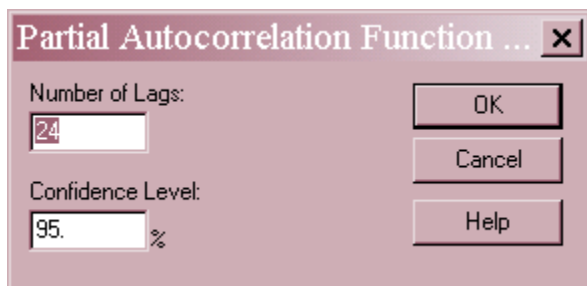
Residual Partial Autocorrelations

If the model fits well, the residual partial autocorrelations should also be insignificant. The *Residual Partial Autocorrelations* pane displays the residual partial autocorrelations together with large lag standard errors and probability limits:

Estimated Partial Autocorrelations for residuals				
Data variable: Traffic				
Model: ARIMA(0,1,1) _x (0,1,1) ₁₂				
Lag	Partial Autocorrelation	Std. Error	Lower 95.0% Prob. Limit	Upper 95.0% Prob. Limit
1	0.020701	0.0803219	-0.157428	0.157428
2	-0.141409	0.0803219	-0.157428	0.157428
3	-0.0735248	0.0803219	-0.157428	0.157428
4	-0.0805431	0.0803219	-0.157428	0.157428
5	-0.0916127	0.0803219	-0.157428	0.157428
6	0.000904756	0.0803219	-0.157428	0.157428
7	0.00389133	0.0803219	-0.157428	0.157428
8	0.004734	0.0803219	-0.157428	0.157428
9	-0.0212021	0.0803219	-0.157428	0.157428
10	-0.092355	0.0803219	-0.157428	0.157428
11	0.0442202	0.0803219	-0.157428	0.157428
12	-0.129774	0.0803219	-0.157428	0.157428
13	0.0568189	0.0803219	-0.157428	0.157428
14	-0.0730993	0.0803219	-0.157428	0.157428
15	0.00455603	0.0803219	-0.157428	0.157428
16	-0.00205713	0.0803219	-0.157428	0.157428
17	-0.0218083	0.0803219	-0.157428	0.157428
18	0.0148837	0.0803219	-0.157428	0.157428
19	0.0466933	0.0803219	-0.157428	0.157428
20	-0.0104909	0.0803219	-0.157428	0.157428
21	-0.10974	0.0803219	-0.157428	0.157428
22	0.0036704	0.0803219	-0.157428	0.157428
23	0.0558175	0.0803219	-0.157428	0.157428
24	-0.104699	0.0803219	-0.157428	0.157428

The StatAdvisor highlights any significant partial autocorrelations in red.

Pane Options



Partial Autocorrelation Function ...

Number of Lags: OK

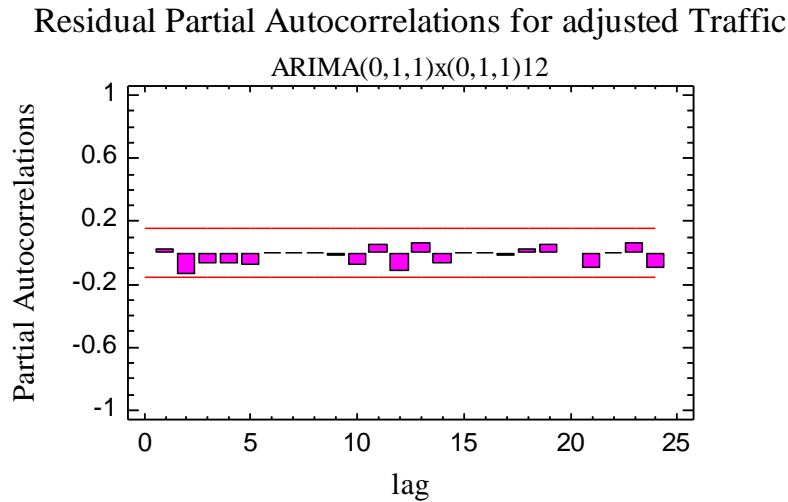
Confidence Level: % Cancel

Help

- **Number of lags:** maximum lag k at which to calculate the partial autocorrelation.
- **Confidence level:** value of $100(1-\alpha)\%$ used to calculate the probability limits.

Residual Partial Autocorrelation Function

The *Residual Partial Autocorrelation Function* plots the residual partial autocorrelations and probability limits:



All of the coefficients should be within the limits, as in the plot above.

Residual Periodogram Table

It is also useful to examine the residuals in the *frequency domain*, by considering how much variability exists at different frequencies. As described in the *Time Series – Descriptive Methods* documentation, the periodogram plots the power at each of the Fourier frequencies. If the residuals are random, there should approximately equal power at all frequencies, which is why a random time series is often called “white noise”.

The *Residual Periodogram* pane displays the following table:

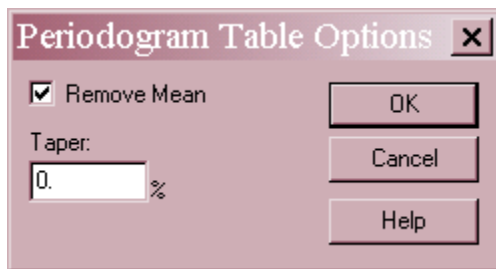
Periodogram for residuals					
Data variable: Traffic					
Model: ARIMA(0,1,1)x(0,1,1)12					
<i>i</i>	Frequency	Period	Ordinate	Cumulative Sum	Integrated Periodogram
0	0.0		5.15304E-32	5.15304E-32	7.71106E-35
1	0.00645161	155.0	0.377934	0.377934	0.000565545
2	0.0129032	77.5	7.63606	8.01399	0.0119922
3	0.0193548	51.6667	2.11531	10.1293	0.0151576
4	0.0258065	38.75	2.96824	13.0975	0.0195993
5	0.0322581	31.0	8.73905	21.8366	0.0326765
6	0.0387097	25.8333	1.13413	22.9707	0.0343736
7	0.0451613	22.1429	10.7466	33.7174	0.050455
8	0.0516129	19.375	4.51085	38.2282	0.0572051
9	0.0580645	17.2222	3.25441	41.4826	0.062075
10	0.0645161	15.5	20.2058	61.6884	0.0923111
11	0.0709677	14.0909	2.03049	63.7189	0.0953496
12	0.0774194	12.9167	5.83907	69.558	0.104087
13	0.083871	11.9231	6.57417	76.1321	0.113925
14	0.0903226	11.0714	0.310867	76.443	0.11439
15	0.0967742	10.3333	19.5081	95.9511	0.143582

The table includes:

- **Frequency:** the i -th Fourier frequency $f_i = i/n$.
- **Period:** the period associated with the Fourier frequency, given by $1/f_i$. This is the number of observations in a complete cycle at that frequency.
- **Ordinate:** the periodogram ordinate $I(f_i)$.
- **Cumulative Sum:** the sum of the periodogram ordinates at all frequencies up to and including the i -th.
- **Integrated Periodogram:** the cumulative sum divided by the sum of the periodogram ordinates at all of the Fourier frequencies. This column represents the proportion of the power in the time series at or below the i -th frequency.

Unlike the periodogram for the original traffic series, there is no large spike at a frequency of once every 12 months.

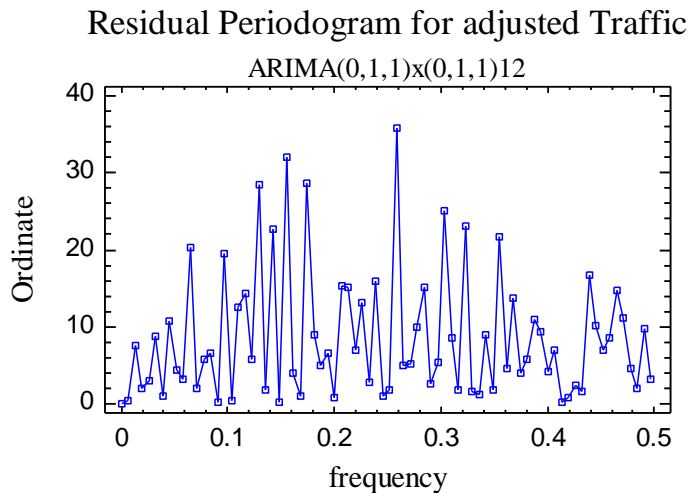
Pane Options



- **Remove mean:** check to subtract the mean from the time series before calculating the periodogram.
- **Taper:** percent of the data at each end of the time series to which a data taper will be applied before the periodogram is calculated. Following Bloomfield (2000), STATGRAPHICS uses a cosine taper that downweights observations close to $i = 1$ and $i = n$. This is useful for correcting bias if the periodogram ordinates are to be smoothed in order to create an estimate of the underlying spectral density function.

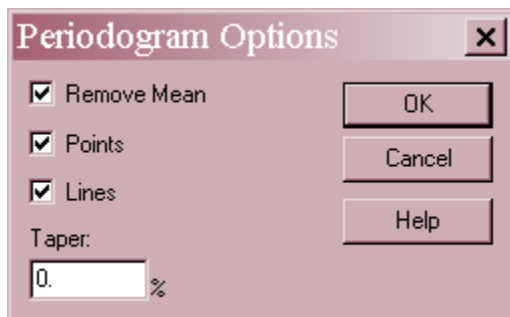
Residual Periodogram Plot

The *Residual Periodogram* plots the periodogram ordinates of the residuals:



If the residuals are random, there should be no noticeable spikes. Allowing for some natural skewness in the distribution of the ordinates, the above plot shows no large peaks.

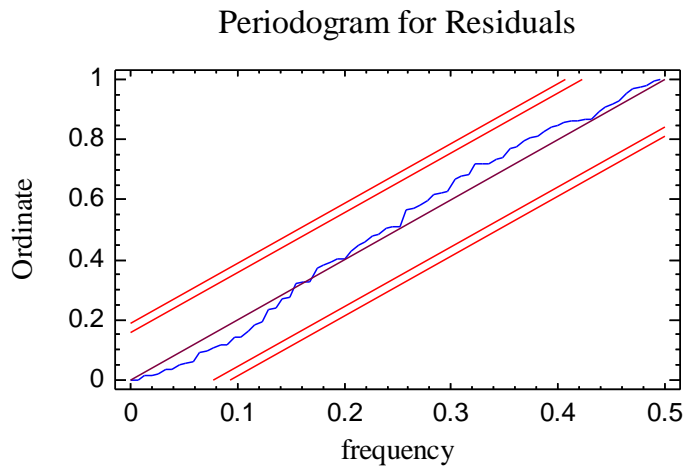
Pane Options



- **Remove mean:** check to subtract the mean from the time series before calculating the periodogram.
- **Points:** if checked, point symbols will be displayed.
- **Lines:** if checked, the ordinates will be connected by a line.
- **Taper:** percent of the data at each end of the time series to which a data taper will be applied before the periodogram is calculated.

Residual Integrated Periodogram

The *Residual Integrated Periodogram* displays the cumulative sums of the residual periodogram ordinates, divided by the sum of the ordinates over all of the Fourier frequencies:



A diagonal line is included on the plot, together with 95% and 99% Kolmogorov-Smirnov bounds. If the residuals are random, the integrated periodogram should fall within those bounds 95% and 99% of the time. For the traffic data, the residuals do appear to be white noise.

Tests For Randomness

The *Tests for Randomness* pane displays the results of additional tests run to determine whether or not the residuals are purely random:

Tests for Randomness of residuals	
Data variable: Traffic	
Model: ARIMA(0,1,1)x(0,1,1) ₁₂	
(1) Runs above and below median	
Median = 0.092978	
Number of runs above and below median = 79	
Expected number of runs = 78.0	
Large sample test statistic $z = 0.0808469$	
P-value = 0.935558	
(2) Runs up and down	
Number of runs up and down = 99	
Expected number of runs = 103.0	
Large sample test statistic $z = 0.670684$	
P-value = 0.50242	
(3) Box-Pierce Test	
Test based on first 24 autocorrelations	
Large sample test statistic = 13.6096	
P-value = 0.914755	

Three tests are performed:

1. *Runs above and below median*: counts the number of times the series goes above or below its median. This number is compared to the expected value for a random time

series. Small P-values (less than 0.05 if operating at the 5% significance level) indicate that the residuals are not purely random.

2. *Runs up and down*: counts the number of times the series goes up or down. This number is compared to the expected value for a random time series. Small P-values indicate that the time series is not purely random.
3. *Box-Pierce Test* or *Ljung-Box Test*: constructs a test statistic based on the first k residual autocorrelations. One of two tests is performed, as specified on the *Forecasting* tab of the *System Preferences* dialog box. For the Box-Pierce test:

$$Q = n \sum_{i=1}^k r_i^2 \quad (39)$$

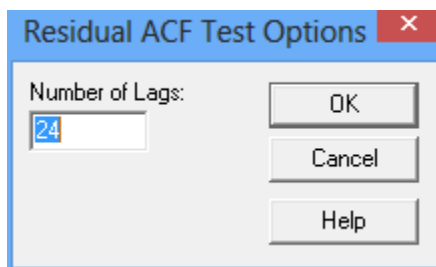
For the Ljung-Box test:

$$Q = n(n+2) \sum_{i=1}^k \frac{r_i^2}{n-i} \quad (40)$$

For either test, the test statistic is compared to a chi-squared distribution with k degrees of freedom. As with the other two tests, small P-values indicate that the residuals are not purely random.

Since the P-values for all three tests are well above 0.05, there is no reason to doubt that the residuals are white noise.

Pane Options



- **Number of Lags**: number of lags k to include in the Box-Pierce or Ljung-Box test.

Residual Crosscorrelations

The *Residual Crosscorrelations* pane displays crosscorrelations between the residuals and a second series, specified using *Pane Options*. The crosscorrelation between one time series Y at time t and a second time series X at time $t-k$ is denoted by $c_{xy}(k)$. A typical use of crosscorrelations is in identifying “leadings indicators” or an input-output relationship. For

example, Box, Jenkins and Reinsel (1994) present data from the input and output of a gas furnace at 9 second intervals, contained in the file *furnace.sgd*. The data consist of:

1. Output series Y: % Co2 in outlet gas
2. Input series X: input gas rate in cubic feet per minute

The output time series is well described by an ARIMA(3,1,0) model.

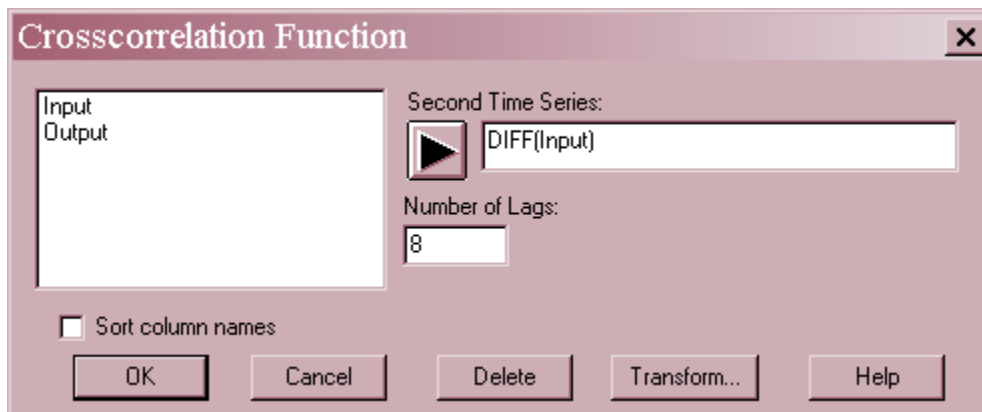
The table below shows the residual autocorrelations for the output model residuals and the similarly differenced input time series:

Estimated Crosscorrelations for residuals with DIFF(Input)
 Data variable: Output
 Model: ARIMA(3,1,0)

Lag	Crosscorrelation
-8	-0.0530088
-7	0.00912287
-6	0.0463566
-5	0.127723
-4	0.13986
-3	0.163528
-2	0.206428
-1	0.174213
0	0.09082
1	-0.0977499
2	-0.364327
3	-0.515241
4	-0.417341
5	-0.237661
6	-0.0153997
7	0.0512618
8	0.0298728

Some large negative correlations are noticeable, peaking at $k = 3$. This suggests that changes in the input gas rate are correlated with the residuals from the fitted output model and could therefore be used to improve the forecasts.

Pane Options

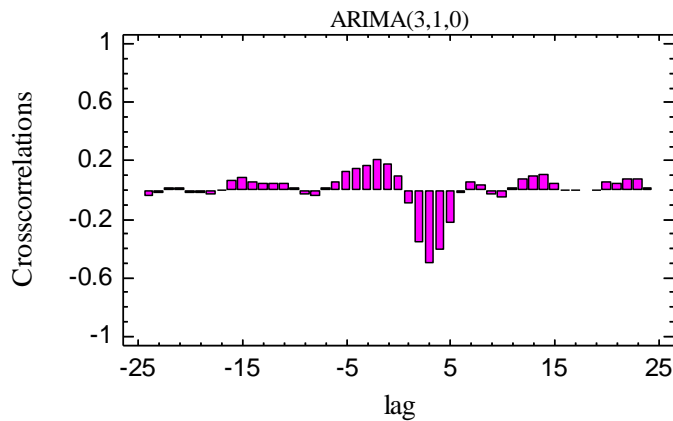


- **Second Time Series:** the observations for the X time series. Note the use of the *DIFF* operator to calculate the first differences of the *Input* column.
- **Number of Lags:** maximum lag k (both positive and negative) at which to calculate the crosscorrelations

Residual Crosscorrelation Plot

The *Residual Crosscorrelation Plot* displays the estimated crosscorrelations:

Estimated Crosscorrelations for Residuals with DIFF(Input)



Note the large negative correlation peaking at lag 3. This implies that the changes in the input are correlated with the residuals from the output model. They could thus be used to help forecast the output values.

Save Results

The following results can be saved to the datasheet:

1. *Data* – the original observations, together with any interpolated replacements for missing values.
2. *Adjusted data* – time series data after any adjustments have been made.
3. *Forecasts* – forecasted values within and beyond the sampling period.
4. *Upper forecast limits* – upper probability limits for the forecasts.
5. *Lower forecast limits* – lower probability limits for the forecasts.
6. *Residuals* – one-step ahead forecast errors.
7. *Autocorrelations* – residual autocorrelations.
8. *Partial autocorrelations* – residual partial autocorrelations.
9. *Crosscorrelations* – crosscorrelations between the residuals and a second time series.
10. *Residual periodogram ordinates* – calculated periodogram ordinates for the residuals.
11. *Fourier frequencies* – Fourier frequencies corresponding to the residual periodogram ordinates.

Calculations

Error Statistics – validation period

RMSE = root mean squared error

$$RMSE = \sqrt{\frac{\sum_{i=1}^m e_{n+i}^2}{m}} \quad (41)$$

MAPE = mean absolute percentage error

$$MAPE = 100 \frac{\sum_{i=1}^m |e_{n+i} / Y_{t+i}|}{m} \% \quad (42)$$

MAE = mean absolute error

$$MAE = \frac{\sum_{i=1}^m |e_{n+i}|}{m} \quad (43)$$

ME = mean error

$$ME = \frac{\sum_{i=1}^m e_{n+i}}{m} \quad (44)$$

MPE = mean percentage error

$$MPE = 100 \frac{\sum_{i=1}^m \frac{e_{n+i}}{Y_{n+i}}}{m} \% \quad (45)$$

Variance function for forecasts

Random walk model

$$\hat{V}(k) = k\hat{\sigma}_a \quad (46)$$

Mean model

$$\hat{V}(k) = \hat{\sigma}_a \left(1 + \frac{1}{n} \right) \quad (47)$$

Moving average model

$$\hat{V}(k) = \hat{\sigma}_a \left(1 + \frac{1}{c} \right) \quad (48)$$

Simple Exponential Smoothing

The variance function is determined from the equivalent ARIMA(0,1,1) model.

$$\hat{V}(k) = \hat{\sigma}_a \left(1 + (k-1)\alpha^2 \right) \quad (49)$$

Brown's Linear Exponential Smoothing

The variance function is determined from the equivalent ARIMA(0,2,2) model.

$$\hat{V}(k) = \hat{\sigma}_a \left(1 + (k-1)\lambda_0^2 + \frac{k(k-1)(2k-1)\lambda_1^2}{6} + \lambda_0\lambda_1k(k-1) \right) \quad (50)$$

where $\lambda_0 = \alpha(2-\alpha)$ and $\lambda_1 = \alpha^2$

Brown's Quadratic Exponential Smoothing

The variance function is determined from the equivalent ARIMA(0,3,3) model.

Holt's Linear Exponential Smoothing

The variance function is determined from the equivalent ARIMA(0,2,2) model.

$$\hat{V}(k) = \hat{\sigma}_a \left(1 + (k-1)\lambda_0^2 + \frac{k(k-1)(2k-1)\lambda_1^2}{6} + \lambda_0\lambda_1k(k-1) \right) \quad (51)$$

where $\lambda_0 = \alpha$ and $\lambda_1 = \alpha\beta$

Winter's Exponential Smoothing

The variance function is determined from the equivalent ARIMA(0,1,s+1)(0,1,0)_s model.

Trend models

Forecast limits are calculated from regression formulas for predicting a new observation at time $t = n + m + k$, including use of Student's t distribution with the appropriate number of degrees of freedom.

ARIMA Models

Calculated following the methods of Box, Jenkins and Reinsel (1994), which involves finding the Ψ function to express the observation at time t in terms of current and previous shocks.