Chapter 467

Exponential Smoothing – Trend & Seasonal

Introduction

This module forecasts seasonal series with upward or downward trends using the Holt-Winters exponential smoothing algorithm. Two seasonal adjustment techniques are available: additive and multiplicative.

Additive Seasonality

Given observations $X_1, X_2, \cdots, X_t$ of a time series, the Holt-Winters additive seasonality algorithm computes an evolving trend equation with a seasonal adjustment that is additive. *Additive* means that the amount of the adjustment is constant for all levels (average value) of the series.

The forecasting algorithm makes use of the following formulas:

$$a_t = \alpha\left(X_t - F_{t-s}\right) + (1-\alpha)\left(a_{t-1} + b_{t-1}\right)$$

$$b_t = \beta\left(a_t - a_{t-1}\right) + (1-\beta)b_{t-1}$$

$$F_t = \gamma\left(X_t - a_t\right) + (1-\gamma)F_{t-s}$$

Here α, β, and γ are smoothing constants which are between zero and one. Again, a_t gives the y-intercept (or level) at time *t*, while b_t is the slope at time *t*. The letter *s* represents the number of periods per year, so the quarterly data is represented by s = 4 and monthly data is represented by s = 12.

The forecast at time *T* for the value at time *T+k* is $a_T + b_T k + F_{[(T+k-1)/s]+1}$. Here *[(T+k-1)/s]* is means the remainder after dividing *T+k-1* by *s*. That is, this function gives the season (month or quarter) that the observation came from.

Multiplicative Seasonality

Given observations $X_1, X_2, \cdots, X_t$ of a time series, the Holt-Winters multiplicative seasonality algorithm computes an evolving trend equation with a seasonal adjustment that is multiplicative. *Multiplicative* means that the amount of the adjustment is varies with the level (average value) of the series. Note that the nature of most economic time series make the multiplicative model more popular than the additive model.

The forecasting algorithm makes use of the following formulas:

$$a_t = \alpha\left(X_t / F_{t-s}\right) + (1-\alpha)\left(a_{t-1} + b_{t-1}\right)$$

$$b_t = \beta\left(a_t - a_{t-1}\right) + (1-\beta)b_{t-1}$$

$$F_t = \gamma\left(X_t / a_t\right) + (1-\gamma)F_{t-s}$$

Here α, β, and γ are smoothing constants which are between zero and one. Again, a_t gives the y-intercept (or level) at time *t*, while b_t is the slope at time *t*. The letter *s* represents the number of periods per year, so the quarterly data is represented by s = 4 and monthly data is represented by s = 12.

The forecast at time *T* for the value at time *T+k* is $(a_T + b_T k)F_{[(T+k-1)/s]+1}$. Here *[(T+k-1)/s]* is means the remainder after dividing *T+k-1* by *s*. That is, this function gives the season (month or quarter) that the observation came from.

Smoothing Constants

Notice that the *smoothing constants* determines how fast the weights of the series decays. The values may be chosen either subjectively or objectively. Values of a smoothing constant near one put almost all weight on the most recent observations. Values of a smoothing constant near zero allow the distant past observations to have a large influence.

Note that α is associated with the level of the series, β is associated with the trend, and γ is associated with the seasonality factors.

When selecting the smoothing constant *subjectively*, you use your own experience with this, and similar, series. Also, specifying the smoothing constant yourself lets you tune the forecast to your own beliefs about the future of the series. If you believe that the mechanism generating the series has recently gone through some fundamental changes, use a smoothing constant value of 0.9 which will cause distant observations to be ignored. If, however, you think the series is fairly stable and only going through random fluctuations, use a value of 0.1.

To select the value of the smoothing constants *objectively*, you search for values that are best in some sense. Our program searches for that values that minimize the size of the combined forecast errors of the currently available series. Three methods of summarizing the amount of error in the forecasts are available: the mean square error (MSE), the mean absolute error (MAE), and the mean absolute percent error (MAPE). The forecast error is the difference between the forecast of the current period made at the last period and the value of the series at the current period. This is written as

$$e_t = X_t - F_{t-1}$$

Using this formulation, we can define the three error-size criterion as follows:

$$MSE = \frac{1}{n}\sum e_t^2$$

$$MAE = \frac{1}{n}\sum |e_t|$$

$$MAPE = \frac{100}{n}\sum \left|\frac{e_t}{X_t}\right|$$

To find the value of the smoothing constants objectively, we select one of these criterion and search for those values of α and β that minimize this function. The program conducts a search for the appropriate values using an efficient grid-searching algorithm.

Initial Values

Winters method requires initialization since the forecast for period one requires the forecast at period zero, which we do not, by definition, have. It also requires the seasonal adjustment factors. Several methods have been proposed for generating starting values. NCSS uses the initialization method described in Bowerman and O'Connell (1993).

Relationship to ARIMA Method

The multiplicative seasonal adjustment model does not have an ARIMA counterpart, while the additive model does.

Assumptions and Limitations

These algorithms are useful for forecasting seasonal time series with (local or global) trend.

Data Structure

The data are entered in a single variable.

Missing Values

When missing values are found in the series, they are either replaced or omitted. The replacement value is the average of the nearest observation in the future and in the past or the nearest non-missing value in the past.

If you do not feel that this is a valid estimate of the missing value, you should manually enter a more reasonable estimate before using the algorithm. These missing value replacement methods are particularly poor for seasonal data. We recommend that you replace missing values manually before using the algorithm.

Procedure Options

This section describes the options available in this procedure.

Variables Tab

Specify the variable(s) on which to run the analysis.

Time Series Variables

Time Series Variable(s)

Specify the variable(s) on which to run the analysis. A separate analysis will be conducted for each variable listed.

Use Logarithms

Specifies that the log (base 10) transformation should be applied to the values of the variable. The forecasts are converted back to there original metric before display.

Missing Values

Choose how missing (blank) values are processed.

The algorithm used in this procedure cannot tolerate missing values since each row is assumed to represent the next point in a time sequence. Hence, when missing values are found, they must be removed either by imputation (filling in with a reasonable value) or by skipping the row and pretending it does not exist.

Whenever possible, we recommend that you replace missing values manually.

Here are the available options.

Average the Adjacent Values

Replace the missing value with the average of the nearest values in the future (below) and in the past (above).

Carry the Previous Value Forward

Replace the missing value with the first non-missing value immediately above (previous) this value.

Omit Row from Calculations

Ignore the row in all calculations. Analyze the data as if the row was not on the database.

Forecasting Options

Number of Forecasts

This option specifies the number of forecasts to be generated.

Seasonal Model Options

Seasonal Adjustment

Select either the Additive or Multiplicative adjustment scheme.

Seasonality Options

Seasons

Specify the number of seasons per year in the series. Use '4' for quarterly data or '12' for monthly data.

First Season

Specify the first season of the series. This value is used to format the reports and plots. For example, if you have monthly data beginning with March, you would enter a '3' here.

First Year

Specify the first year of the series. This value is used to format the reports and plots.

Smoothing Constant Search Options

Search Method

This option specifies whether a search is conducted for the best values of the smoothing constants and what the criterion for the search will be.

- **Specified Value**
  No search is conducted. The values of the smoothing constants given in the next options are used.

- **Search on MSE**
  A search is conducted to find the values of the smoothing constants that minimize MSE.

- **Search on MAE**
  A search is conducted to find the values of the smoothing constants that minimize MAE.

- **Search on MAPE**
  A search is conducted to find the values of the smoothing constants that minimize MAPE.

Smoothing Constant Search Options – Pre-Specified Smoothing Constants

Alpha Smoothing Constant

When the Search Method is set to Specified Value, this option specifies the value of alpha. Alpha is the smoothing constant for the level of the series. The limits of this value are zero and one. Usually, a value between 0.1 and 0.3 are used. As the value gets closer to one, more and more weight is given to recent observations.

Beta Smoothing Constant

When the Search Method is set to Specified Value, this option specifies the value of beta. Beta is the smoothing constant for the trend. The limits of this value are zero and one. Usually, a value between 0.1 and 0.3 are used. As the value gets closer to one, more and more weight is given to recent observations.

Gamma Smoothing Constant

When the Search Method is set to Specified Value, this option specifies the value of gamma. Gamma is the smoothing constant for the seasonal factors. The limits of this value are zero and one. Usually, a value between 0.1 and 0.3 are used. As the value gets closer to one, more and more weight is given to recent observations.

Reports Tab

The following options control which reports are displayed.

Select Reports

Summary Report

This option specifies whether the indicated report is displayed.

Forecast Report

This option specifies which parts of the series are listed on the numeric reports: the original data and forecasts, just the forecasts, or neither.

Report Options

Precision

Specify the precision of numbers in the report. Single precision will display seven-place accuracy, while the double precision will display thirteen-place accuracy. Note that all reports are formatted for single precision only.

Variable Names

Specify whether to use variable names or (the longer) variable labels in report headings.

Page Title

Specify a title to be shown at the top of the reports.

Plots Tab

This section controls the forecast and residual plots.

Select Plots

Forecast Plot - Residual Plot

Each of these options specifies whether the indicated plot is displayed. Click the plot format button to change the plot settings.

Horizontal Axis Variable

Horizontal Variable

This option controls the spacing on the horizontal axis when missing or filtered values occur.

Your choices are

Actual Row Number

Use the actual row number of each row from the dataset along the horizontal axis.

Constructed Date

Construct a date value from the sequence (relative row) number and the *Seasonality Options* settings. Any missing or filtered values are skipped when forming the sequence number.

Storage Tab

The forecasts and residuals may be stored on the current dataset for further analysis. These options let you designate which statistics (if any) should be stored by designating which columns should receive the statistics.

Note that existing data is replaced. Be careful that you do not specify columns that contain important data.

Data Storage Columns

Forecasts

The forecasts are stored in this column.

Residuals

The residuals are stored in this column.

Example 1 – Trend & Seasonal Exponential Smoothing

This section presents an example of how to generate forecasts of a series using Winters multiplicative seasonal model. The data in the Sales dataset will be used. We will forecast the values of the Sales variable for the next twelve months.

You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the Exponential Smoothing – Trend / Seasonal window.

1 Open the Sales dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Sales.NCSS**.
- Click **Open**.

2 Open the Exponential Smoothing – Trend / Seasonal window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Exponential Smoothing - Trend/Seasonal** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- On the Exponential Smoothing – Trend / Seasonal window, select the **Variables tab**.
- Double-click in the **Time Series Variable** box. This will bring up the variable selection window.
- Select **Sales** from the list of variables and then click **Ok**.
- Enter **1970** in the **First Year** box.

4 Specify the reports.

- On the Exponential Smoothing – Trend / Seasonal window, select the **Reports tab**.
- Select **Data and Forecasts** in the **Forecast Report** list box.

5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

Forecast Summary Section

Forecast Summary Section

Variable	Sales
Number of Rows	144
Missing Values	None
Mean	174.2847
Pseudo R-Squared	0.980145
Mean Square Error	16.10279
Mean |Error|	3.114085
Mean |Percent Error|	1.786407
Forecast Method	Winter's with multiplicative seasonal adjustment.
Search Iterations	532
Search Criterion	Mean Square Error
Alpha	0.3496813
Beta	5.572607E-05
Gamma	6.69852E-11
Intercept (A)	138.3665
Slope (B)	0.5871831
Season 1 Factor	0.9028255
Season 2 Factor	0.8556558
Season 3 Factor	0.9714928
Season 4 Factor	0.997539
Season 5 Factor	1.028944
Season 6 Factor	1.028335
Season 7 Factor	0.9977484
Season 8 Factor	1.005503
Season 9 Factor	0.9752801
Season 10 Factor	1.023642
Season 11 Factor	1.007328
Season 12 Factor	1.205706

This report summarizes the forecast equation.

Variable

The name of the variable for which the forecasts are generated.

Number of Rows

The number of rows that were in the series. This is provided to allow you to double-check that the correct series was used.

Missing Values

If missing values were found, this option lists the method used to estimate them.

Mean

The mean of the variable across all time periods.

Pseudo R-Squared

This value generates a statistic that acts like the R-Squared value in multiple regression. A value near zero indicates a poorly fitting model, while a value near one indicates a well fitting model. The statistic is calculated as follows:

$$R^2 = 100\left(1 - \frac{SSE}{SST}\right)$$

where SSE is the sum of square residuals and SST is the total sum of squares after correcting for the mean.

Mean Square Error

The average squared residual (MSE) is a measure of how closely the forecasts track the actual data. The statistic is popular because it shows up in analysis of variance tables. However, because of the squaring, it tends to exaggerate the influence of outliers (points that do not follow the regular pattern).

Mean |Error|

The average absolute residual (MAE) is a measure of how closely the forecasts track the actual data without the squaring.

Mean |Percent Error|

The average percent absolute residual (MAPE) is a measure of how closely the forecasts track the actual data put on a percentage basis.

Forecast Method

This line shows which of the two possible seasonal adjustment algorithms was selected.

Search Iterations

This line shows how many iterations were needed to find the best values for the smoothing constants.

Search Criterion

If a search was made to find the best values of the smoothing constants, this row gives the criterion used during the search.

Alpha

The value of the smoothing constant alpha that was used to generate the forecasts.

Beta

The value of the smoothing constant beta that was used to generate the forecasts.

Gamma

The value of the smoothing constant gamma that was used to generate the forecasts.

Intercept (A)

The value of the y-intercept for time period one!

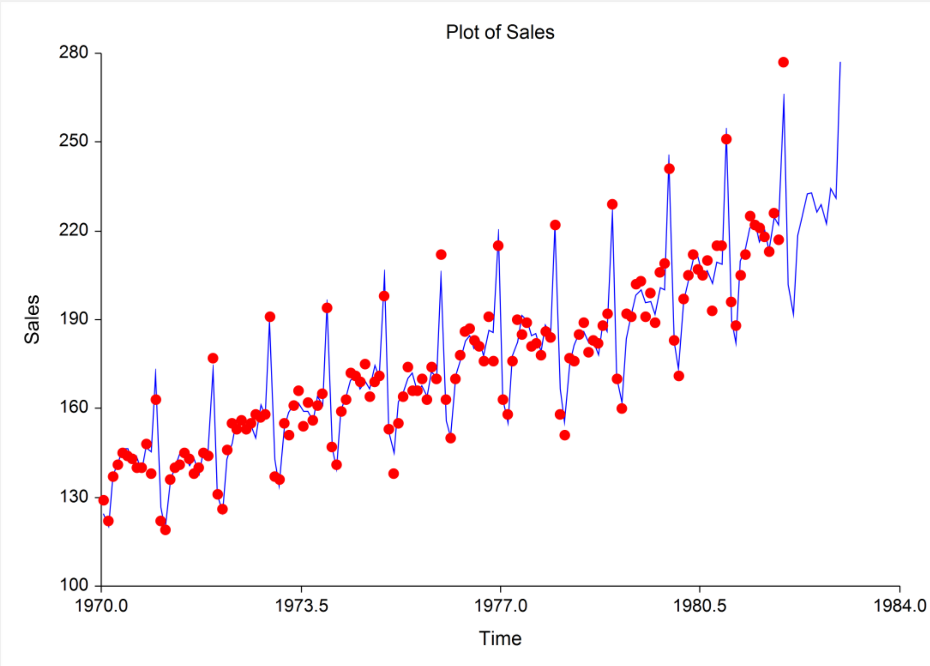
Slope (B)

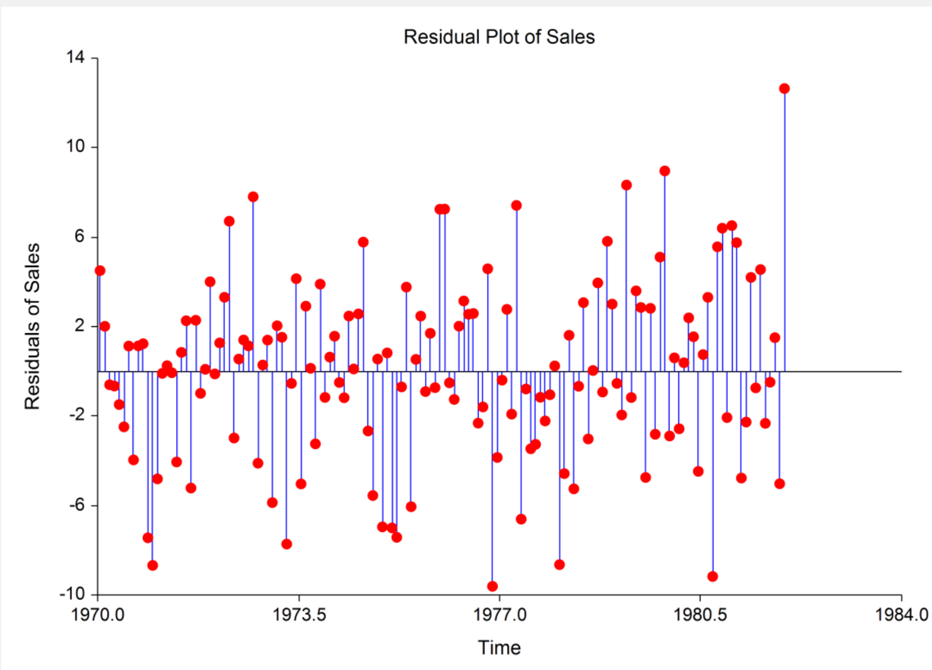
The value of the slope.

Season (1-12) Factor

The values of the multiplicative seasonal factors.

Forecast and Residuals Plots





Forecast Plot

The forecast plot lets you analyze how closely the forecasts track the data. The plot also shows the forecasts at the end of the data series.

Residual Plot

This plot lets you analyze the residuals themselves. You are looking for patterns, outliers, or any other information that may help you improve the forecasting model. The first thing to compare is the scale of the Residual Plot versus the scale of the Forecast Plot. If your forecasting algorithm is working well, the vertical scale of the Residual Plot will be much less than the scale of the Forecast Plot.

Forecasts Section

Forecasts Section

Row No.	Date	Forecast Sales	Actual Sales	Residuals
1	1970 1	124.4976	129	4.502408
2	1970 2	119.9876	122	2.012394
3	1970 3	137.6008	137	-0.600761
4	1970 4	141.66	141	-0.6600129
5	1970 5	146.486	145	-1.486042
6	1970 6	146.4839	144	-2.483881
7	1970 7	141.8699	143	1.130069
8	1970 8	143.9612	140	-3.961171
9	1970 9	138.8632	140	1.136814
10	1970 10	146.7673	148	1.232683
11	1970 11	145.4439	138	-7.443867
12	1970 12	171.6791	163	-8.679084
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
140	1981 8	220.3259	218	-2.325934
141	1981 9	213.4872	213	-0.4872138
142	1981 10	224.4956	226	1.504373
143	1981 11	222.0268	217	-5.026803
144	1981 12	264.3555	277	12.64447
145	1982 1	201.7888		
146	1982 2	191.7484		
147	1982 3	218.2773		
148	1982 4	224.7152		
149	1982 5	232.394		
150	1982 6	232.8604		
151	1982 7	226.5199		
152	1982 8	228.8708		
153	1982 9	222.5643		
154	1982 10	234.2018		
155	1982 11	231.0607		
156	1982 12	277.2727		

This section shows the values of the forecasts, the dates, the actual values, and the residuals.