# A11Wb Excel ETS seasonal exponential smoothing

The latest versions of Excel have a new forecasting method incorporated. This is the Excel version of the so-called state space models. These models are just a more advanced version of the conventional exponential smoothing methods. Imagine that you have a model where the Trend and Seasonal components can be:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Seasonal** | | |
| **Trend** | N (None) | A (Additive) | M (Multiplicative) |
| N (None) | N, N | N, A | N, M |
| A (Additive) | A, N | A, A | A, M |
| Ad (Additive damped) | Ad, N | Ad, A | Ad, M |
| M (Multiplicative) | M, N | M, A | M, M |
| Md (Multiplicative damped) | Md, N | Md, A | Md, M |

Table W11.1 Various exponential smoothing models

The methods we already covered in previous chapters are:

1. NN (the simple exponential smoothing),
2. AN (Holt‘s linear method),
3. AA (additive Holt-Winters method,
4. AM (multiplicative Holt-Winters method).

The exponential notation we have used so far is based on the so-called recursive formulae for exponential smoothing. The alternative class of equations for the same models can be expressed in the error correction format. This is the format that state-space models use.

What is typical for all state-space models is that each model consists of two components. The first one is the measurement equation that describes the observed data. The second one is the transition equations that describes how the unobserved components (that is the trend, level and the seasonal component) change the state over time. Clearly, this is the reason why this approach is called the state-space modelling. There are a variety of state space models, and they are often called the ETS models (**E** stands for error, **T** for trend and **S** for Seasonality). Excel implemented only one type of these models, which is the A, A, A model. This means that Trend is Additive, the Slope is Additive, and the Seasonal component is Additive, which is effectively an additive Holt-Winters method.

To keep this explanation simple, we will not go into any related equations. We will just show how Excel implemented the new =FORECAST.ETS() function to illustrate the application.

**Example W11.1**

We are again using the same time series as in textbook Examples 11.4-6, which is the quarterly CO2 emissions in ppm measured at the Mauna Loa observatory in Hawaii from Q1 2013 to Q4 2017. Figure W10.1 contains our data.

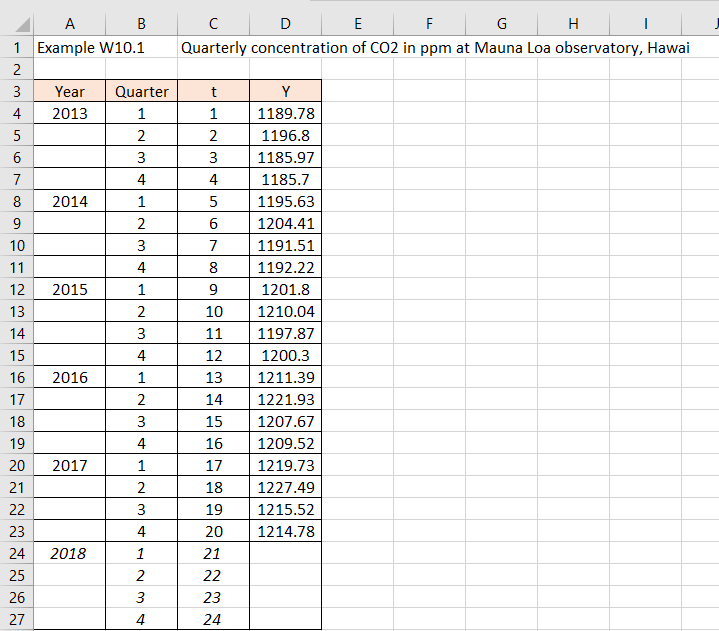


Figure W11.1

If we go to Excel Data tab, under the Forecast section (see Figure W10.2), this is where the algorithm that Microsoft deployed for this function resides. This means you do not have to use the function manually, it is used in the same way as many Data Analysis algorithms, but for some reason it is not placed there, but in the Data tab.

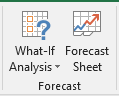


Figure W11.2 Excel Forecast tab

Highlight the cells C3:D23 from Figure W11.1 (the periods and the time series values) and click on the Forecast Sheet as in Figure W10.2. The dialogue box will immediately open as in Figure W11.3.

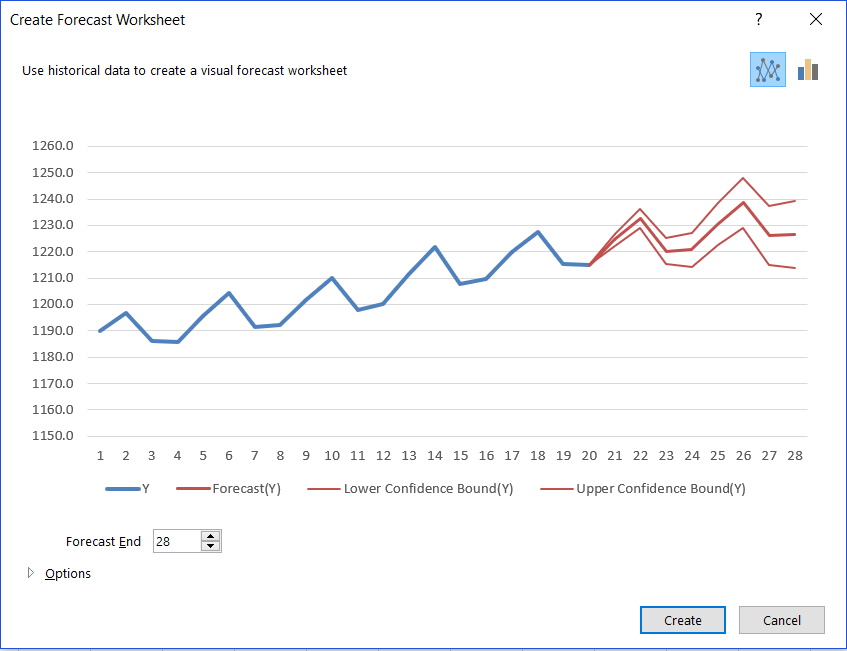


Figure W11.3

At the bottom of the dialogue box there is an Option button, and after we clicked on it, the box will now look as in Figure W11.4.

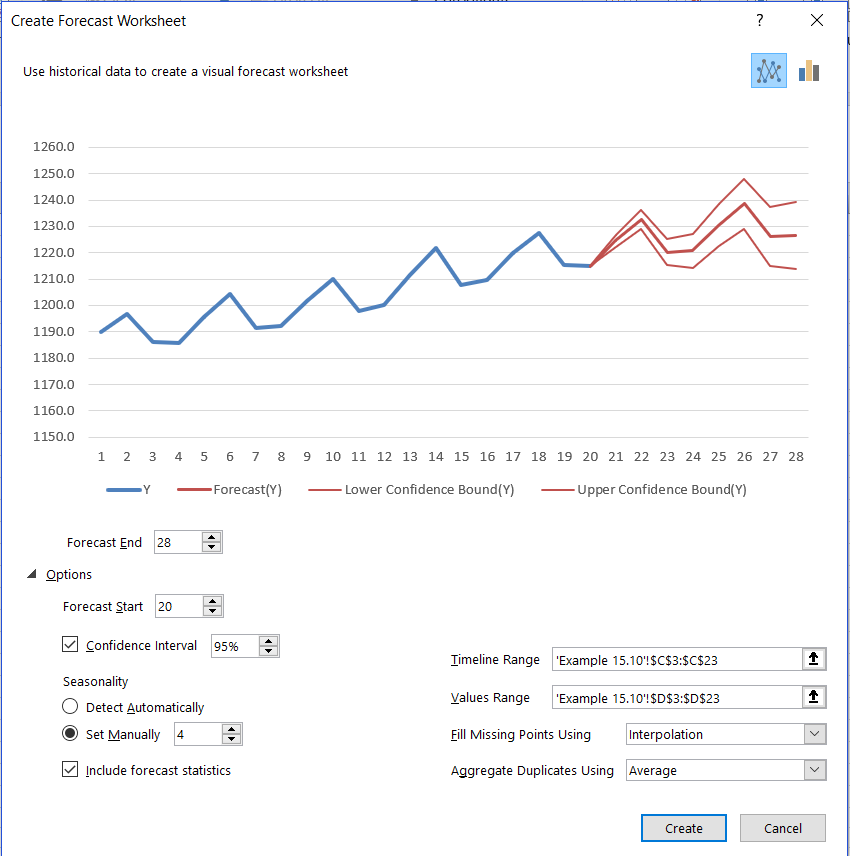


Figure W11.4

Our timeline is in C3:C23 and the actual data in D3:D23. The forecast will start at period 20 and finish at period 28, and even before you click on anything, Excel is showing you the graph of everything, including the 95% confidence interval for the forecasts, as per the dialogue box.

There are a few other buttons on this dialogue box that need brief explanation. On the right-hand side, we have Fill Missing Points and the default is the Interpolations method. There are a few other methods available, but if you do not have gaps in your time series, which you should not, then no need to worry about it. However, if you are using this as a forecasting tool on your dataset that came via data mining, and you do not have clear visibility of it, then this is a useful function.

Aggregate Duplicates Using is another option and it comes with the default option as Average. Like the Missing Points, this function is valuable for the datasets that come from data mining. If by mistake (or by design) you end up with two or more different values for the same time stamp, then this function will take their average to resolve the conflict and base all calculations on this average value.

On the left-hand side of the dialogue box, besides the forecasting horizon and the level of confidence interval, you can allow the algorithm to Detect Seasonality automatically, or by setting it manually, as we did to 4. And finally, you should click on “Include forecast statistics”, as this will give you some interesting details. When you press the Create button, Excel creates a new sheet with the printout as per Figure W11.5.

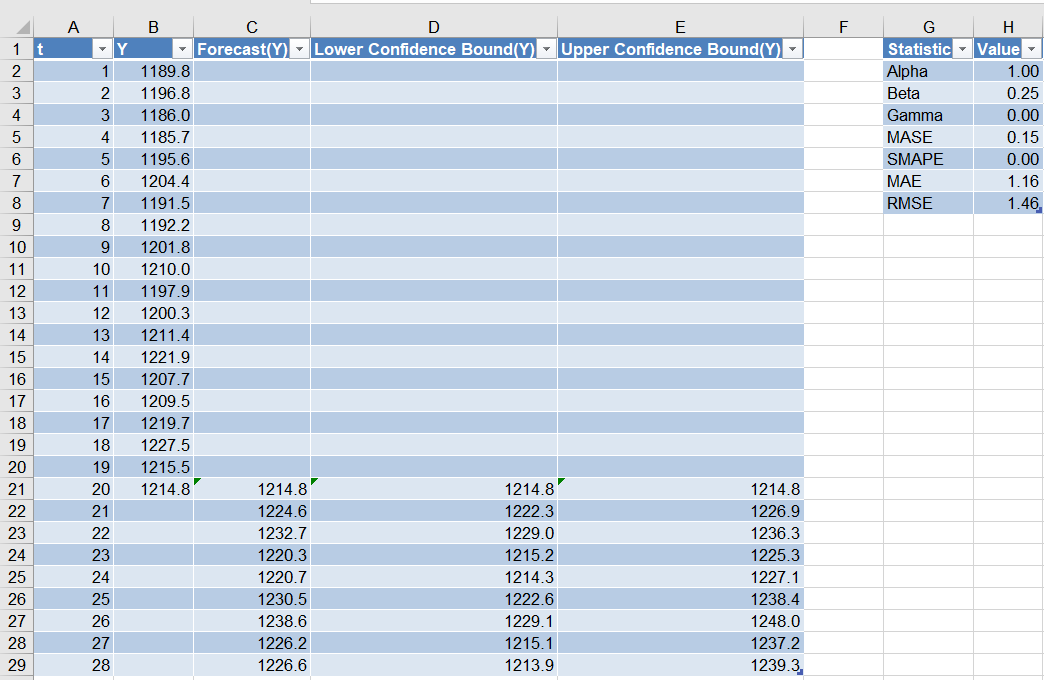


Figure W11.5

What we get are the future forecasts with the 95% confidence corridor and the graph (not shown here). In addition, because we included “Include forecast statistics”, we get a little table (cells G1:H8) that contains the three smoothing constants (alpha, beta and gamma), as well four different error measurements (MASE, SMAPE, MAE and RMSE). We have not covered MASE and SMAPE in this textbook but we did cover MAE (we call it MAD) and RMSE.

**>>**

Just one explanation, which might prove useful if you try to replicate this Excel approach by building your own ETS model manually. Rather than minimizing the sum of squared errors, or the mean squared error (MSE) to optimize the model, ETS uses as the criterion called the **maximization of the likelihood**. This word likelihood means different things, depending what model is used. In the case of additive error models (as here in Excel), maximizing the likelihood is the same as minimizing the MSE. For multiplicative error models, maximizing likelihood is the same as minimizing the sum of squared relative errors, but the errors are relative to one step ahead forecasts and not to observations in the time series.

If you look at the formulae in C22:E29, you will see that beside the function =FORECAST.ETS(), another new Excel function is deployed. This is =FORECAST.ETS.CONFINT().

**Excel solution**

t Cells A2:A29 Values

Y Cells B2:B21 Values

Forecast(Y) Cell C21 Formula:=B21

Cell C22 Formula: =FORECAST.ETS(A22,$B$2:$B$21,$A$2:$A$21,4,1)

Copy down from C22:C29

Lower Confidence Bound (Y)

Cell D21 Formula:=B21

Cell D22 Formula: =C22-FORECAST.ETS.CONFINT(A22,$B$2:$B$21,$A$2:$A$21,0.95,4,1)

Copy down D22:D29

Upper Confidence Bound (Y)

Cell E21 Formula:=B21

Cell E22 Formula: =C22+ FORECAST.ETS.CONFINT(A22,$B$2:$B$21,$A$2:$A$21,0.95,4,1)

Copy down E22:E29

The third new function =FORECAST.ETS.STAT() produces the values for the table in G2:H8 as illustrated in Figure W11.6.

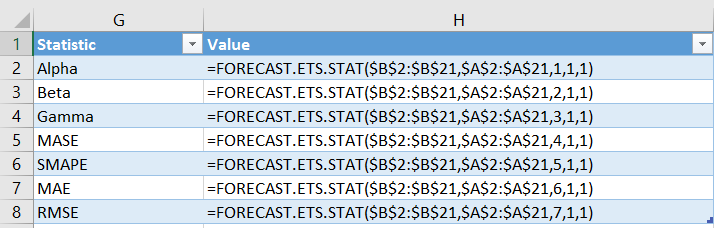


Figure W11.6 The application of the new Excel FORECAST.ETS functions

We can Copy/Paste the function in cell C22 backwards to C2. If we do this, we will get the values that this model has fitted to the actual observations. If we chart our data, together with the fitted model values, future forecasts and the prediction interval, we get a graph as in Figure W11.7. In fact, this graph was produced automatically by Excel and the only modifications we made was to include the historical model values and changed the prediction lines to dotted lines.

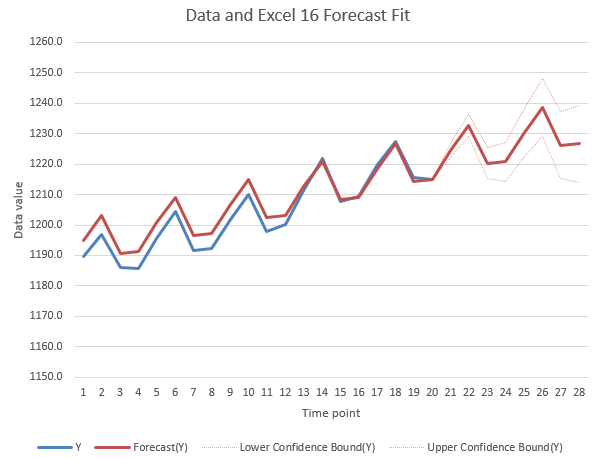


Figure W11.7

As we can see, this is very credible model. Although it did not fit the earlier data too well, towards the end it produces almost perfect fit values. If you compare the future forecasts with what we achieved in textbook Example 10.6, you will see that they are remarkably close.