Interpreting Measures of Error

Scale Dependent Errors

Scale dependent errors, such as mean error (ME) mean percentage error (MPE), mean absolute error (MAE) and root mean squared error (RMSE), are based on a set scale, which for us is our time series, and cannot be used to make comparisons that are on a different scale. For example, we wouldn't take these error values from a time series model of the sheep population in Scotland and compare it to corn production forecast in the United States.

- **Mean Error (ME)** shows the average of the difference between actual and forecasted values.
- **Mean Percentage Error (MPE)** shows the average of the percent difference between actual and forecasted values. Both the ME and MPE will help indicate whether the forecasts are biased to be disproportionately positive or negative.
- **Root Mean Squared Error (RMSE)** represents the sample standard deviation of the differences between predicted values and observed values. These individual differences are called residuals when the calculations are performed over the data sample that was used for estimation, and are called prediction errors when computed out-of-sample. This is a great measurement to use when comparing models as it shows how many deviations from the mean the forecasted values fall.
- **Mean Absolute Error (MAE)** takes the sum of the absolute difference from actual to forecast and averages them. It is less sensitive to the occasional very large error because it does not square the errors in the calculation.

Percentage Errors

Percentage errors, like MAPE, are useful because they are scale independent, so they can be used to compare forecasts between different data series, unlike scale dependent errors. The disadvantage is that it cannot be used if the series has zero values.

- **Mean Absolute Percentage Error (MAPE)** is also often useful for purposes of reporting, because it is expressed in generic percentage terms it will make sense even to someone who has no idea what constitutes a "big" error in terms of dollars spent or widgets sold.

Scale-Free Errors

Scale-free errors were introduced more recently to offer a scale-independent measure that doesn't have many of the problems of other errors like percentage errors.

- **Mean Absolute Scaled Error (MASE)** is another relative measure of error that is applicable only to time series data. It is defined as the mean absolute error of the model divided by the the mean absolute value of the first difference of the series. Thus, it measures the relative reduction in error compared to a naive model. Ideally its value will be significantly less than 1 but is relative to comparison across other models for the same

series. Since this error measurement is relative and can be applied across models, it is accepted as one of the best metrics for error measurement.
To read a bit more of the subject, see this 4-page paper by Rob Hyndman, Statistics professor at Monash University in Australia.

In Alteryx you can find these measures in the ARIMA or ETS tools labelled as In-sample error measures. Example below:

4 **Summary of Time Series Exponential Smoothing Model ETS**

5 Method:
ETS(M,N,M)

6 In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-272.3281196	12097.816118	9938.3494152	-0.2211276	3.5936211	0.4093906	0.0315054

7

When you forecast values using ARIMA and ETS and compare these forecasted results in the TS Compare tool you will find these measures in the Accuracy Measures section. Example below:

3 Accuracy Measures:

Model	ME	RMSE	MAE	MPE	MAPE	MASE	NA
ETS	-7417.239	10631.74	10073.75	-2.8168	3.8462	0.5013	NA
ARIMA	-21118.476	22645.43	21118.48	-8.1431	8.1431	1.051	NA

4