Time Series Forecasting and Testing

Modelling vs Forecasting

Until now we have

- used all the observed data
- created a model that tries to describe the time series

This is important if we want to understand the **generative mechanisms** of the data (i.e., what is the seasonal component, the trend, etc)

But we may also be interested in trying to **forecast and predict** future values!

CAMBRIDGE SPARK

Modelling vs Forecasting

Until now we have

- used all the observed data
- created a model that tries to describe the time series

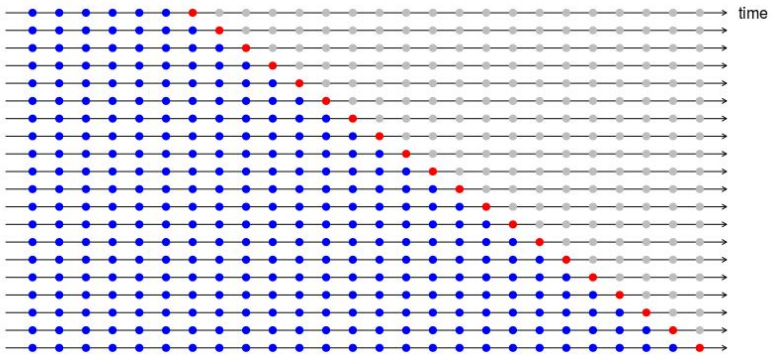
If we are interested in forecasting, we need a way to **assess the performance of the model**: one thing we could do is train on the first 90% of the time series and try using the model to generate values for the next 10%.

Similar to cross-validation: but we need to take extra care with time-series!

Cross-validating on time series

In this procedure:

- we need to account for the **chronological ordering of the data**
- the original time series is split multiple times in train and test set
- the training sets consist only of observations that occurred prior to the observations in the test set
- each iteration is called a *fold*



CAMBRIDGE SPARK

Hands-on session

time_series_testing.ipynb

CAMBRIDGE SPARK