Time series analysis: Exponential Smoothing

Forecasting accuracy

Let us define a forecasting error $e_t = Y_t - F_t$.
We may then define some forecasting accuracy measures:
Mean Error, Mean Absolute Error, Mean Squared Error

$$\text{ME} = \frac{1}{n}\sum_{t=1}^{n} e_t$$

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n} |e_t|$$

$$\text{MSE} = \frac{1}{n}\sum_{t=1}^{n} e_t^2.$$

Forecasting accuracy

The value of ME, MAE, MSE depend on the scale of data.
This makes difficult to compare different models.
We may define the percentage error and related measures.

$$\mathsf{PE}_t = \frac{Y_t - F_t}{Y_t} 100$$

$$\mathsf{MPE} = \frac{1}{n} \sum_{t=1}^{n} \mathsf{PE}_t$$

$$\mathsf{MAPE} = \frac{1}{n} \sum_{t=1}^{n} |\mathsf{PE}_t|$$

Short-term forecasting: simple exponential smoothing

The simple exponential smoothing is defined as

$$F_{t+1} = F_t + \alpha(Y_t - F_t)$$

where α is a constant term taking values between 0 e 1.
The new forecast F_{t+1} is the old forecast F_t with an adjustment for the error that occurred in the last forecast.

Short-term forecasting: simple exponential smoothing

An equivalent way to express the simple exponential smoothing is

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t$$

The new forecast F_{t+1} is a weighted average of the last observation, Y_t, and the last forecast, F_t.

Short-term forecasting: simple exponential smoothing

Why exponential smoothing?

$$\begin{aligned} F_{t+1} &= \alpha Y_t + (1-\alpha)[\alpha Y_{t-1} + (1-\alpha) F_{t-1}] \\ &= \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + (1-\alpha)^2 F_{t-1} \end{aligned}$$

so that we obtain

$$\begin{aligned} F_{t+1} &= \alpha Y_t + \alpha(1-\alpha) Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} \\ &\quad + \alpha(1-\alpha)^3 Y_{t-3} + ... + \alpha(1-\alpha)^{t-1} Y_1 + (1-\alpha)^t F_1 \end{aligned}$$

Short-term forecasting: simple exponential smoothing

Initialization of the process

$$F_2 = \alpha Y_1 + (1 - \alpha) F_1$$

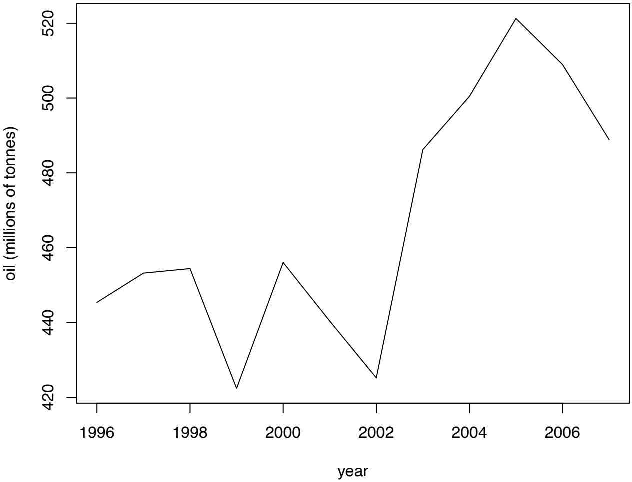
Since F_1 is not available, typically we use the first observation, $Y_1 = F_1$.

Short-term forecasting: simple exponential smoothing

- A crucial point in exponential smoothing concerns choosing a suitable value for α.
- A higher value for α is more sensitive to a change in the data structure, while a lower value generates a ‘flat’ forecast.
- A suitable selection for α is that minimizing the SSE.

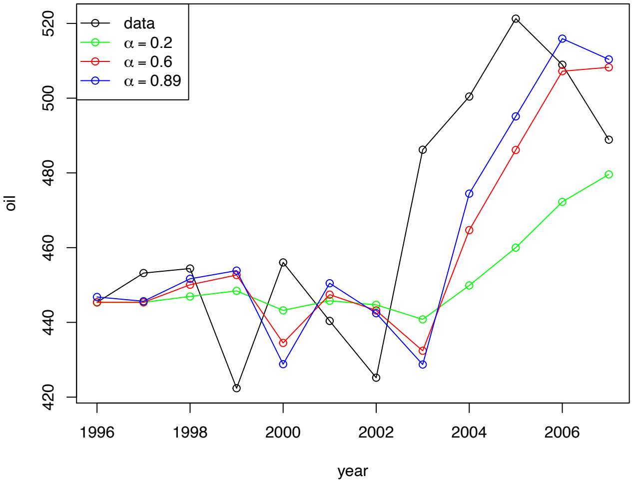
Example

Oil production in Saudi Arabia



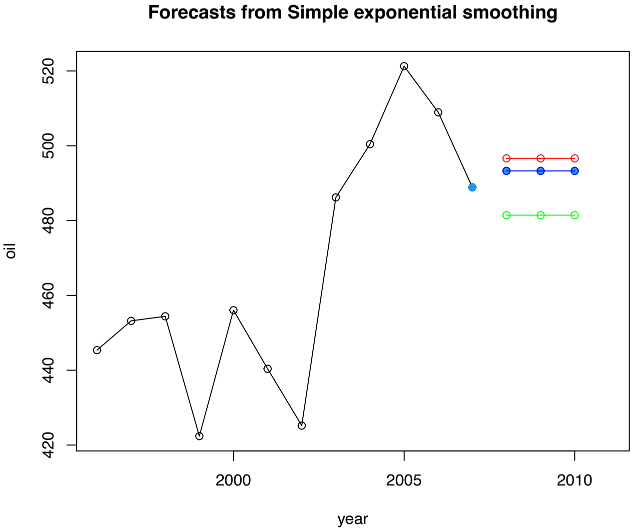
Example

Oil production in Saudi Arabia: prediction with $\hat{\alpha} = 0.2$, $\hat{\alpha} = 0.6$, $\hat{\alpha} = 0.89$



Example

Oil production in Saudi Arabia: forecasting with $\hat{\alpha} = 0.2$, $\hat{\alpha} = 0.6$, $\hat{\alpha} = 0.89$



Short-term forecasting: Holt's exponential smoothing

- The Holt's linear trend method is a useful extension to allow the forecasting of data with a trend.
- This method involves a forecast equation and two smoothing equations (one for the level and one for the trend)

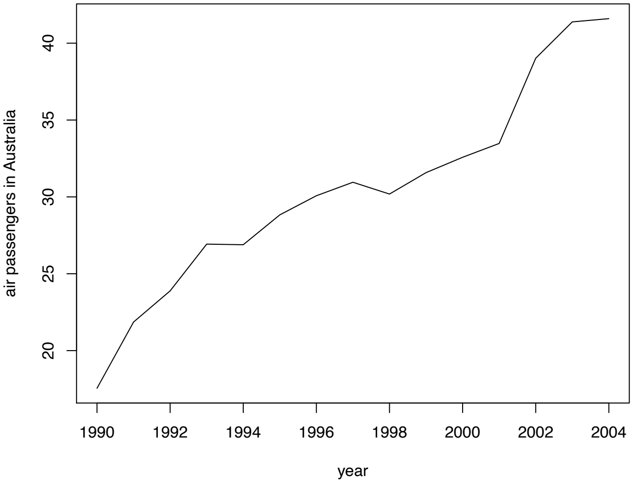
$$
\begin{aligned}
L_t &= \alpha Y_t + (1-\alpha)(L_{t-1} + b_{t-1}) \\
b_t &= \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1} \\
F_{t+m} &= L_t + b_t m
\end{aligned}
$$

Short-term forecasting: Holt's exponential smoothing

- This exponential smoothing is a double smoothing.
- The forecast function is no longer flat but trending.
- The m-step-ahead forecast is equal to the last estimated level plus times the last estimated trend value.
- Hence the forecasts are a linear function of m.

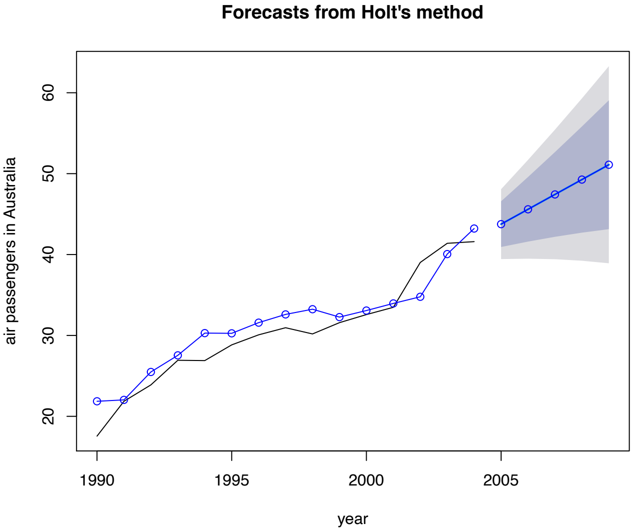
Example

Airline passengers in Australia



Example

Airline passengers in Australia: Forecasting with Holt's method



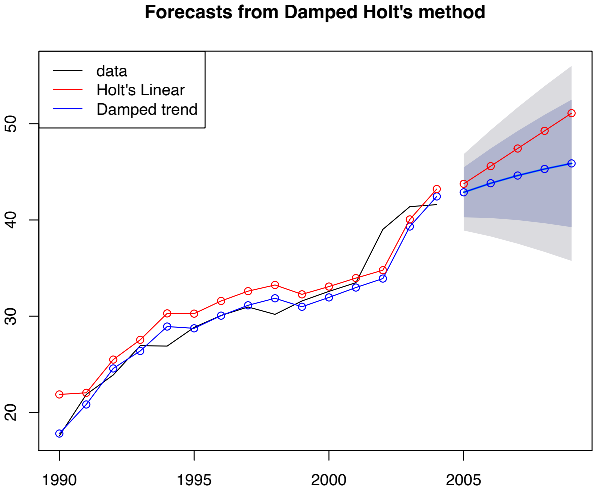
Short-term forecasting: Damped trend methods

- The forecasts generated by Holt's linear method display a constant trend (increasing or decreasing) indefinitely into the future.
- Empirical evidence indicates that these methods tend to over-forecast, especially for longer forecast horizons.
- A useful extension includes a damping parameter $0 < \phi < 1$

$$\begin{aligned} L_t &= \alpha Y_t + (1-\alpha)(L_{t-1} + \phi b_{t-1}) \\ b_t &= \beta(L_t - L_{t-1}) + (1-\beta)\phi b_{t-1} \\ F_{t+m} &= L_t + (\phi + \phi^2 + \cdots + \phi^m) b_t \end{aligned}$$

Example

Airline passengers in Australia: Forecasting with Holt's and Damped trend methods



Short-term forecasting: Holt-Winters' exponential smoothing

- If the data have no trend or seasonality, then exponential smoothing is appropriate
- If the data show a linear trend, Holt's linear method is appropriate
- But if the data are seasonal these methods cannot handle the problem well
- Holt-Winters' method is suitable when both trend and seasonality are present in the data

Short-term forecasting: Holt-Winters' exponential smoothing

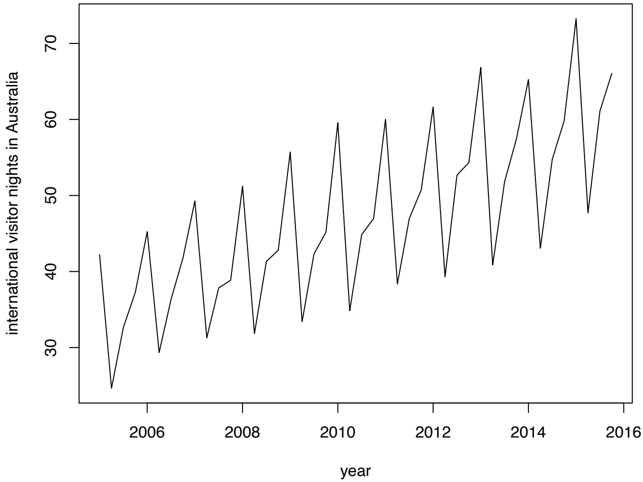
Holt-Winters' multiplicative seasonality

$$
\begin{aligned}
L_t &= \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1}) \\
b_t &= \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1} \\
S_t &= \gamma \frac{Y_t}{L_t} + (1-\gamma)S_{t-s} \\
F_{t+m} &= (L_t + b_t m)S_{t-s+m}
\end{aligned}
$$

Note: the value S_{t-s} is used because S_t cannot be calculated until L_t is known

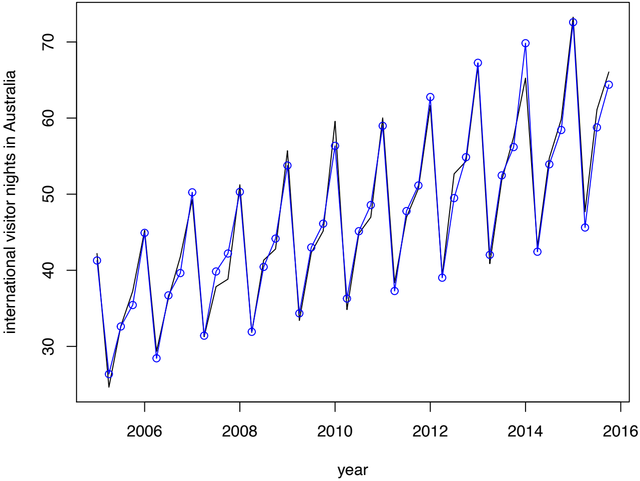
Example

International visitor nights in Australia



Example

International visitor nights in Australia: prediction with Holt-Winters'



Example

International visitor nights in Australia: forecasting with Holt-Winters'