



# Forecasting: principles and practice

Rob J Hyndman

1.4 Exponential smoothing

# Outline

- 1 Simple exponential smoothing
- 2 Trend methods
- 3 Lab session 7
- 4 Seasonal methods
- 5 Lab session 8
- 6 Taxonomy of exponential smoothing methods

# Simple methods

Time series  $y_1, y_2, \dots, y_T$ .

## Random walk forecasts

$$\hat{y}_{T+h|T} = y_T$$

# Simple methods

Time series  $y_1, y_2, \dots, y_T$ .

## Random walk forecasts

$$\hat{y}_{T+h|T} = y_T$$

## Average forecasts

$$\hat{y}_{T+h|T} = \frac{1}{T} \sum_{t=1}^T y_t$$

# Simple methods

Time series  $y_1, y_2, \dots, y_T$ .

## Random walk forecasts

$$\hat{y}_{T+h|T} = y_T$$

## Average forecasts

$$\hat{y}_{T+h|T} = \frac{1}{T} \sum_{t=1}^T y_t$$

- Want something in between that weights most recent data more highly.
- Simple exponential smoothing uses a weighted moving average with weights that decrease exponentially.

# Simple Exponential Smoothing

## Forecast equation

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \dots$$

where  $0 \leq \alpha \leq 1$ .

# Simple Exponential Smoothing

## Forecast equation

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \dots$$

where  $0 \leq \alpha \leq 1$ .

Weights assigned to observations for:				
Observation	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$
$y_T$	0.2	0.4	0.6	0.8
$y_{T-1}$	0.16	0.24	0.24	0.16
$y_{T-2}$	0.128	0.144	0.096	0.032
$y_{T-3}$	0.1024	0.0864	0.0384	0.0064
$y_{T-4}$	$(0.2)(0.8)^4$	$(0.4)(0.6)^4$	$(0.6)(0.4)^4$	$(0.8)(0.2)^4$
$y_{T-5}$	$(0.2)(0.8)^5$	$(0.4)(0.6)^5$	$(0.6)(0.4)^5$	$(0.8)(0.2)^5$

# Simple Exponential Smoothing

## Component form

Forecast equation  $\hat{y}_{t+h|t} = \ell_t$

Smoothing equation  $\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1}$

- $\ell_t$  is the level (or the smoothed value) of the series at time  $t$ .
- $\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$   
Iterate to get exponentially weighted moving average form.

## Weighted average form

$$\hat{y}_{T+1|T} = \sum_{j=0}^{T-1} \alpha(1 - \alpha)^j y_{T-j} + (1 - \alpha)^T \ell_0$$



# Optimisation

- Need to choose value for  $\alpha$  and  $\ell_0$
- Similarly to regression — we choose  $\alpha$  and  $\ell_0$  by minimising SSE:

$$\text{SSE} = \sum_{t=1}^T (y_t - \hat{y}_{t|t-1})^2.$$

- Unlike regression there is no closed form solution — use numerical optimization.

# Example: Oil production

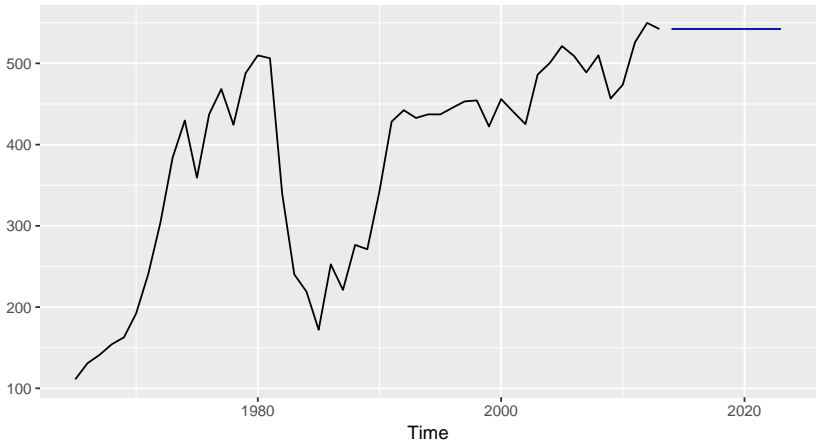
```
fc <- ses(oil, h=5)
summary(fc[["model"]])
```

```
## Simple exponential smoothing
##
## Call:
##   ses(y = oil, h = 5)
##
##   Smoothing parameters:
##     alpha = 0.9999
##
##   Initial states:
##     l = 110.8832
##
##   sigma:  49.05
##
##   AIC  AICc  BIC
## 576.2 576.7 581.8
```

# Example: Oil production

```
oil %>% ses(PI=FALSE) %>% autoplot
```

Forecasts from Simple exponential smoothing



# Outline

- 1 Simple exponential smoothing
- 2 Trend methods
- 3 Lab session 7
- 4 Seasonal methods
- 5 Lab session 8
- 6 Taxonomy of exponential smoothing methods

# Holt's linear trend

## Component form

Forecast  $\hat{y}_{t+h|t} = \ell_t + hb_t$

Level  $\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$

Trend  $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1},$

# Holt's linear trend

## Component form

Forecast  $\hat{y}_{t+h|t} = \ell_t + hb_t$

Level  $\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$

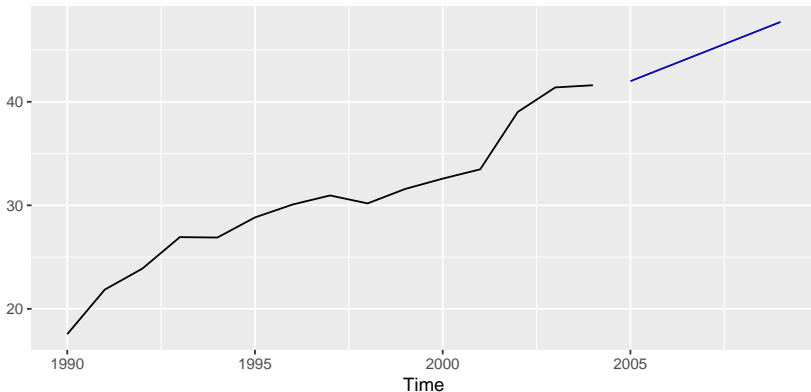
Trend  $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1},$

- Two smoothing parameters  $\alpha$  and  $\beta^*$   
( $0 \leq \alpha, \beta^* \leq 1$ ).
- $\ell_t$  level: weighted average between  $y_t$  and one-step ahead forecast for time  $t$ ,  
( $\ell_{t-1} + b_{t-1} = \hat{y}_{t|t-1}$ )
- $b_t$  slope: weighted average of  $(\ell_t - \ell_{t-1})$  and  $b_{t-1}$ , current and previous estimate of slope.

# Holt's method in R

```
window(ausair, start=1990, end=2004) %>%  
  holt(h=5, PI=FALSE) %>%  
  autoplot()
```

Forecasts from Holt's method



# Damped trend method

## Component form

$$\hat{y}_{t+h|t} = \ell_t + (\phi + \phi^2 + \dots + \phi^h)b_t$$

$$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}.$$



# Damped trend method

## Component form

$$\hat{y}_{t+h|t} = \ell_t + (\phi + \phi^2 + \dots + \phi^h)b_t$$

$$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}.$$

- Damping parameter  $0 < \phi < 1$ .
- If  $\phi = 1$ , identical to Holt's linear trend.
- As  $h \rightarrow \infty$ ,  $\hat{y}_{T+h|T} \rightarrow \ell_T + \phi b_T / (1 - \phi)$ .
- Short-run forecasts trended, long-run forecasts constant.

## Example: Sheep in Asia

```
livestock2 <- window(livestock, start=1970, end=2000)
fc1 <- ses(livestock2)
fc2 <- holt(livestock2)
fc3 <- holt(livestock2, damped = TRUE)
```

```
accuracy(fc1, livestock)
```

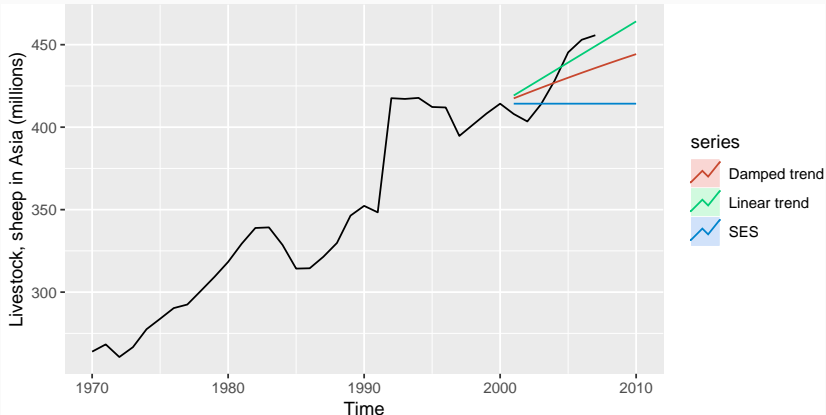
```
accuracy(fc2, livestock)
```

```
accuracy(fc3, livestock)
```

	SES	Linear trend	Damped trend
Test RMSE	25.46	11.88	14.73
Test MAE	20.38	10.71	13.30
Test MAPE	4.60	2.54	3.07
Test MASE	2.26	1.19	1.48

# Example: Sheep in Asia

```
autoplot(window(livestock, start=1970)) +  
  autolayer(fc1, series="SES", PI=FALSE) +  
  autolayer(fc2, series="Linear trend", PI=FALSE) +  
  autolayer(fc3, series="Damped trend", PI=FALSE) +  
  ylab("Livestock, sheep in Asia (millions)")
```



# Outline

- 1 Simple exponential smoothing
- 2 Trend methods
- 3 Lab session 7
- 4 Seasonal methods
- 5 Lab session 8
- 6 Taxonomy of exponential smoothing methods

# Lab Session 7

# Outline

- 1 Simple exponential smoothing
- 2 Trend methods
- 3 Lab session 7
- 4 Seasonal methods
- 5 Lab session 8
- 6 Taxonomy of exponential smoothing methods

# Holt-Winters additive method

Holt and Winters extended Holt's method to capture seasonality.

## Component form

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)}$$

$$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1})$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$$

$$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m},$$

- $k = \text{integer part of } (h - 1)/m$ . Ensures estimates from the final year are used for forecasting.
- Parameters:  $0 \leq \alpha \leq 1$ ,  $0 \leq \beta^* \leq 1$ ,  $0 \leq \gamma \leq 1 - \alpha$  and  $m = \text{period of seasonality (e.g. } m = 4 \text{ for quarterly data)}$ .

# Holt-Winters multiplicative method

For when seasonal variations are changing proportional to the level of the series.

## Component form

$$\hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)}.$$

$$\ell_t = \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(\ell_{t-1} + b_{t-1})$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$$

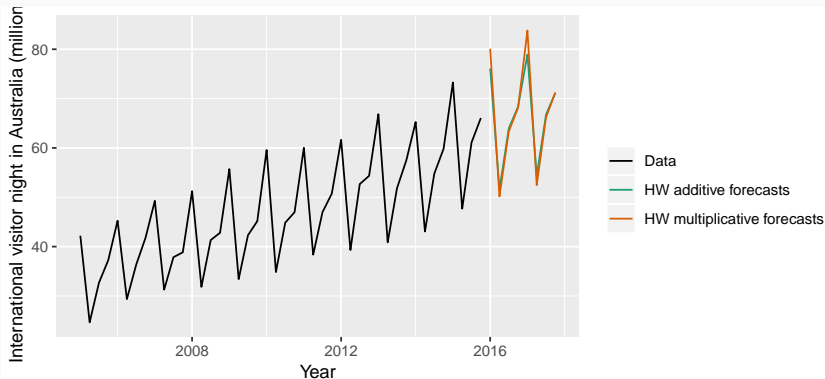
$$s_t = \gamma \frac{y_t}{(\ell_{t-1} + b_{t-1})} + (1 - \gamma)s_{t-m}$$

- $k$  is integer part of  $(h - 1)/m$ .
- **Additive:**  $s_t$  in absolute terms: within each year  $\sum_i s_i \approx 0$ .
- **Multiplicative:**  $s_t$  in relative terms: within each year  $\sum_i s_i \approx m$ .



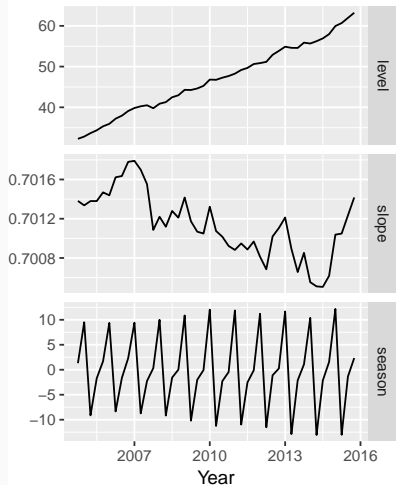
# Example: Visitor Nights

```
aust <- window(austourists,start=2005)
fc1 <- hw(aust,seasonal="additive")
fc2 <- hw(aust,seasonal="multiplicative")
```

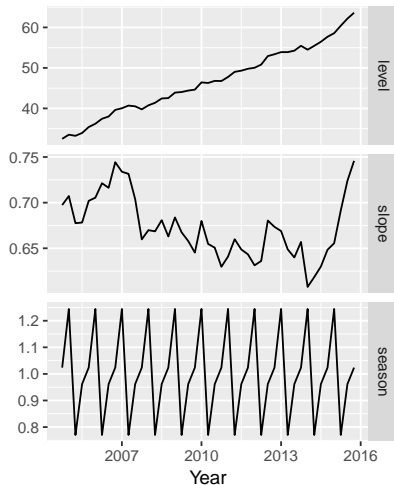


# Estimated components

Additive states



Multiplicative states



# Holt-Winters damped method

Often the single most accurate forecasting method for seasonal data:

$$\hat{y}_{t+h|t} = [\ell_t + (\phi + \phi^2 + \dots + \phi^h)b_t]s_{t+h-m(k+1)}$$

$$\ell_t = \alpha(y_t/s_{t-m}) + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}$$

$$s_t = \gamma \frac{y_t}{(\ell_{t-1} + \phi b_{t-1})} + (1 - \gamma)s_{t-m}$$

# Outline

- 1 Simple exponential smoothing
- 2 Trend methods
- 3 Lab session 7
- 4 Seasonal methods
- 5 Lab session 8
- 6 Taxonomy of exponential smoothing methods

# Lab Session 8

# Outline

- 1 Simple exponential smoothing
- 2 Trend methods
- 3 Lab session 7
- 4 Seasonal methods
- 5 Lab session 8
- 6 Taxonomy of exponential smoothing methods

# Exponential smoothing methods

Trend Component		Seasonal Component		
		N (None)	A (Additive)	M (Multiplicative)
N	(None)	(N,N)	(N,A)	(N,M)
A	(Additive)	(A,N)	(A,A)	(A,M)
A <sub>d</sub>	(Additive damped)	(A <sub>d</sub> ,N)	(A <sub>d</sub> ,A)	(A <sub>d</sub> ,M)

(N,N): Simple exponential smoothing

(A,N): Holt's linear method

(A<sub>d</sub>,N): Additive damped trend method

(A,A): Additive Holt-Winters' method

(A,M): Multiplicative Holt-Winters' method

(A<sub>d</sub>,M): Damped multiplicative Holt-Winters' method

# Recursive formulae

Trend	Seasonal		
	N	A	M
N	$\hat{y}_{t+h t} = \ell_t$	$\hat{y}_{t+h t} = \ell_t + s_{t-m+h_m^+}$	$\hat{y}_{t+h t} = \ell_t s_{t-m+h_m^+}$
	$\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1}$	$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)\ell_{t-1}$	$\ell_t = \alpha(y_t/s_{t-m}) + (1 - \alpha)\ell_{t-1}$
		$s_t = \gamma(y_t - \ell_{t-1}) + (1 - \gamma)s_{t-m}$	$s_t = \gamma(y_t/\ell_{t-1}) + (1 - \gamma)s_{t-m}$
A	$\hat{y}_{t+h t} = \ell_t + hb_t$	$\hat{y}_{t+h t} = \ell_t + hb_t + s_{t-m+h_m^+}$	$\hat{y}_{t+h t} = (\ell_t + hb_t)s_{t-m+h_m^+}$
	$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$	$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1})$	$\ell_t = \alpha(y_t/s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1})$
	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$
		$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}$	$s_t = \gamma(y_t/(\ell_{t-1} - b_{t-1})) + (1 - \gamma)s_{t-m}$
Ad	$\hat{y}_{t+h t} = \ell_t + \phi_h b_t$	$\hat{y}_{t+h t} = \ell_t + \phi_h b_t + s_{t-m+h_m^+}$	$\hat{y}_{t+h t} = (\ell_t + \phi_h b_t)s_{t-m+h_m^+}$
	$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$	$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$	$\ell_t = \alpha(y_t/s_{t-m}) + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$
	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}$	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}$	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}$
		$s_t = \gamma(y_t - \ell_{t-1} - \phi b_{t-1}) + (1 - \gamma)s_{t-m}$	$s_t = \gamma(y_t/(\ell_{t-1} - \phi b_{t-1})) + (1 - \gamma)s_{t-m}$



# R functions

- Simple exponential smoothing: no trend.

```
ses(y)
```

- Holt's method: linear trend.

```
holt(y)
```

- Damped trend method.

```
holt(y, damped=TRUE)
```

- Holt-Winters methods

```
hw(y, damped=TRUE, seasonal="additive")
```

```
hw(y, damped=FALSE, seasonal="additive")
```

```
hw(y, damped=TRUE, seasonal="multiplicative")
```

```
hw(y, damped=FALSE, seasonal="multiplicative")
```

- Combination of no trend with seasonality not possible using these functions.