

Exponentially weighted forecasts

FORECASTING IN R



Rob J. Hyndman

Professor of Statistics at Monash
University

Simple exponential smoothing

Forecasting Notation:

$\hat{y}_{t+h|t}$ = point forecast of \hat{y}_{t+h} given data y_1, \dots, y_t

Forecast Equation:

$$\hat{y}_{t+h|t} = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots$$

where $0 \leq \alpha \leq 1$

Simple exponential smoothing

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

- ℓ_t is the level (or the smoothed value) of the series at time t
- We choose α and ℓ_0 by minimizing SSE:

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

Example: oil production

```
oildata <- window(oil, start = 1996)      # Oil Data  
fc <- ses(oildata, h = 5)                 # Simple Exponential Smoothing  
summary(fc)
```

```
Forecast method: Simple exponential smoothing
```

```
Model Information:
```

```
Simple exponential smoothing
```

```
Call:
```

```
ses(y = oildata, h = 5)
```

```
Smoothing parameters:
```

```
alpha = 0.8339
```

```
Initial states:
```

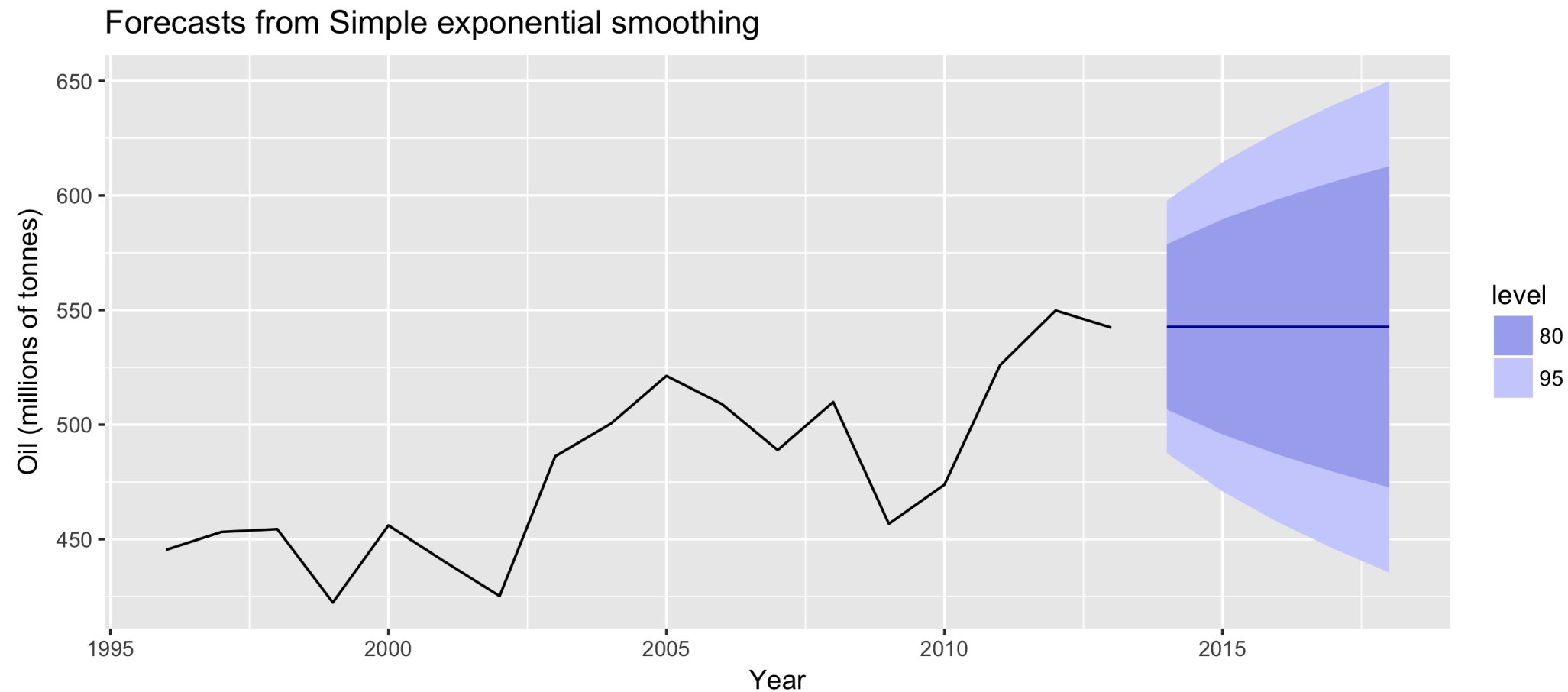
```
l = 446.5759
```

```
sigma: 28.12
```

```
*** Truncated due to space
```

Example: oil production

```
autoplot(fc) +  
  ylab("Oil (millions of tonnes)") + xlab("Year")
```



Let's practice!
FORECASTING IN R

Exponential smoothing methods with trend

FORECASTING IN R



Rob J. Hyndman

Professor of Statistics at Monash
University

Holt's linear trend

	Simple exponential smoothing
Forecast	$\hat{y}_{t+h t} = \ell_t$
Level	$\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1}$

	Holt's linear trend
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	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$

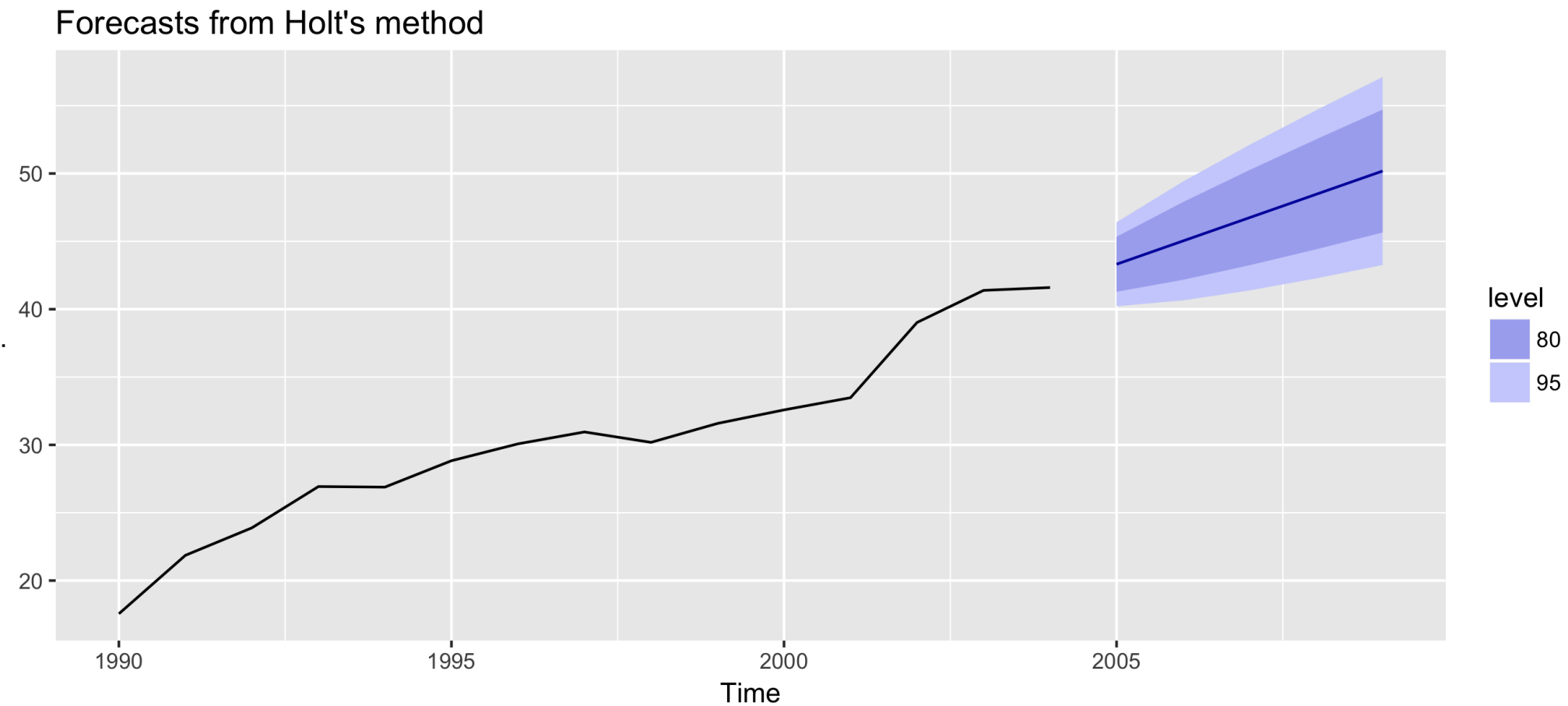
Holt's linear trend

	Holt's linear trend
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 α and β^* where $0 \leq \alpha$ and $\beta^* \leq 1$
- Choose $\alpha, \beta^*, \ell_0, b_0$ to minimize SSE

Holt's method in R

```
airpassengers %>% holt(h = 5) %>% autoplot
```



Damped trend method

Component form

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

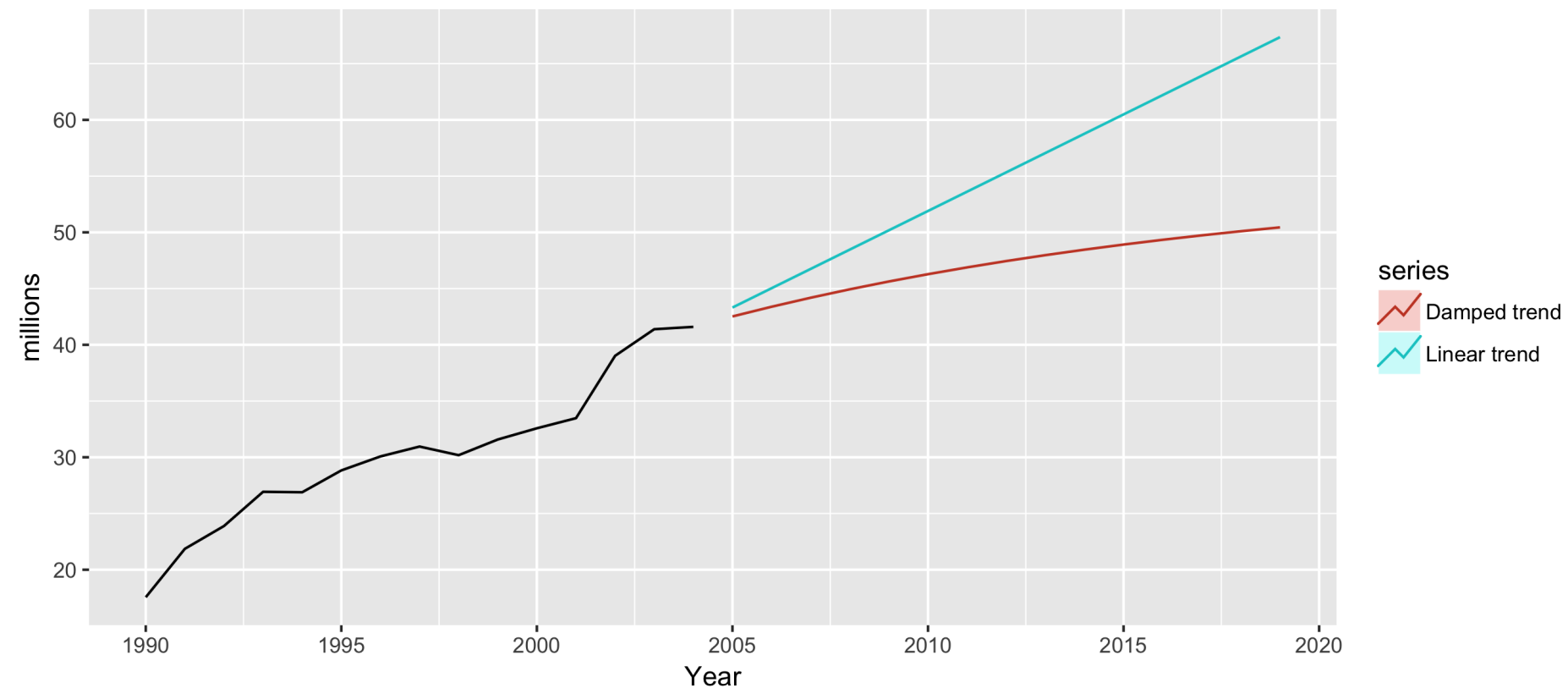
$$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$$

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

- Damping parameter $0 < \phi < 1$
- If $\phi = 1$, identical to Holt's linear trend
- Short-run forecasts trended, long-run forecasts constant

Example: air passengers

```
fc1 <- holt(airpassengers, h = 15, PI = FALSE)
fc2 <- holt(airpassengers, damped = TRUE, h = 15, PI = FALSE)
autoplot(airpassengers) + xlab("Year") + ylab("millions") +
  autolayer(fc1, series="Linear trend") +
  autolayer(fc2, series="Damped trend")
```



Let's practice!
FORECASTING IN R

Exponential smoothing methods with trend and seasonality

FORECASTING IN R

Rob J. Hyndman

Professor of Statistics at Monash
University



Holt-Winters' additive method

Holt-Winters additive method

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t-m+h_m^+}$$

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

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

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

- $s_{t-m+h_m^+}$ = seasonal component from final year of data
- Smoothing parameters:
 $0 \leq \alpha \leq 1, 0 \leq \beta^* \leq 1, 0 \leq \gamma \leq 1 - \alpha$
- m = period of seasonality (e.g. $m = 4$ for quarterly data)
- Seasonal component averages **zero**

Holt-Winters' multiplicative method

Holt-Winters multiplicative method

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

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

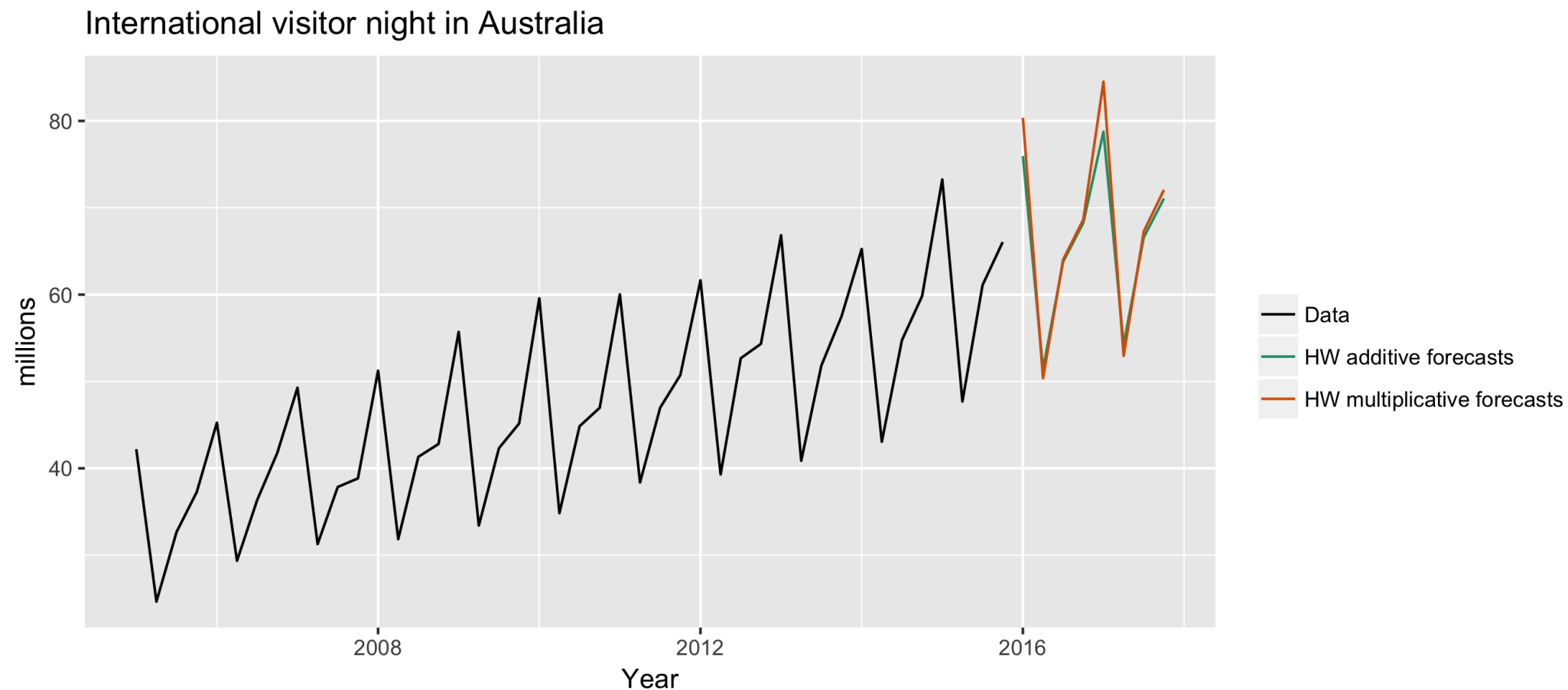
$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$$

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

- Seasonal component averages **one**

Example: Visitor Nights

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



Taxonomy of exponential smoothing methods

	Seasonal Component		
Trend Component	N (None)	A (Additive)	M (Multiplicative)
N (None)	(N, N)	(N, A)	(N, M)
A (Additive)	(A, N)	(A, A)	(A, M)
A _d (Additive Damped)	(A _d , N)	(A _d , N)	(A _d , N)

Taxonomy of exponential smoothing methods

	Seasonal Component		
Trend Component	N (None)	A (Additive)	M (Multiplicative)
N (None)	(N, N)	(N, A)	(N, M)
A (Additive)	(A, N)	(A, A)	(A, M)
A _d (Additive Damped)	(A _d , N)	(A _d , N)	(A _d , N)

(N, N)	Simple exponential smoothing	ses()
(A, N)	Holt's linear method	holt()
(A _d , N)	Additive damped trend method	hw()
(A, A)	Additive Holt-Winter's method	hw()
(A, M)	Multiplicative Holt-Winter's method	hw()
(A _d , M)	Damped multiplicative Holt-Winter's	hw()

Let's practice!
FORECASTING IN R

State space models for exponential smoothing

FORECASTING IN R



Rob J. Hyndman

Professor of Statistics at Monash
University

Innovations state space models

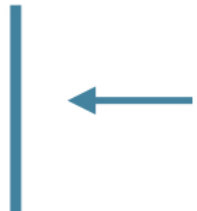
- Each exponential smoothing method can be written as an **"innovations state space model"**
 - Trend = {N, A, A_d}

Innovations state space models

- Each exponential smoothing method can be written as an **"innovations state space model"**
 - Trend = $\{N, A, A_d\}$
 - Seasonal = $\{N, A, M\}$


Innovations state space models

- Each exponential smoothing method can be written as an **"innovations state space model"**

- Trend = {N, A, A_d}
 - Seasonal = {N, A, M}
- 
- 3 x 3 = 9 possible
exponential smoothing
methods

Innovations state space models

- Each exponential smoothing method can be written as an **"innovations state space model"**

- Trend = {N, A, A_d}
 - Seasonal = {N, A, M}
 - Error = {A, M}
- 
- 3 x 3 = 9 possible
exponential smoothing
methods

Innovations state space models

- Each exponential smoothing method can be written as an "innovations state space model"
 - Trend = {N, A, A_d}
 - Seasonal = {N, A, M}
 - Error = {A, M}
- 3 x 3 = 9 possible exponential smoothing methods
- 9 x 2 = 18 possible state space models
- ETS models: Error, Trend, Seasonal

ETS models

- Parameters: estimated using the "**likelihood**", the probability of the data arising from the specified model
- For models with additive errors, this is **equivalent to minimizing SSE**
- Choose the best model by minimizing a corrected version of Akaike's Information Criterion (AIC_c)

Example: Australian air traffic

```
ets(ausair)
```

```
ETS(M,A,N)
```

```
Call:
```

```
ets(y = ausair)
```

```
Smoothing parameters:
```

```
alpha = 0.9999
```

```
beta = 0.0176
```

```
Initial states:
```

```
l = 6.5242
```

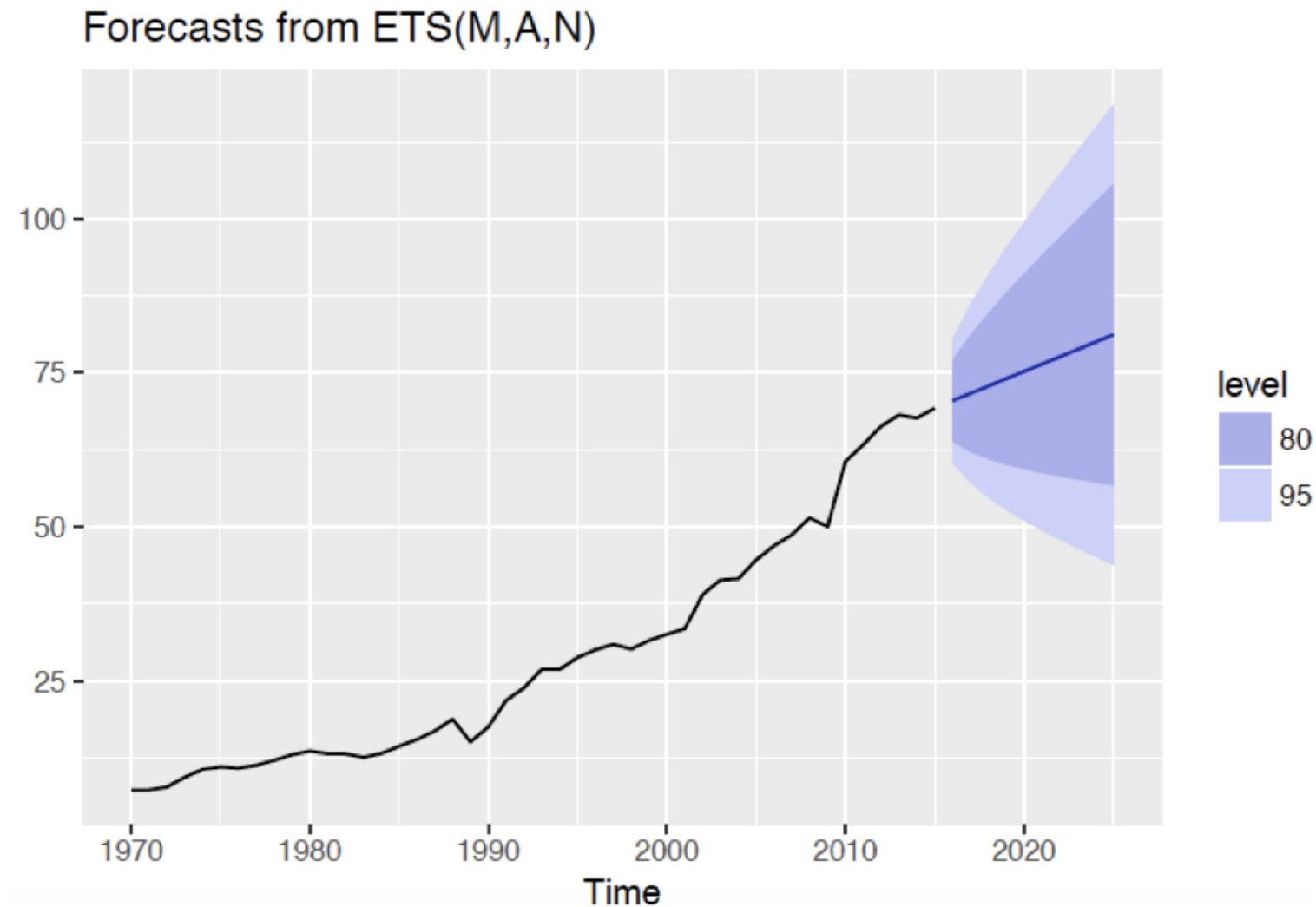
```
b = 0.7584
```

```
sigma: 0.0729
```

AIC	AICc	BIC
234.5273	236.0273	243.6705

Example: Australian air traffic

```
ausair %>% ets() %>% forecast() %>% autoplot()
```



Example: Monthly cortecosteroid drug sales

```
ets(h02)
```

```
ETS(M,Ad,M)
```

```
Call:
```

```
ets(y = h02)
```

```
Smoothing parameters:
```

```
alpha = 0.2173
```

```
beta  = 2e-04
```

```
gamma = 1e-04
```

```
phi   = 0.9756
```

```
Initial states:
```

```
l = 0.3996
```

```
b = 0.0098
```

```
s=0.8675 0.8259 0.7591 0.7748 0.6945 1.2838
```

```
1.3366 1.1753 1.1545 1.0968 1.0482 0.983
```

```
sigma: 0.0647
```

```
AIC
```

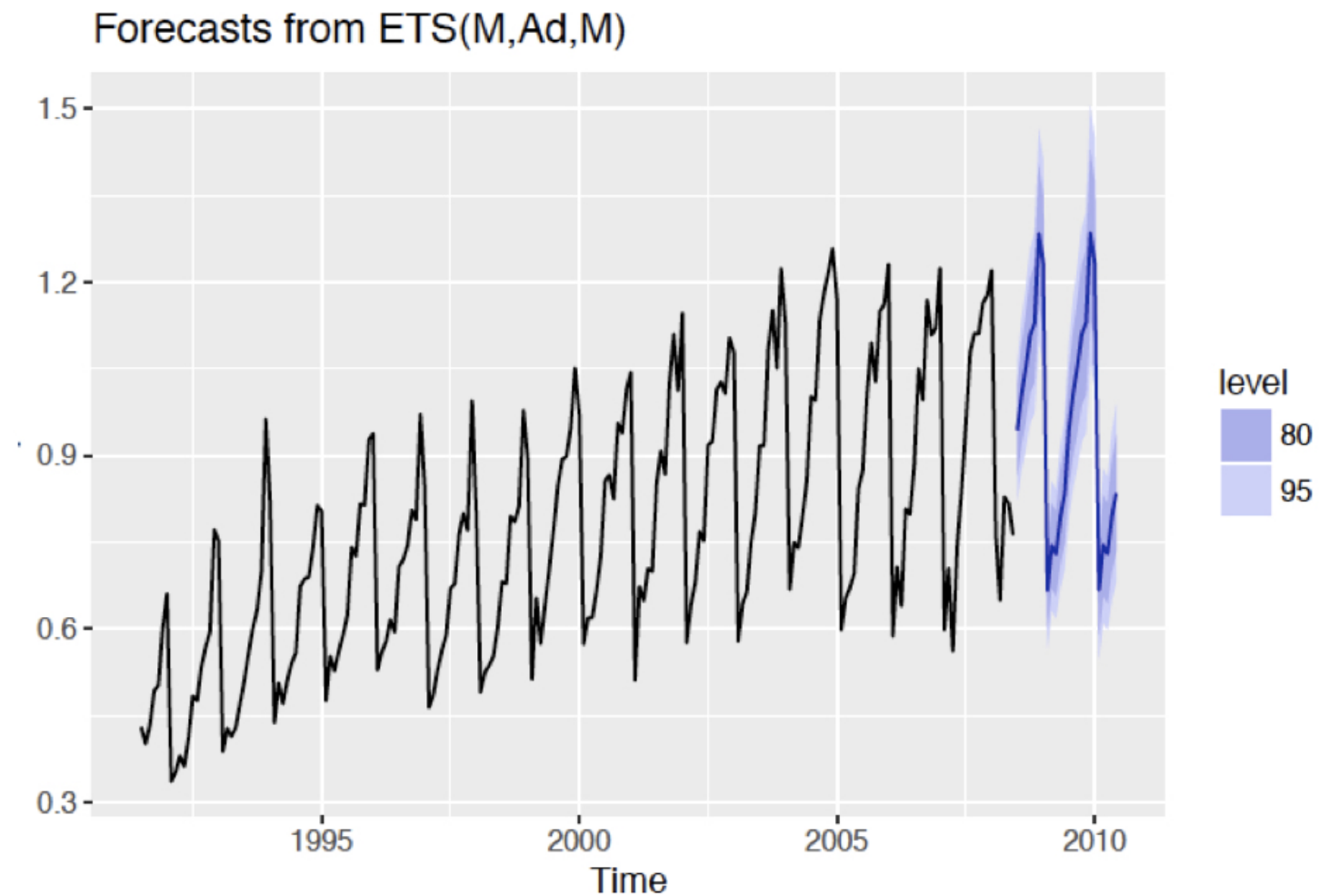
```
AICc
```

```
BIC
```

```
-123.21905 -119.52175 -63.49289
```

Example: Monthly cortecosteroid drug sales

```
h02 %>% ets() %>% forecast() %>% autoplot()
```



Let's practice!
FORECASTING IN R