



MONASH  
University

MONASH  
BUSINESS  
SCHOOL

# ETC3550/ETC5550 Applied forecasting

## Ch8. Exponential smoothing

[af.numbat.space](http://af.numbat.space)



# ETS models

**General notation**    E T S : ExponenTial Smoothing



Error   **T**rend   **S**eason

**Error:** Additive ("A") or multiplicative ("M")

# ETS models

**General notation**    E T S : ExponenTial Smoothing



The diagram shows three arrows pointing from the words 'Error', 'Trend', and 'Season' below to the letters 'E', 'T', and 'S' respectively in the 'ETS' part of the notation above.

Error Trend Season

**Error:** Additive ("A") or multiplicative ("M")

**Trend:** None ("N"), additive ("A"), multiplicative ("M"), or damped ("Ad" or "Md").

# ETS models

**General notation**    E T S : ExponenTial Smoothing



Error Trend Season

**Error:** Additive ("A") or multiplicative ("M")

**Trend:** None ("N"), additive ("A"), multiplicative ("M"), or damped ("Ad" or "Md").

**Seasonality:** None ("N"), additive ("A") or multiplicative ("M")

# ETS(A,N,N): SES with additive errors

## ETS(A,N,N) model

Observation equation

$$y_t = l_{t-1} + \varepsilon_t$$

State equation

$$l_t = l_{t-1} + \alpha \varepsilon_t$$

where  $\varepsilon_t \sim \text{NID}(0, \sigma^2)$ .

- “innovations” or “single source of error” because equations have the same error process,  $\varepsilon_t$ .
- Measurement equation: relationship between observations and states.
- State equation(s): evolution of the state(s) through time.

# ETS(A,A,N)

Holt's methods method with additive errors.

Forecast equation  
Observation equation  
State equations

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

$$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$$

$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$$

$$b_t = b_{t-1} + \beta\varepsilon_t$$

■ Forecast errors:  $\varepsilon_t = y_t - \hat{y}_{t|t-1}$

Holt-Winters additive method with additive errors.

Forecast equation	$\hat{y}_{t+h t} = \ell_t + hb_t + s_{t+h-m(k+1)}$
Observation equation	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$
State equations	$\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$
	$b_t = b_{t-1} + \beta\varepsilon_t$
	$s_t = s_{t-m} + \gamma\varepsilon_t$

- Forecast errors:  $\varepsilon_t = y_t - \hat{y}_{t|t-1}$
- $k$  is integer part of  $(h - 1)/m$ .

Holt-Winters multiplicative method with multiplicative errors.

Forecast equation	$\hat{y}_{t+h t} = (\ell_t + hb_t)s_{t+h-m(k+1)}$
Observation equation	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t)$
State equations	$\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$

- Forecast errors:  $\varepsilon_t = (y_t - \hat{y}_{t|t-1})/\hat{y}_{t|t-1}$
- $k$  is integer part of  $(h - 1)/m$ .



# ETS model specification

```
ETS(y ~ error("A") + trend("N") + season("N"))
```

By default, optimal values for  $\alpha$ ,  $\beta$ ,  $\gamma$ , and the states at time 0 are used.

The values for  $\alpha$ ,  $\beta$  and  $\gamma$  can be specified:

```
trend("A", alpha = 0.5, beta = 0.2)
trend("A", alpha_range = c(0.2, 0.8), beta_range = c(0.1, 0.4))
season("M", gamma = 0.04)
season("M", gamma_range = c(0, 0.3))
```

# Exponential smoothing methods

		Seasonal Component		
		N	A	M
Trend Component		(None)	(Additive)	(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

# Exponential smoothing methods

		Seasonal Component		
		N	A	M
Trend Component		(None)	(Additive)	(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'

There are also multiplicative trend methods (not recommended).

# ETS models

## Additive Error

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

## Multiplicative Error

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

# ETS models

## Additive Error

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

## Multiplicative Error

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

# AIC and cross-validation

Minimizing the AIC assuming Gaussian residuals is asymptotically equivalent to minimizing one-step time series cross validation MSE.

# Automatic forecasting

## From Hyndman et al. (IJF, 2002):

- Apply each model that is appropriate to the data. Optimize parameters and initial values using MLE (or some other criterion).
- Select best method using AICc:
- Produce forecasts using best method.
- Obtain forecast intervals using underlying state space model.

Method performed very well in M3 competition.

# Residuals

## Response residuals

$$\hat{e}_t = y_t - \hat{y}_{t|t-1}$$

## Innovation residuals

Additive error model:

$$\hat{\varepsilon}_t = y_t - \hat{y}_{t|t-1}$$

Multiplicative error model:

$$\hat{\varepsilon}_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}}$$