Complex exponential Smoothing

Ivan Svetunkov
Nikolaos Kourentzes
Robert Fildes

18 March 2015

LANCASTER UNIVERSITY Management School

Lancaster Centre for Forecasting
www.forecasting-centre.com

Introduction

- Exponential Smoothing methods performed very well in many competitions:
  - M-Competitions in 1982 and 2000,
  - Competition on telecommunication data in 1998 and 2008,
  - Tourism forecasting competition in 2011.

- In practice forecasters usually use:
  - SES for the level time series,
  - Holt's method for trend time series,
  - Holt-Winters method for a trend-seasonal data.

Introduction

- Holt's method is not performing consistently. Examples:
  - M-Competitions;
  - Taylor, 2008;
  - Gardner & Diaz-Saiz, 2008;
  - Acar & Gardner, 2012.

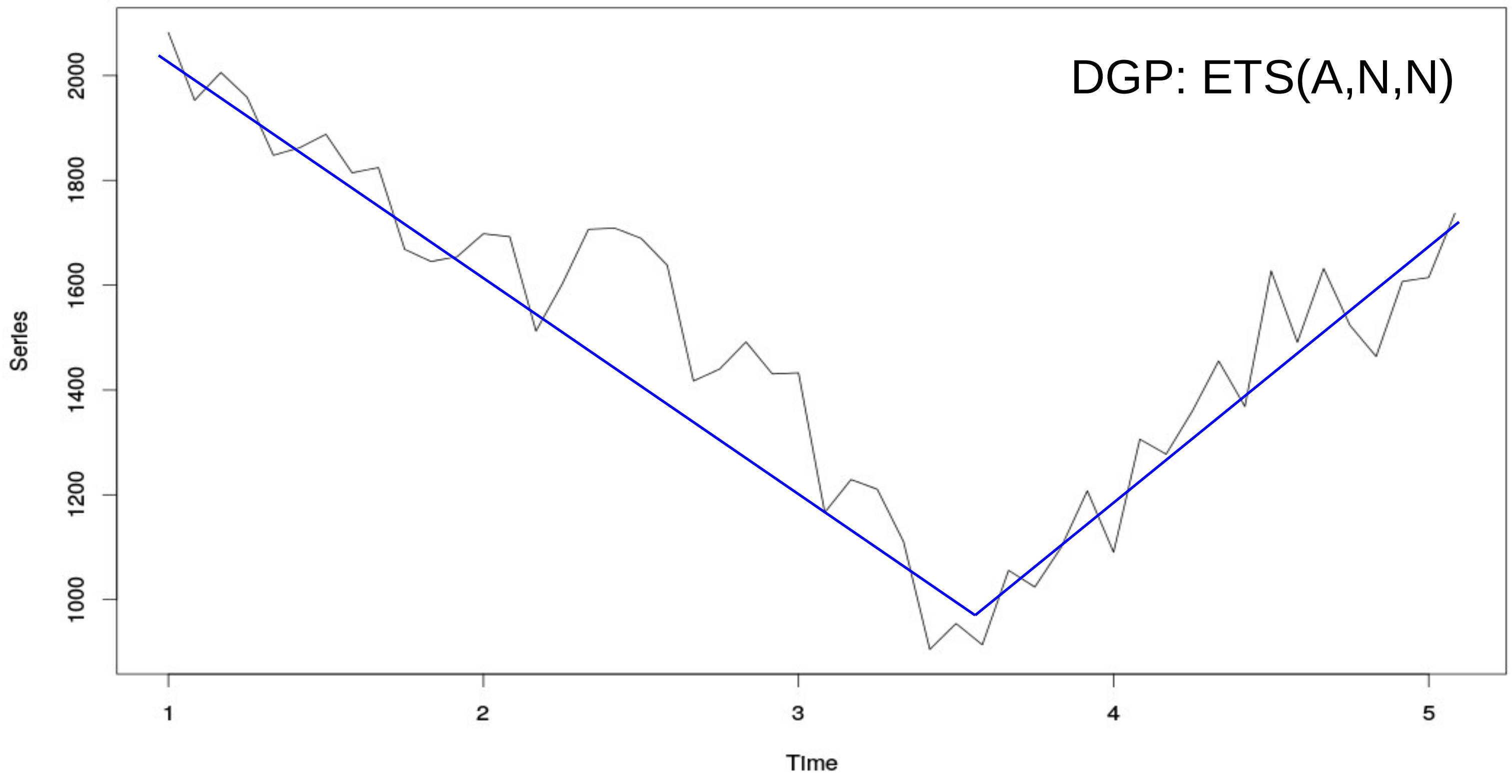
- Holt's method is still very popular in publications:
  - Gelper et. al, 2010;
  - Maia & de Carvalho, 2011.

Introduction

- Several modifications for different types of trends were proposed over the years:
  - Multiplicative trend (Pegels, 1969);
  - Damped trend (Gardner & McKenzie, 1985);
  - Damped multiplicative trend (Taylor, 2003);
  - Prior data transformation using cross-validation (Bermudez et. al., 2009).
- Model selection procedure based on IC is usually used.

Introduction

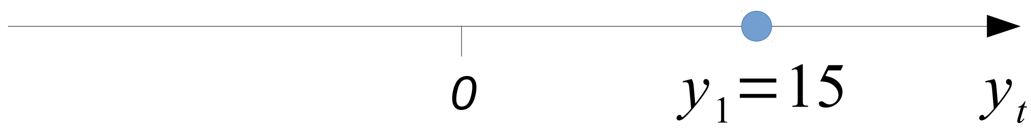
- But the trend is unobservable and arbitrary!



Lancaster Centre for Forecasting
www.forecasting-centre.com

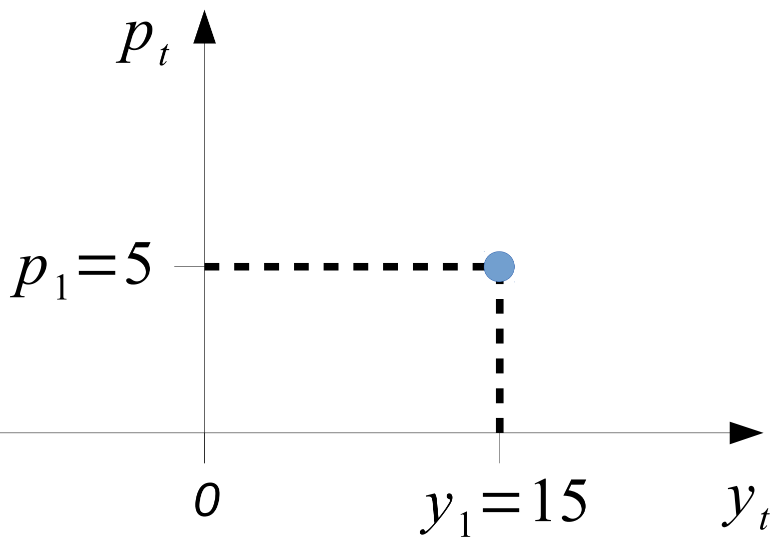
Reminder

- y_t is the real number, actual value,



- $y_t + ip_t$ is the complex number

$i^2 = -1$



Remark

- *The fact that imaginary numbers has hitherto been surrounded by mysterious obscurity, is to be attributed largely to an ill adapted notation.*
- *If "+1", "-1", and "√-1" had been called "direct", "inverse" and "lateral" units, instead of "positive", "negative" and "imaginary", such an obscurity would have been out of the question.*

Carl Friedrich Gauss

Lancaster Centre for Forecasting
www.forecasting-centre.com

New approach

- We propose a different approach to time series modelling.

$$y_t + i\,p_t$$

- where p_t is information potential

- Instead of:

$$y_t = f\left(y_{t-1}, y_{t-2}, \ldots, x_1, x_{2,} \ldots\right) + \varepsilon_t$$

- forecasting model now has a form:

$$y_t + i\,p_t = f\left(y_{t-1} + i\,p_{t-1}, y_{t-2} + i\,p_{t-2}, \ldots, x_1, x_{2,} \ldots\right) + \varepsilon_t + i\,\xi_t$$

Lancaster Centre for Forecasting
www.forecasting-centre.com

Theoretical framework

- Simple exponential smoothing:

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t$$

- Principle of CES: smooth level and combine it with information potential estimate.
- Basic form of CES:

$$\hat{y}_t + i\,\hat{p}_t = (\alpha_0 + i\,\alpha_1)(y_{t-1} + i\,\varsigma_{t-1}) + (1 - \alpha_0 + i - i\,\alpha_1)(\hat{y}_{t-1} + i\,\hat{p}_{t-1})$$

$$\varsigma_t = y_t - \hat{y}_t \qquad \varsigma_t = y_{t-s} - \hat{y}_{t-s} \qquad \varsigma_t = \Delta y_t \qquad \varsigma_t = f(x_{1,t}, x_{2,t}, \dots)$$

Lancaster Centre for Forecasting
www.forecasting-centre.com

Theoretical framework

$$\hat{y}_t + i\hat{p}_t = (\alpha_0 + i\alpha_1)(y_{t-1} + i\varsigma_{t-1}) + (1 - \alpha_0 + i - i\alpha_1)(\hat{y}_{t-1} + i\hat{p}_{t-1})$$

- Complex variables -> system of real variables:

$$\begin{cases} \hat{y}_t = (\alpha_0 y_{t-1} + (1-\alpha_0)\hat{y}_{t-1}) - (\alpha_1 \varsigma_{t-1} + (1-\alpha_1)\hat{p}_{t-1}) \\ \hat{p}_t = (\alpha_0 \varsigma_{t-1} + (1-\alpha_0)\hat{p}_{t-1}) + (\alpha_1 y_{t-1} + (1-\alpha_1)\hat{y}_{t-1}) \end{cases}$$

- Final forecast of CES consists of two parts:
  - smoothed level,
  - information potential part.

State-space form

- Any exponential smoothing method has an underlying statistical model.

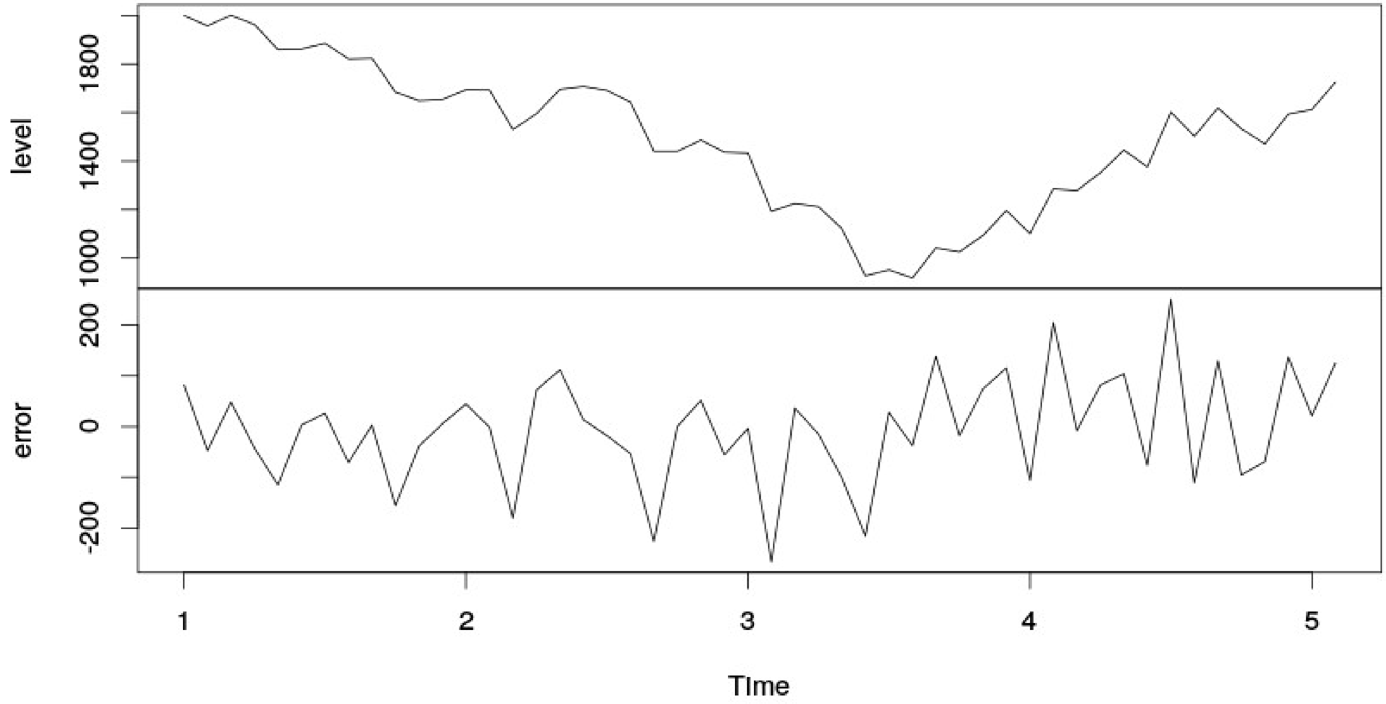
- State-space model with Single Source of Error.

- Every time series consists of components:
    - level,
    - trend,
    - seasonality,
    - error.

Lancaster Centre for Forecasting
www.forecasting-centre.com

State-space form

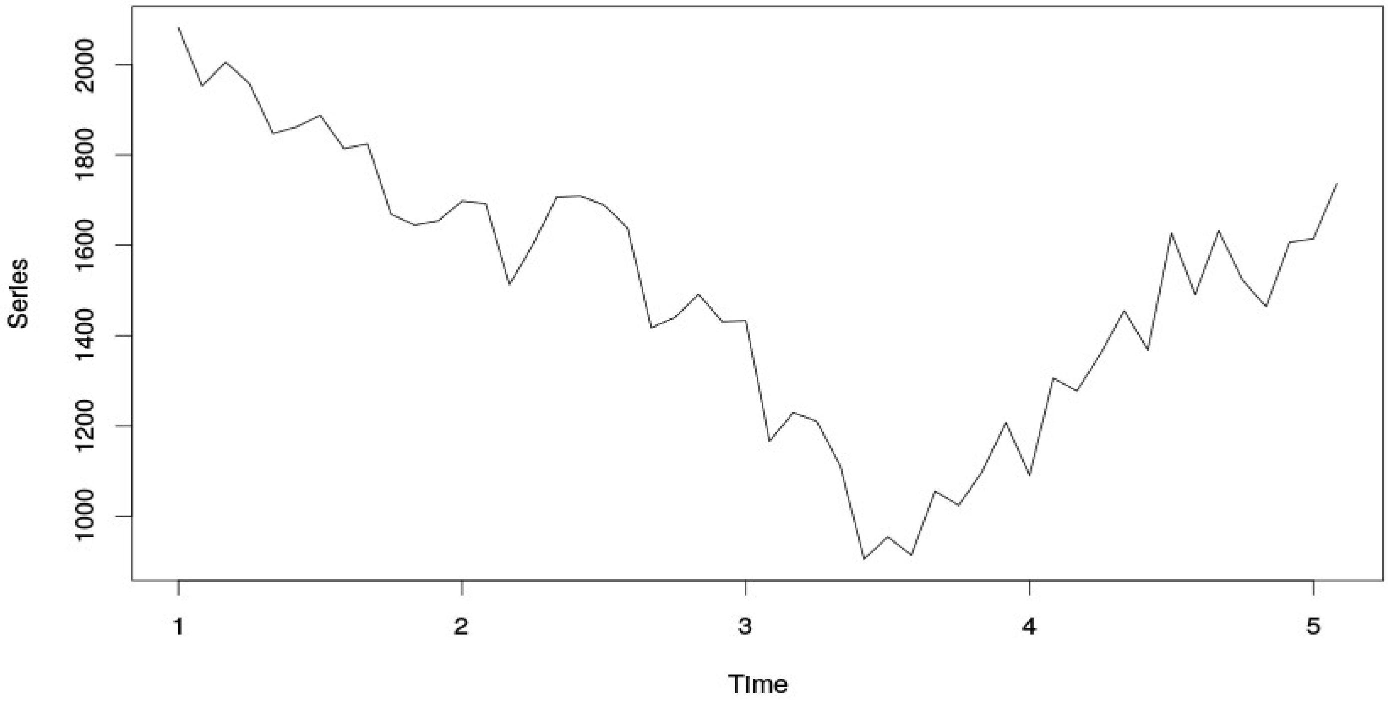
- Any time series model consists of:
  - transition equation: $x_t = F x_{t-1} + g \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2)$



Lancaster Centre for Forecasting
www.forecasting-centre.com

State-space form

- Any time series model consists of:
  - measurement equation: $y_t = w' x_{t-1} + \varepsilon_t$



State-space form

- State-space model with Single Source of Error:
  - measurement equation: $y_t = w' x_{t-1} + \varepsilon_t$
  - transition equation: $x_t = F x_{t-1} + g \varepsilon_t$

State-space form

- State-space form of CES:
  - measurement equation:

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

  - transition equation:

$$\begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 - (1 - \alpha_1) \\ 1 \quad 1 - \alpha_0 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} -\alpha_1 \\ \alpha_0 \end{pmatrix} \varsigma_t + \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \varepsilon_t$$

State-space form

- State-space form:

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

$$\begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 & -(1-\alpha_1) \\ 1 & 1-\alpha_0 \end{pmatrix} \begin{pmatrix} l_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} -\alpha_1 \\ \alpha_0 \end{pmatrix} \varsigma_t + \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \varepsilon_t$$

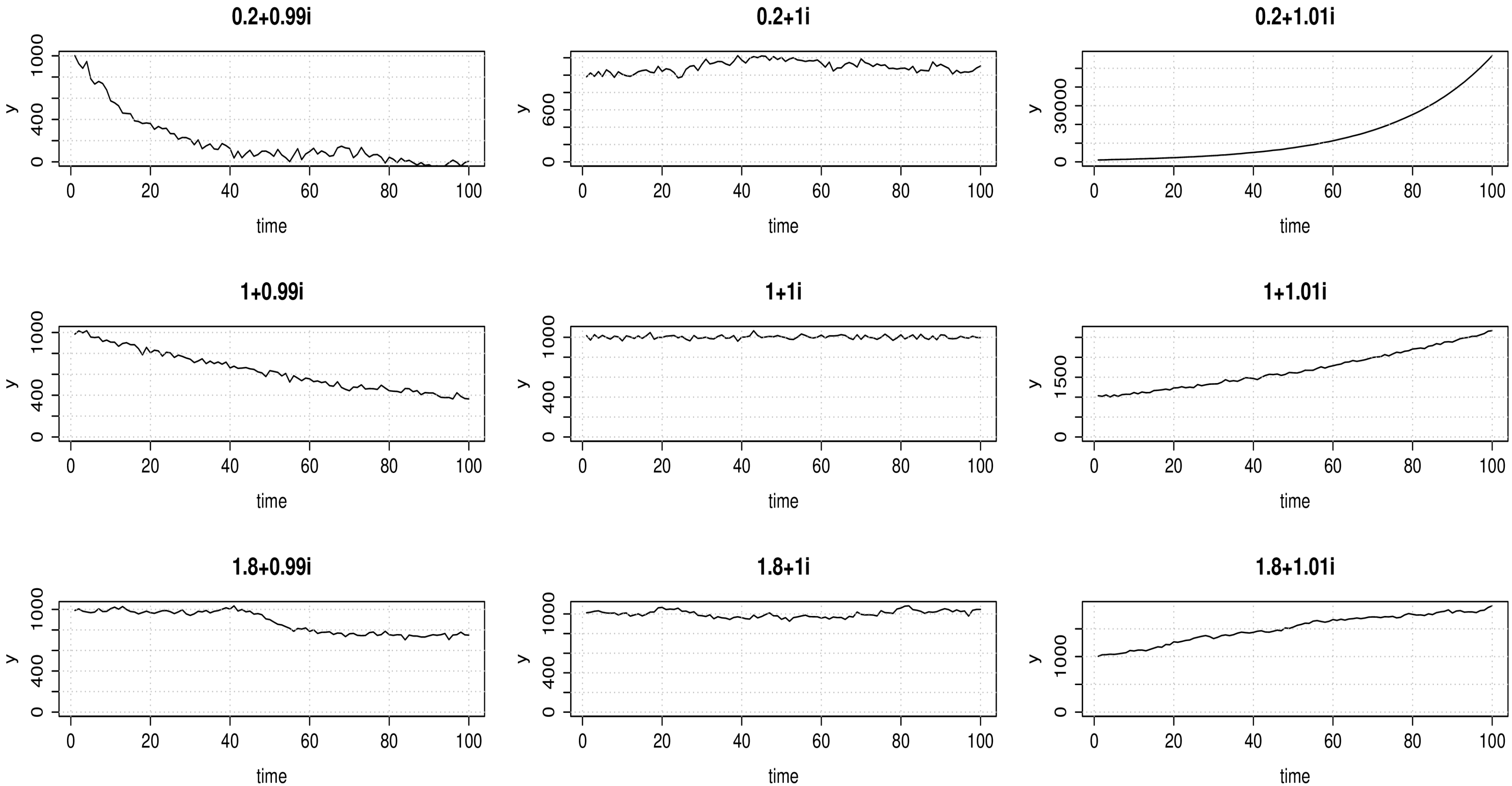
- Likelihood function:

$$f(\alpha_0 + i\alpha_{1,} \sigma^2 | y) = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^T \exp\left(-\frac{1}{2} \sum_{t=1}^{T} \left(\frac{\varepsilon_t}{\sigma} \right)^2 \right)$$

- Maximizing it is equivalent to minimization of SSE:

$$SSE = \sum_{t=1}^{T} \varepsilon_t^2$$

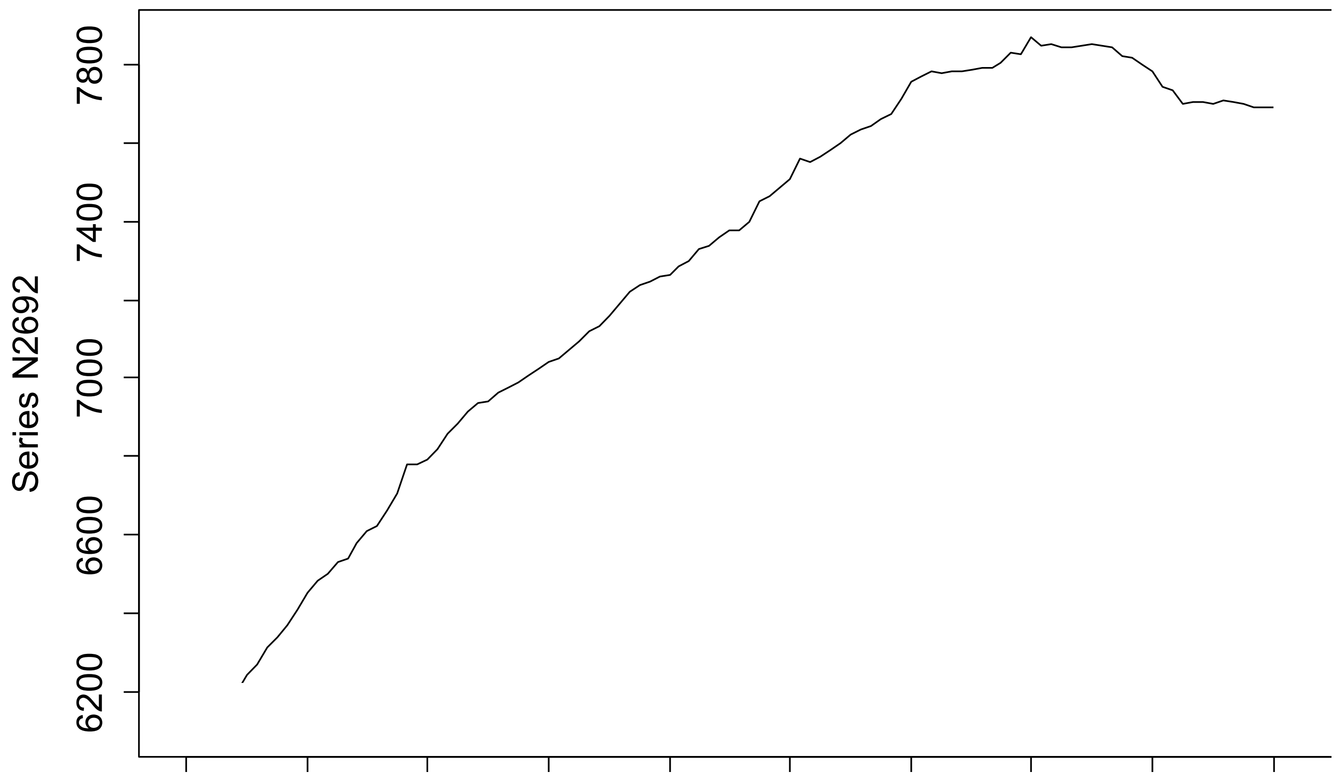
Time series simulation



Lancaster Centre for Forecasting
www.forecasting-centre.com

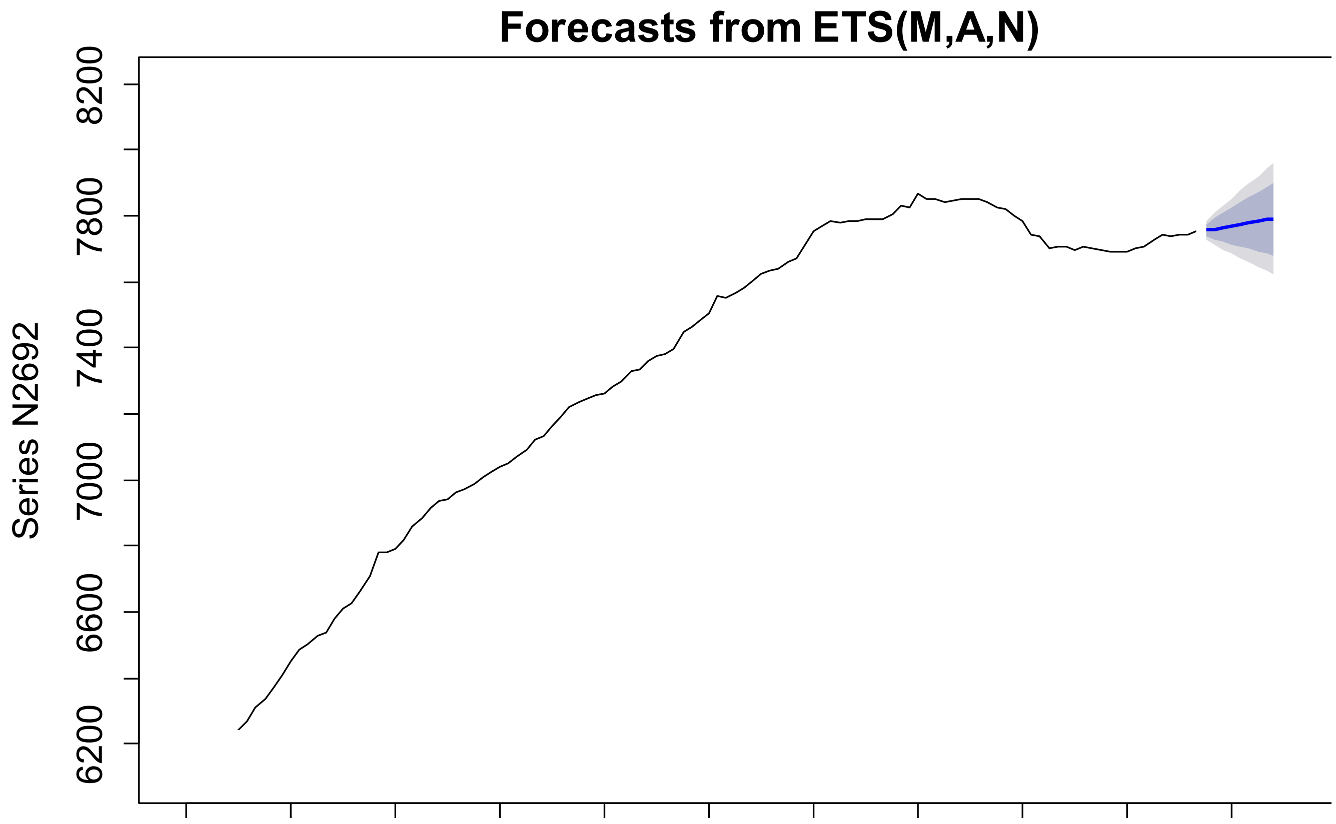
Example. Trended series

- Series N2692 from M3



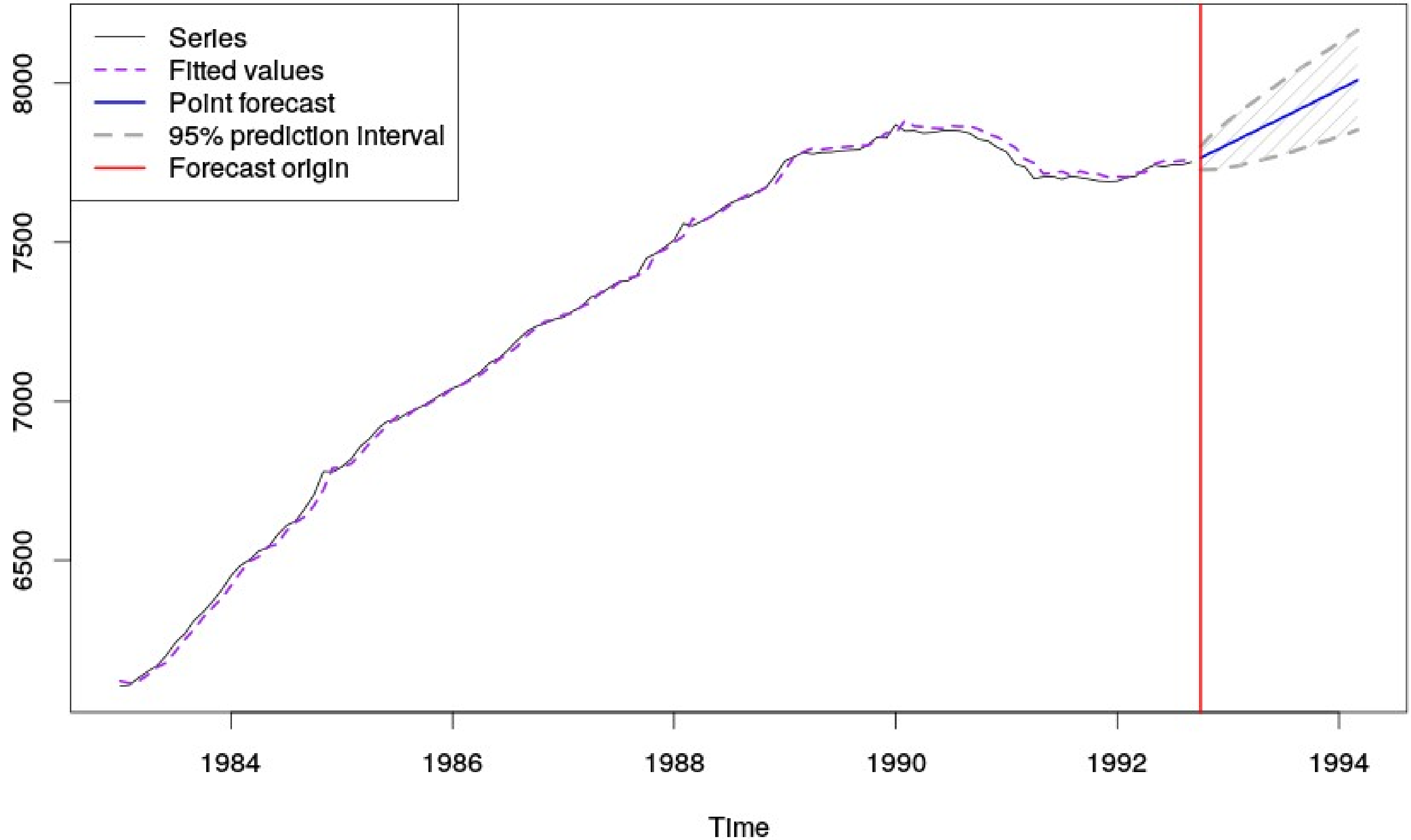
Example. Trended series

- ETS(M,A,N)



Example. Trended series

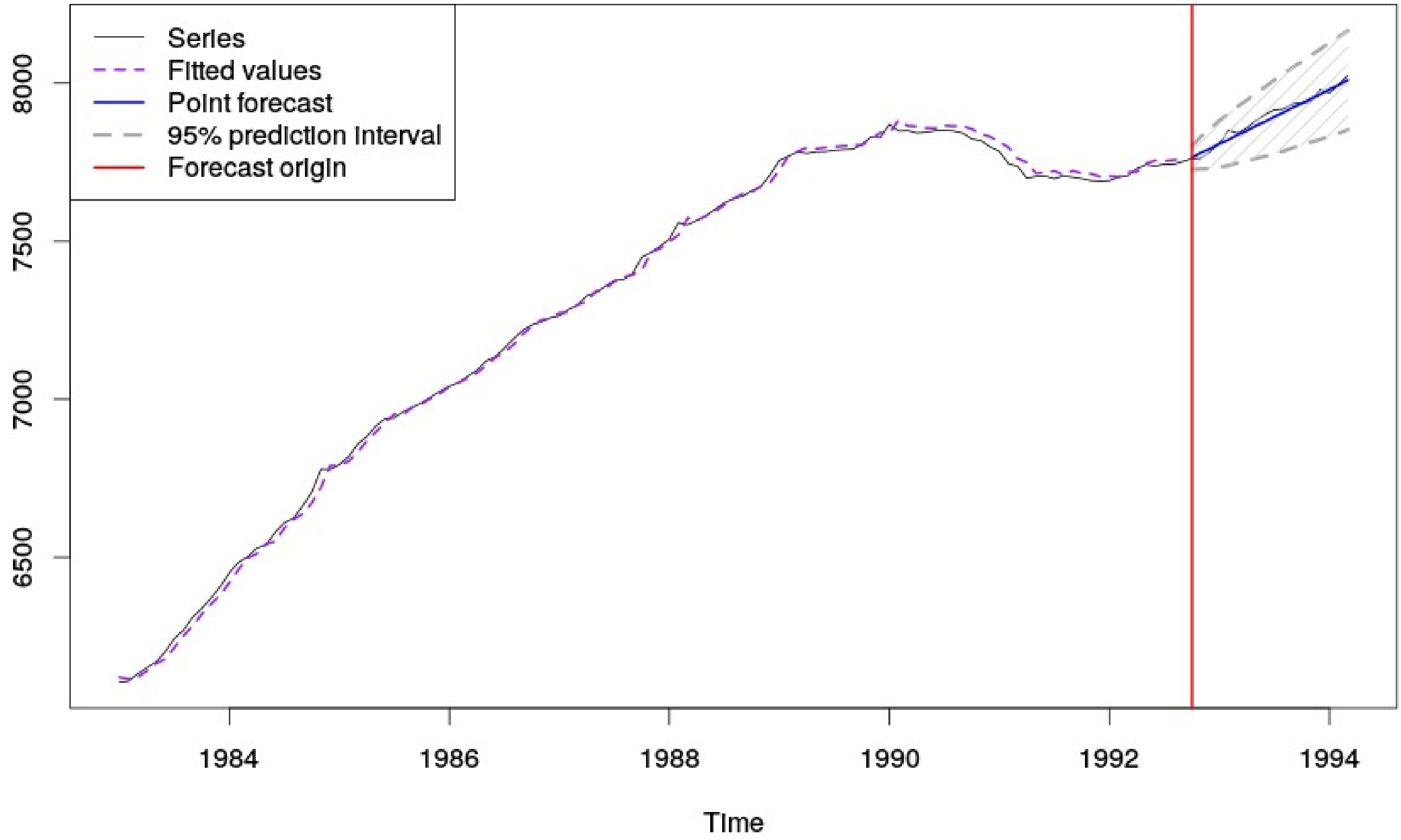
- CES



$$\alpha_0 + i\alpha_1 = 2.00056 + 1.00364i$$

Example. Trended series

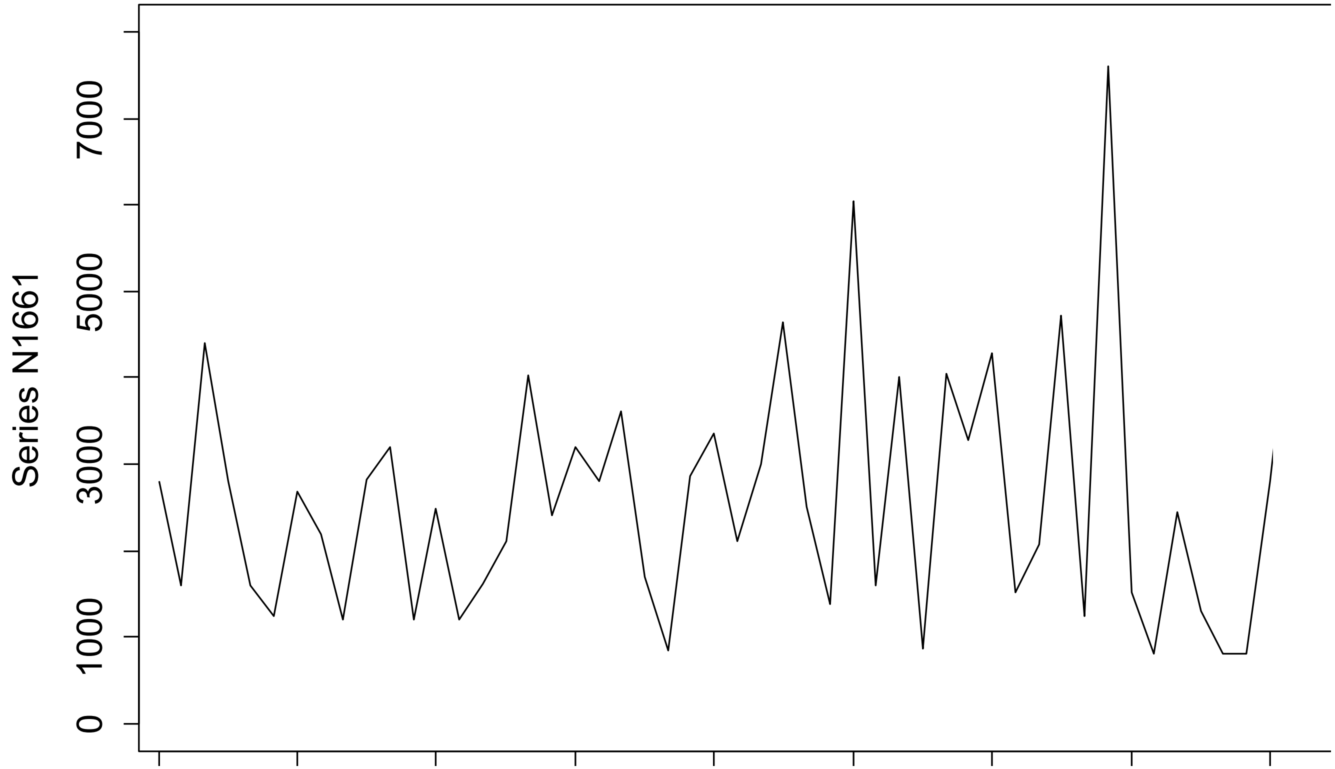
- CES



Lancaster Centre for Forecasting
www.forecasting-centre.com

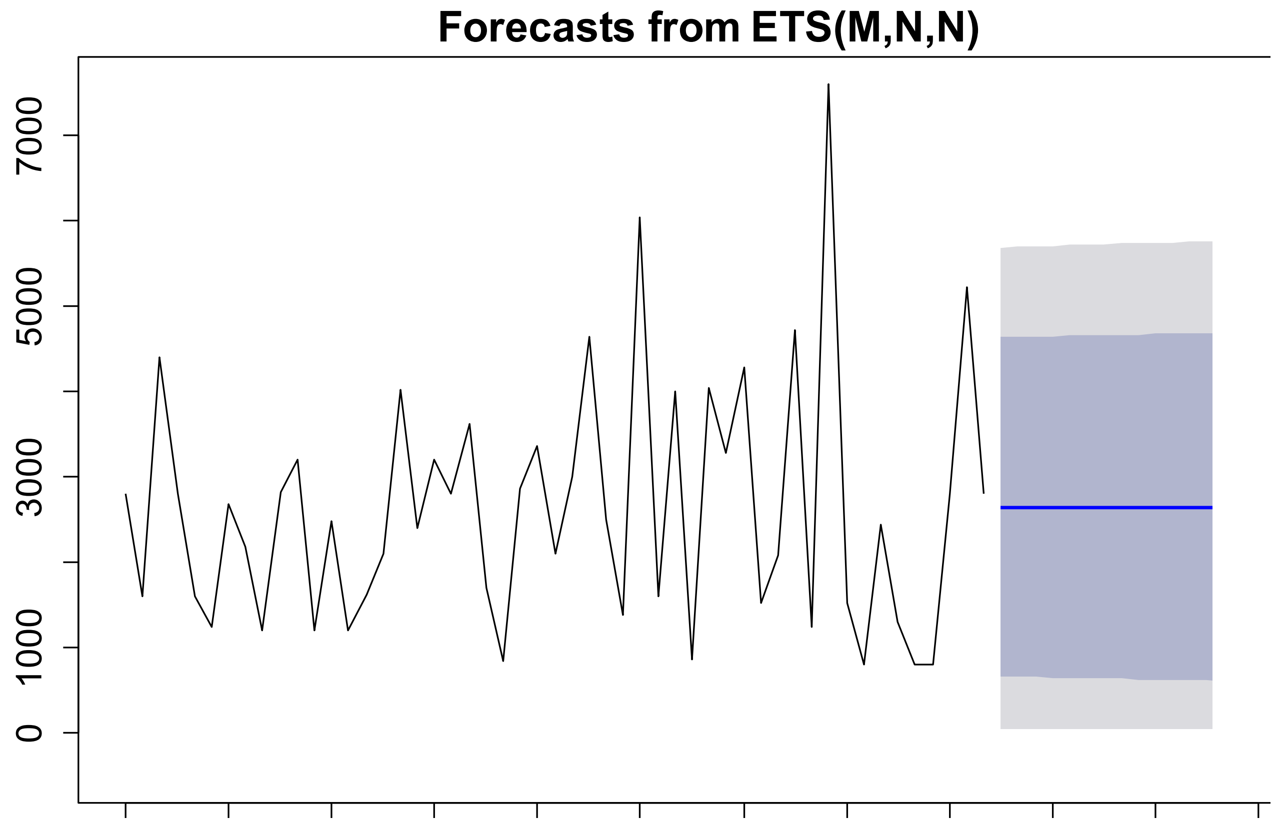
Example. Stationary series

- Series N1661 in M3



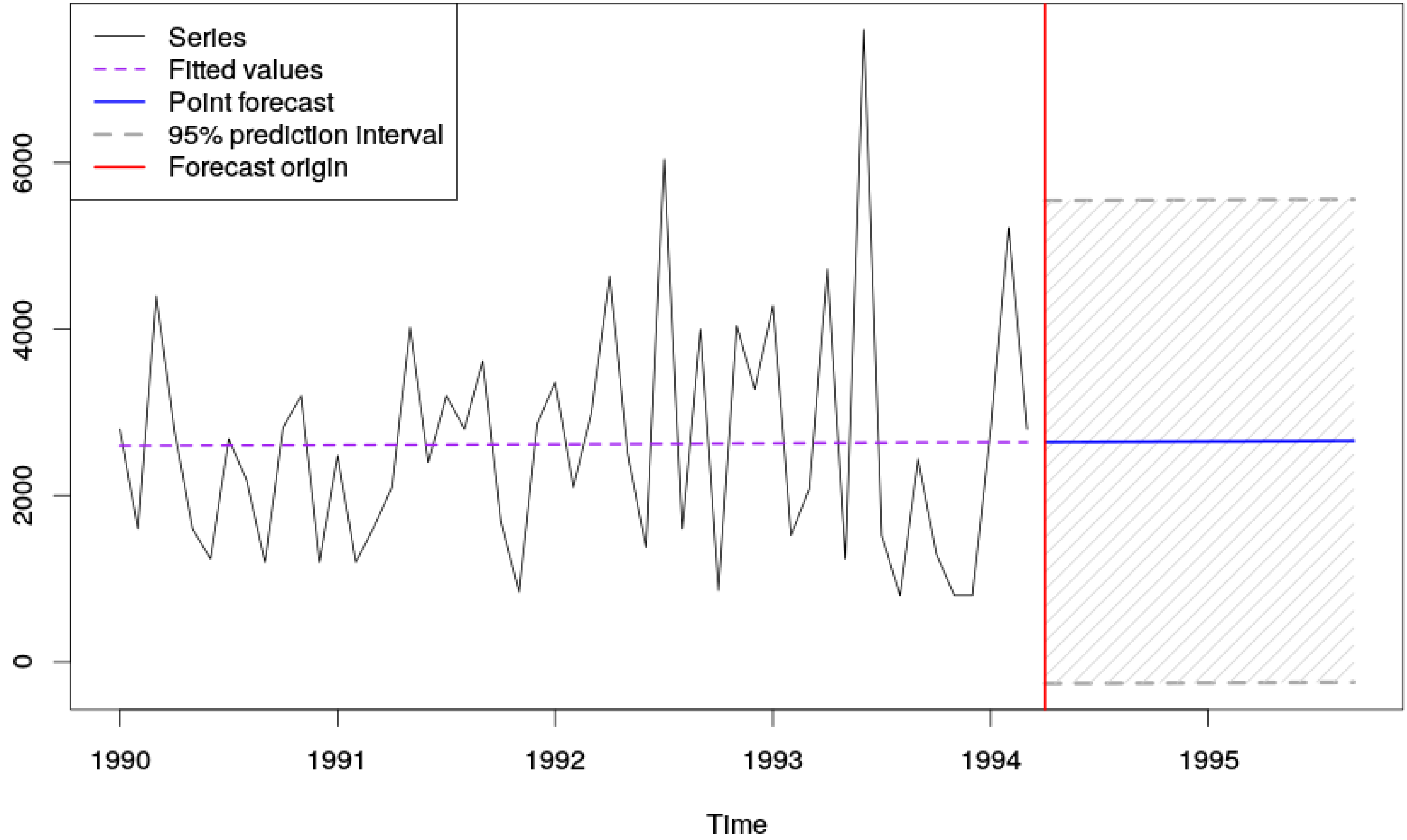
Example. Stationary series

- ETS(M,N,N)



Example. Stationary series

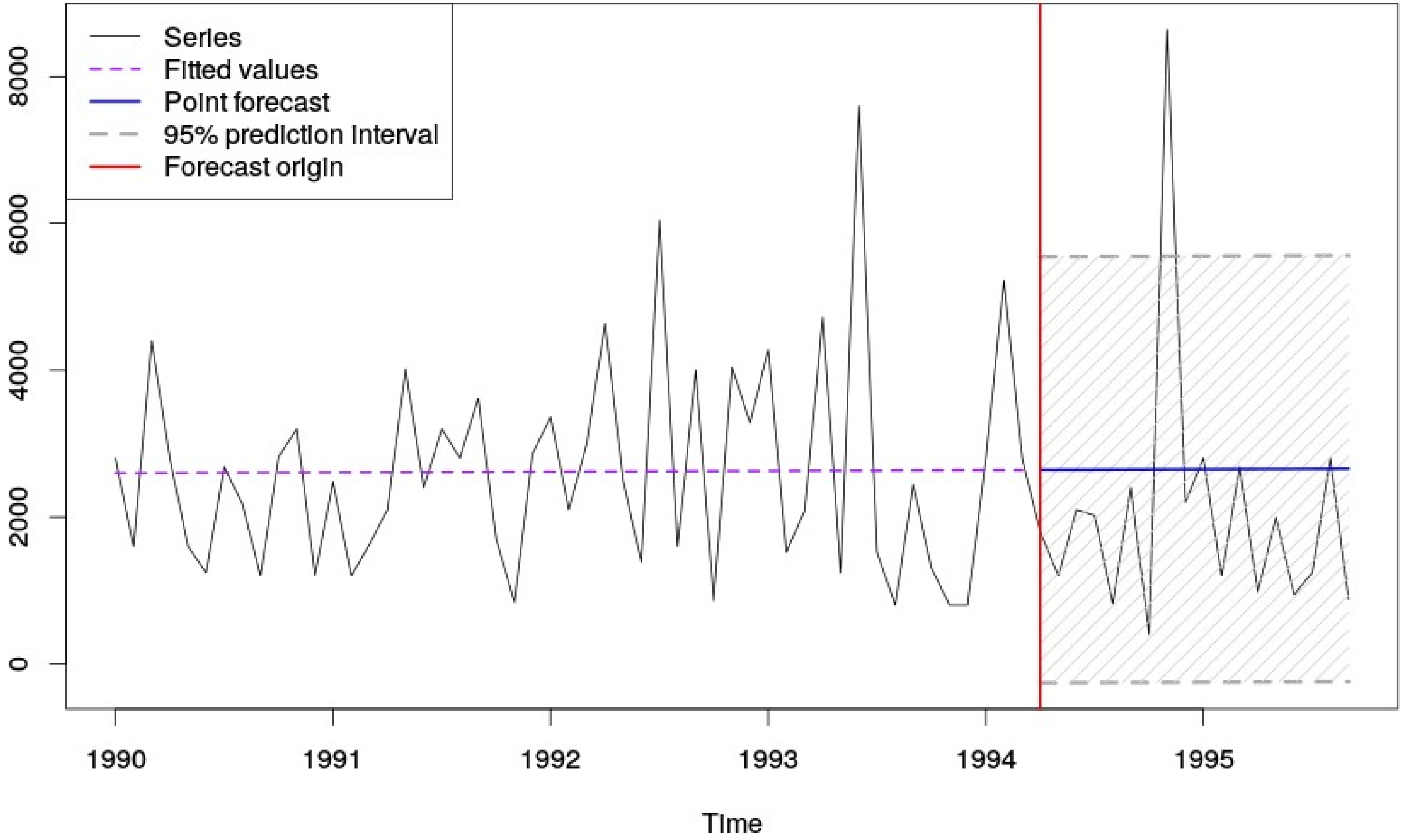
- CES



$$\alpha_0 + i\alpha_1 = 0.99999 + 1.00034\, i$$

Example. Stationary series

- CES



Empirical results: setup

- M3-Competition data. 3003 time series.
- Rolling origin.
- Automated ETS was used to split data into categories:
  - level non-seasonal,
  - level seasonal,
  - trend non-seasonal,
  - trend seasonal.

Empirical results: setup

- M3-Competition data. 3003 time series.
- Rolling origin.
- Automated ETS was used to split data into categories.

Series type	Number of series		Overall	Forecasting horizon	Rolling origin horizon
	Level series	Trend series			
year	255	390	645	6	12
quart	306	450	756	8	16
month	686	742	1428	18	24
other	61	113	174	8	16
Overall	1308	1695	3003		

Lancaster Centre for Forecasting
www.forecasting-centre.com

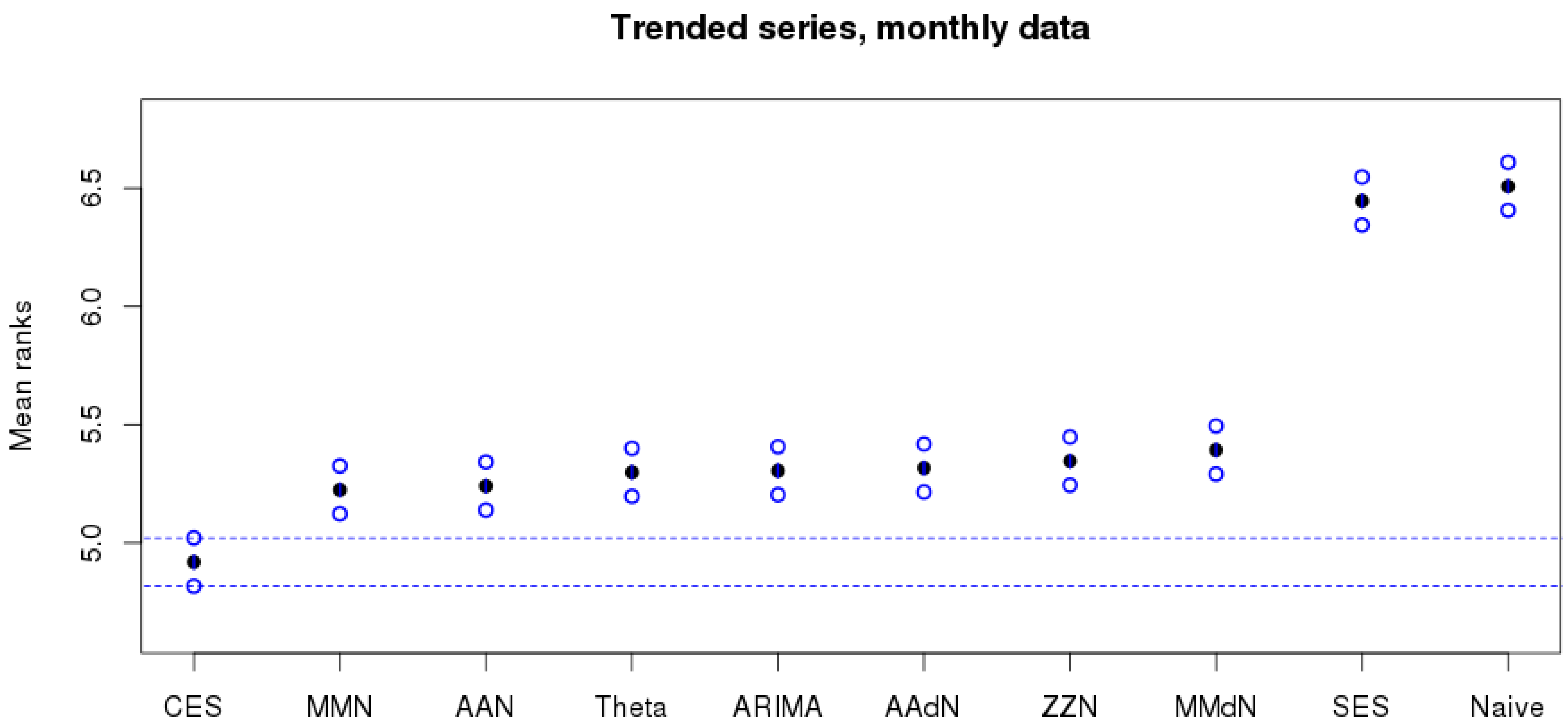
Empirical results: competitors

1. Naive (Naive),
2. Simple exponential smoothing (SES),
3. Holt’s additive trend (AAN),
4. Pegels’ multiplicative trend (MMN),
5. State-space ETS with AICc model selection (ZZN),
6. Gardner’s Damped trend (AAdN),
7. Taylor’s Damped multiplicative trend (MMdN),
8. Theta using Hyndman & Billah, 2003 (Theta),
9. Hyndman & Khandakar 2008 Auto ARIMA (ARIMA),
10. **Complex exponential smoothing** (CES).

Empirical results

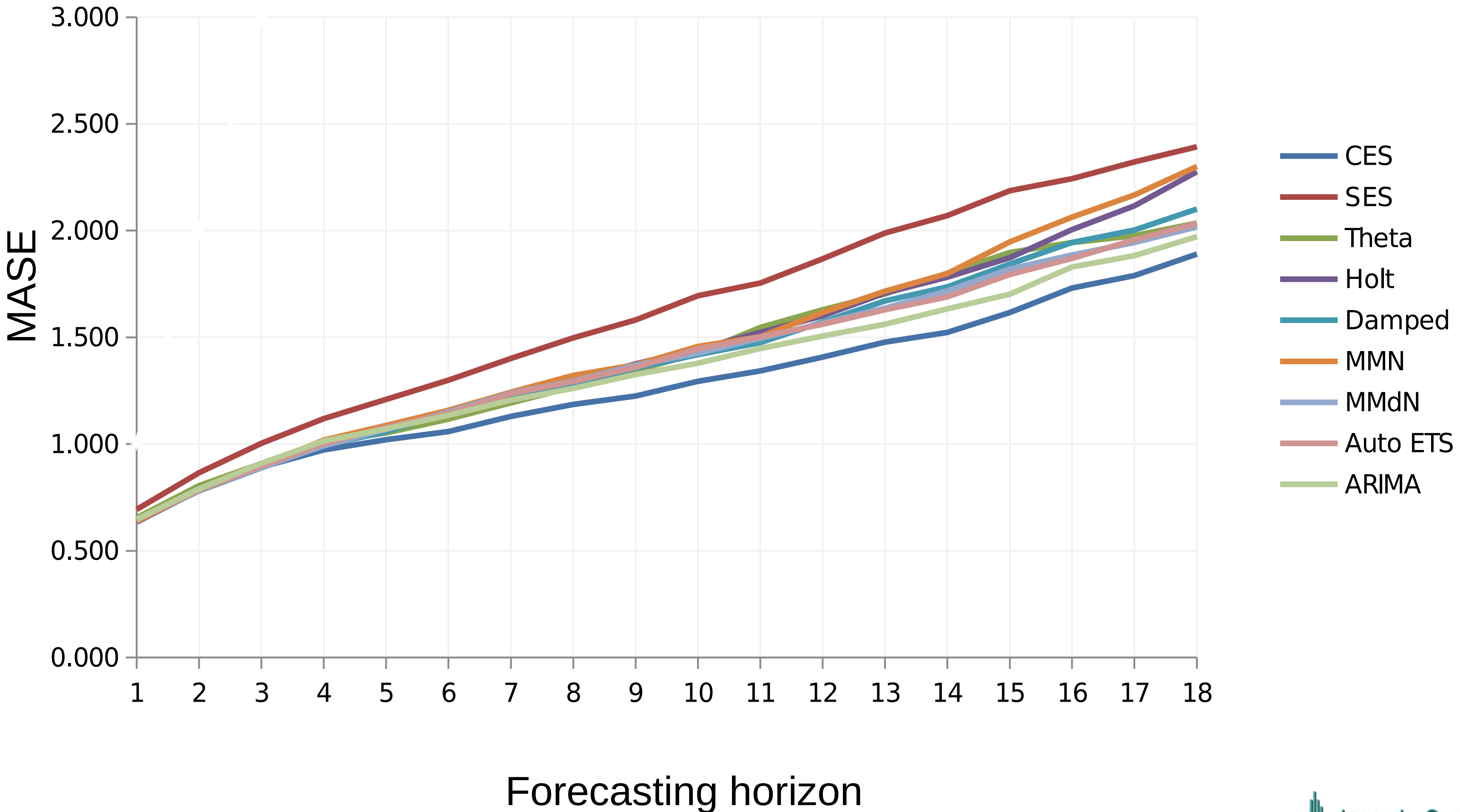
- MASE was calculated for each of the horizons from each of the origins,
- Nemenyi test was conducted to compare methods for each of the series type.
- General results for CES:
  - at least as good as SES on level series,
  - outperforms MMN and AAN on level series,
  - at least as good as MMN and AAN on trend series,
  - outperforms all the methods on monthly trend series.

Empirical results. Nemenyi test



Lancaster Centre for Forecasting
www.forecasting-centre.com

Empirical results



Lancaster Centre for Forecasting
www.forecasting-centre.com

Conclusions

- CES
  - is flexible,
  - has an underlying statistical model,
  - is able to identify trends and levels,
  - does it better than Holt and Pegels,
  - is at least as good as SES,
  - outperforms all the other methods on monthly data,
  - is more accurate on long-term horizons.

Thank you!

Ivan Svetunkov,
email: i.svetunkov@lancaster.ac.uk

Lancaster Centre for
Forecasting
www.forecasting-centre.com