Complex Exponential Smoothing for Time Series Forecasting

Ivan Svetunkov
Nikolaos Kourentzes

24 June 2015, ISF 2015

LANCASTER UNIVERSITY Management School

Lancaster Centre for Forecasting
www.forecasting-centre.com

Introduction

- Exponential Smoothing methods are very popular in forecasting

- They 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.

Introduction

- Hyndman et al., 2008 proposed a taxonomy that includes:
  - 2 types of error terms (additive and multiplicative);
  - 5 types of trend components (none, additive, multiplicative, damped additive and damped multiplicative);
  - 3 types of seasonality (none, additive, multiplicative).
- In theory it leads to 30 types of ETS models
- Model selection procedure based on IC is widely used

Lancaster Centre for Forecasting
www.forecasting-centre.com

Introduction

- Kolassa, 2011 demonstrated that combination of ETS models based on AIC outperforms single ETS
- Kourentzes et al., 2014 proposed MAPA, combining different ETS models across temporal aggregation levels
- The underlying model may be more complex than single ETS
- Model selection procedure may not work properly

Lancaster Centre for Forecasting
www.forecasting-centre.com

Introduction

- We proposed Complex Exponential Smoothing (Svetunkov, Kourentzes, 2015)
- Now we propose a modification of CES for seasonal data

CES method

- CES is based on idea that any time series consists of:

$$y_t + i\,p_t$$

- Using complex variables theory we derived the original CES method:

$$\hat{y}_{t+1} + i\,\hat{p}_{t+1} = \left(\alpha_0 + i\,\alpha_1\right)\left(y_t + i\,p_t\right) + \left(1 - \alpha_0 + i - i\,\alpha_1\right)\left(\hat{y}_t + i\,\hat{p}_t\right)$$

Lancaster Centre for Forecasting
www.forecasting-centre.com

CES model

- Information potential needs to be approximated:

$$p_t = \varepsilon_t$$

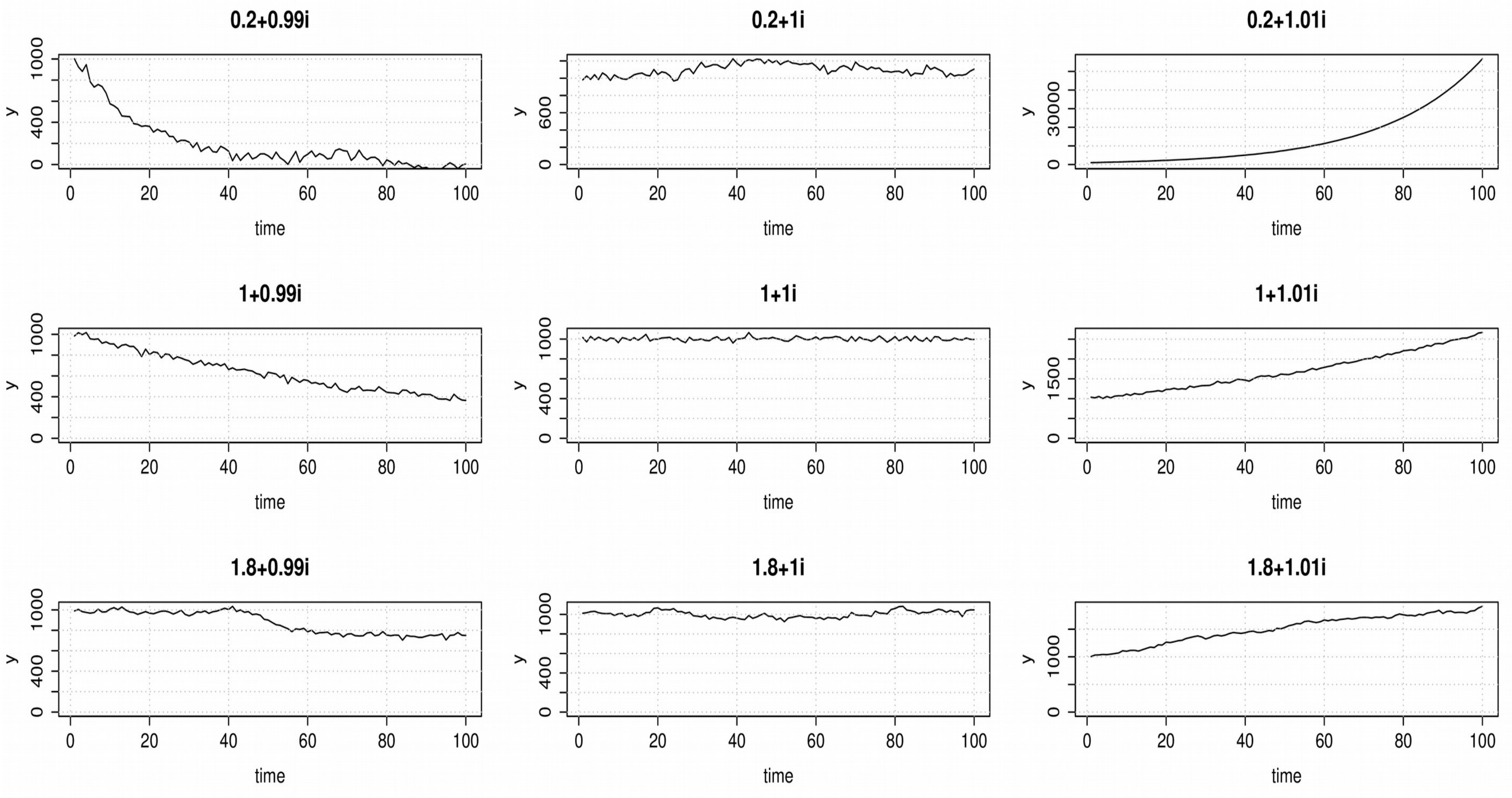
- Splitting CES into components allows to derive the following state-space model:

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

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

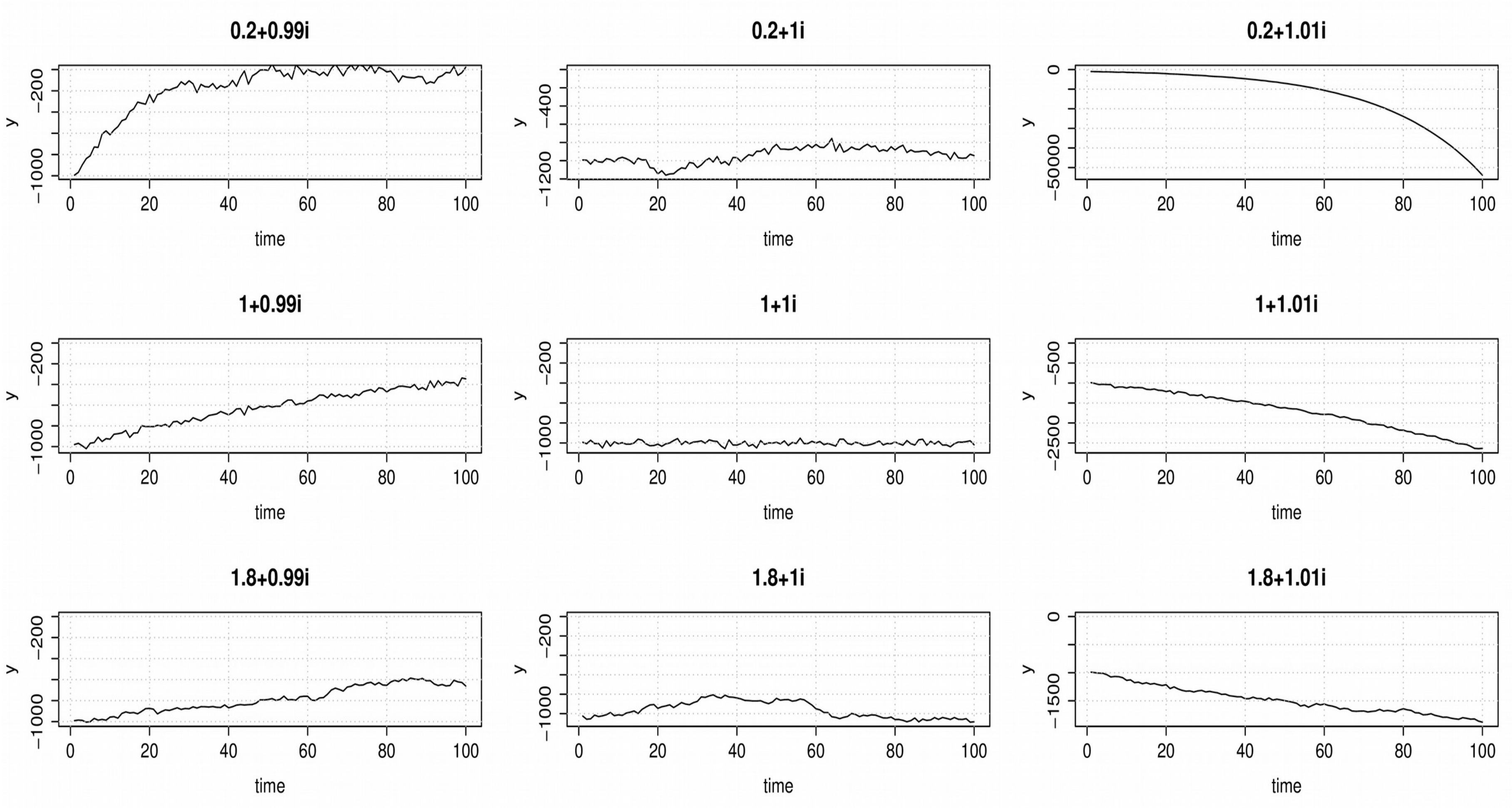
Lancaster Centre for Forecasting
www.forecasting-centre.com

CES trajectories



Lancaster Centre for Forecasting
www.forecasting-centre.com

CES trajectories



Lancaster Centre for Forecasting
www.forecasting-centre.com

Simple seasonal CES model

- Using *t-m* instead of *t-1* leads to simple seasonal model:

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

$$\begin{pmatrix} l_t \\ c_t \end{pmatrix} = \begin{pmatrix} 1 & -(1-\alpha_1) \\ 1 & 1-\alpha_0 \end{pmatrix} \begin{pmatrix} l_{t-m} \\ c_{t-m} \end{pmatrix} + \begin{pmatrix} \beta_0 - \beta_1 \\ \beta_0 + \beta_1 \end{pmatrix} \varepsilon_t$$

- This model can produce all the trajectories seasonally when level is close to zero
- It retains all the properties of the original CES

General seasonal CES model

- Combining the original CES with the simple seasonal:

$$y_t = l_{0,t-1} + l_{1,t-m} + \varepsilon_t$$

$$\begin{pmatrix} l_{0,t} \\ c_{0,t} \end{pmatrix} = \begin{pmatrix} 1 & -(1-\alpha_1) \\ 1 & 1-\alpha_0 \end{pmatrix} \begin{pmatrix} l_{0,t-1} \\ c_{0,t-1} \end{pmatrix} + \begin{pmatrix} \alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \end{pmatrix} \varepsilon_t$$

$$\begin{pmatrix} l_{1,t} \\ c_{1,t} \end{pmatrix} = \begin{pmatrix} 1 & -(1-\beta_1) \\ 1 & 1-\beta_0 \end{pmatrix} \begin{pmatrix} l_{1,t-m} \\ c_{1,t-m} \end{pmatrix} + \begin{pmatrix} \beta_0 - \beta_1 \\ \beta_0 + \beta_1 \end{pmatrix} \varepsilon_t$$

- This model can produce all trend and seasonality types

Lancaster Centre for Forecasting
www.forecasting-centre.com

General seasonal CES model

- The model has the same structure as state-space ETS:

$$y_t = w' x_{t-1} + \varepsilon_t$$
$$x_t = F x_{t-1} + g \varepsilon_t$$

- where:

$$x_t = \begin{pmatrix} l_{0,t} \\ c_{0,t} \\ l_{1,t} \\ c_{1,t} \end{pmatrix} \quad x_{t-1} = \begin{pmatrix} l_{0,t-1} \\ c_{0,t-1} \\ l_{1,t-m} \\ c_{1,t-m} \end{pmatrix} \quad F = \begin{pmatrix} 1 & -(1-\alpha_1) & 0 & 0 \\ 1 & (1-\alpha_0) & 0 & 0 \\ 0 & 0 & 1 & -(1-\beta_1) \\ 0 & 0 & 1 & (1-\beta_0) \end{pmatrix}$$

$$w = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad g = \begin{pmatrix} \alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ \beta_0 - \beta_1 \\ \beta_0 + \beta_1 \end{pmatrix}$$

Lancaster Centre for Forecasting
www.forecasting-centre.com

Model selection in CES

- Likelihood function can be derived:

$$L(g, \sigma^2 | y) = \left(\frac{1}{\sigma \sqrt{2\pi}} \right)^T \exp\left(-\frac{1}{2} \sum_{t=1}^{T} \left(\frac{\varepsilon_t}{\sigma} \right)^2 \right)$$

- Any IC can be used. For example, AIC:

$$AIC = 2k - 2\log\left(L\left(g, \sigma^2 | y\right)\right)$$

- The number of coefficients for these models are:
  - non-seasonal CES: 2 + 2,
  - general seasonal CES: 2 + 4 + 2m

Lancaster Centre for Forecasting
www.forecasting-centre.com

Simulation experiment

- ETS was used as DGP,
- 100 observations in each of the 9 groups,
- ETS, CES and ARIMA were applied to the data,
- “ets” and “auto.arima” from “forecast” package in R,
- “ces.auto” from “CES” package for R (https://github.com/config-i1/CES)
- Number of successfully identified characteristics was calculated.

Simulation experiment

DGP	CES	ETS			ARIMA		
		Overall	Trend	Seasonal	Overall	Trend	Seasonal
$N(5000, 50^2)$	100	99	99	100	56	97	57
ETS(ANN)	100	48	91	95	27	44	51
ETS(MNN)	100	50	94	98	27	50	38
ETS(AAN)	100	67	90	88	45	99	45
ETS(MMN)	100	51	90	93	27	92	31
ETS(ANA)	100	49	82	100	47	47	98
ETS(AAA)	100	80	95	100	88	88	100
ETS(MNM)	100	30	57	100	59	59	96
ETS(MMM)	100	32	91	100	79	79	90
Average	100	56	88	97	51	73	67

Table 1: The percentage of the forecasting models chosen correctly for each data generating process (DGP).

Lancaster Centre for Forecasting
www.forecasting-centre.com

Experiment on monthly M3 data

- Rolling origin on 1428 monthly series,
- Forecasting horizon – 18 observations,
- RO horizon – 24 observations,
- ETS, CES and ARIMA were applied to the data,
- MASE was calculated for each observation.

Lancaster Centre for Forecasting
www.forecasting-centre.com

Experiment on monthly M3 data

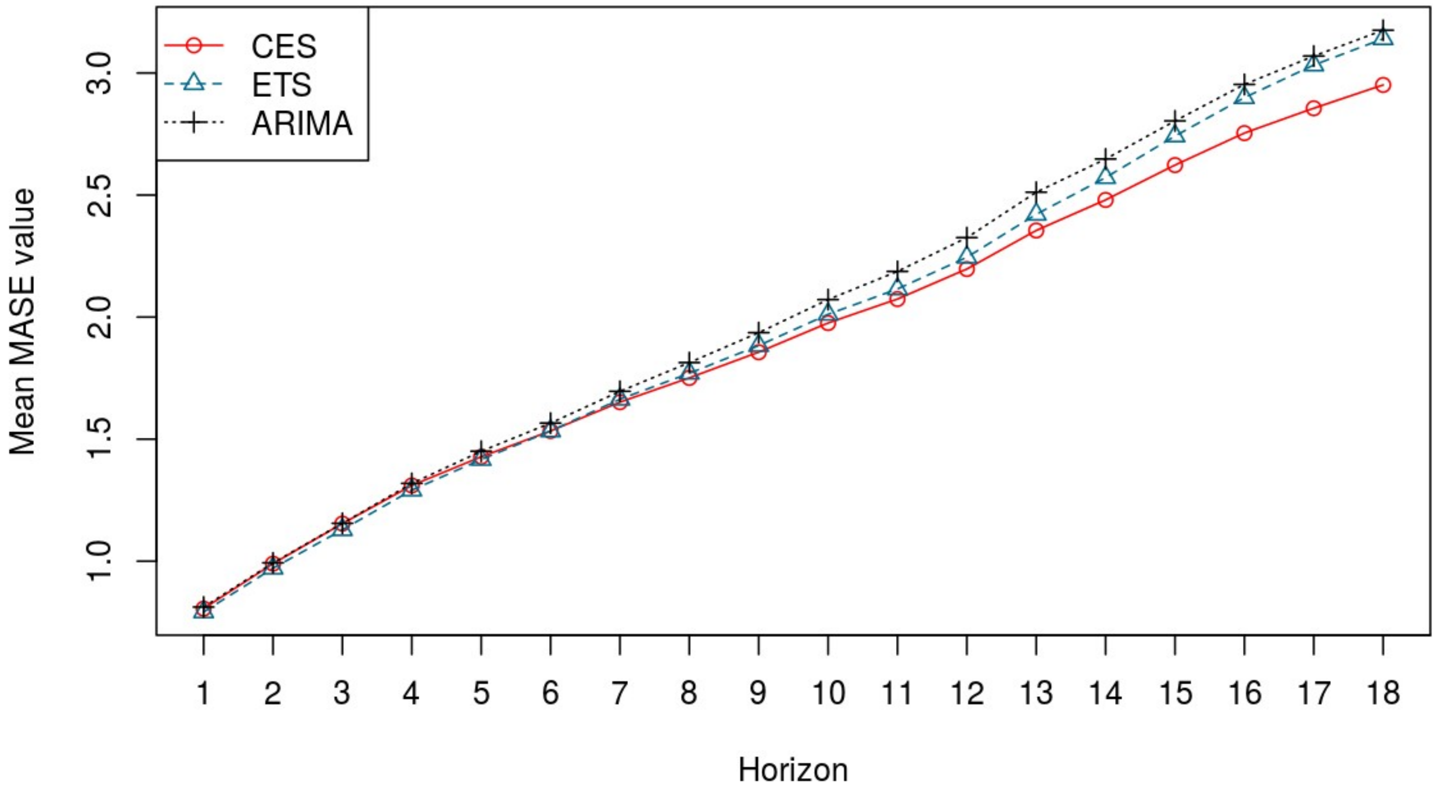
	CES	ETS	ARIMA
Minimum	0.134	**0.084**	0.098
25% quantile	0.665	**0.664**	0.703
Median	**1.049**	1.058	1.093
75% quantile	**2.178**	2.318	2.224
Maximum	**28.440**	53.330	59.343
Mean	**1.922**	1.934	1.967

Table 2: MASE values of competing methods. The smallest values are in bold.

- The difference was statistically significant.

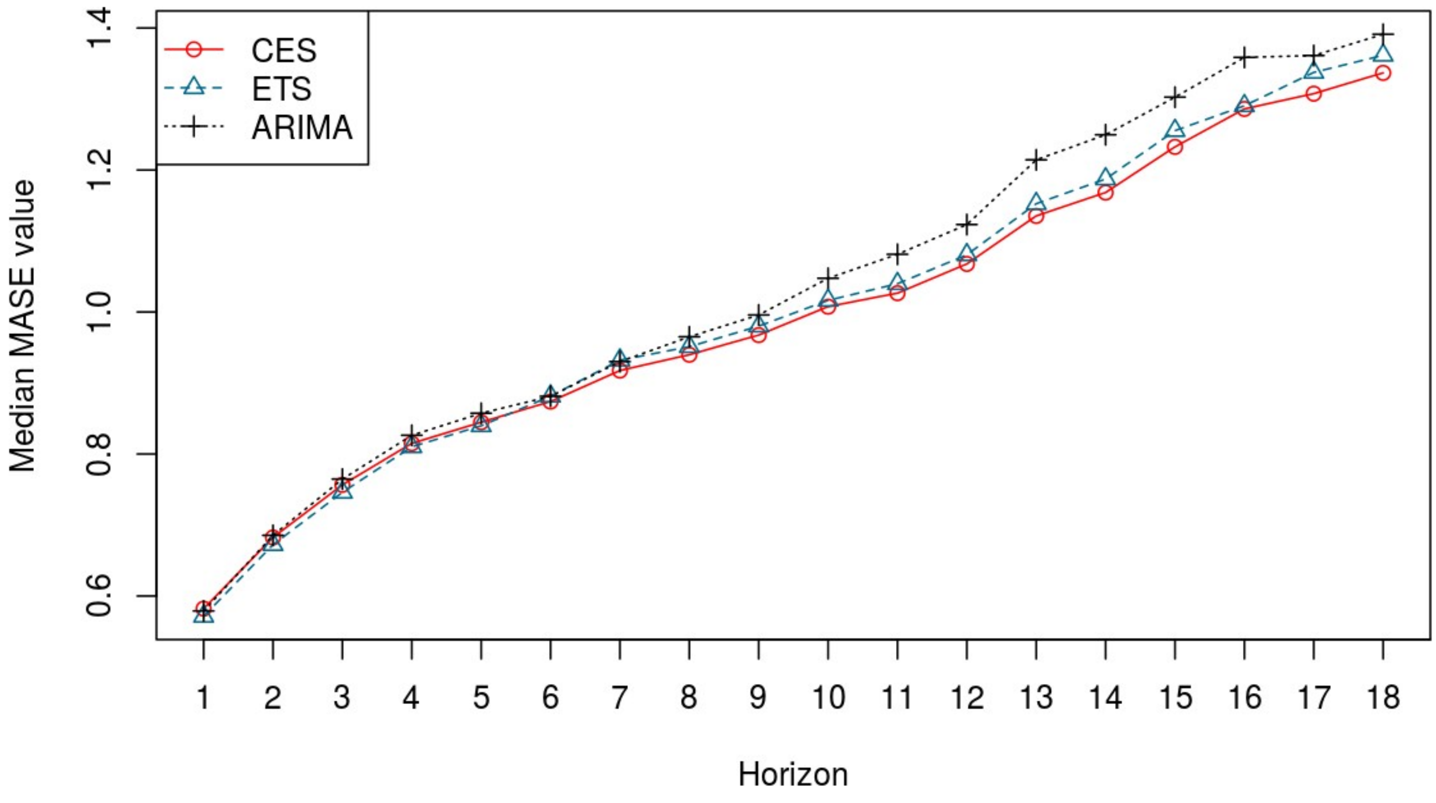
Lancaster Centre for Forecasting
www.forecasting-centre.com

Experiment on monthly M3 data



Lancaster Centre for Forecasting
www.forecasting-centre.com

Experiment on monthly M3 data



Lancaster Centre for Forecasting
www.forecasting-centre.com

Conclusions

- CES
  - can forecast seasonal and non-seasonal data,
  - is able to approximate big variety of trends,
  - can produce additive and multiplicative seasonality,
  - can produce a new type of seasonality,
  - has an efficient model selection mechanism,
  - performs better than ETS and ARIMA on monthly data from M3,
  - makes accurate forecasts on longer horizons.

Lancaster Centre for Forecasting
www.forecasting-centre.com

Thank you!

Ivan Svetunkov,

Lancaster University Management School,
Lancaster Centre for Forecasting

email: i.svetunkov@lancaster.ac.uk

Lancaster Centre for Forecasting
www.forecasting-centre.com