Automatic Forecasting Support System for Business Analytics applications based on Unobserved Components models

DJ Pedregal, MA Villegas, D Villegas

Universidad de Castilla-La Mancha
ETSII (Ciudad Real)
PREDILAB
Diego.Pedregal@uclm.es

EURO2018, Valencia, July 2018

Outline

Objectives of this work:

- Contribute to the dissemination of Unobserved Components models (UC) to a wider audience.

Objectives of this work:

- Contribute to the dissemination of Unobserved Components models (UC) to a wider audience.
- Develop a general Automatic Forecasting Support System based on UC models. This is the first time that automatic identification of UC has been proposed.

Objectives of this work:

- Contribute to the dissemination of Unobserved Components models (UC) to a wider audience.
- Develop a general Automatic Forecasting Support System based on UC models. This is the first time that automatic identification of UC has been proposed.
- Compare the new system with other approaches, like different implementations of other methods, mainly ARIMA and ExponenTial Smoothing.

Objectives of this work:

- Contribute to the dissemination of Unobserved Components models (UC) to a wider audience.
- Develop a general Automatic Forecasting Support System based on UC models. This is the first time that automatic identification of UC has been proposed.
- Compare the new system with other approaches, like different implementations of other methods, mainly ARIMA and ExponenTial Smoothing.
- Show it forecasting automatically on the 166 products of a food franchise chain in Spain.

Objectives of this work:

- Contribute to the dissemination of Unobserved Components models (UC) to a wider audience.
- Develop a general Automatic Forecasting Support System based on UC models. This is the first time that automatic identification of UC has been proposed.
- Compare the new system with other approaches, like different implementations of other methods, mainly ARIMA and ExponenTial Smoothing.
- Show it forecasting automatically on the 166 products of a food franchise chain in Spain.
- Disseminate **SSpace**, a MATLAB toolbox implementing most methods used in this work (Villegas and Pedregal, 2018; `https://bitbucket.org/predilab/sspace-matlab`).

Exponential Smoothing continues to be the most used modeling technique in business and industry, at least in areas ranging from inventory management and scheduling to planning. Several reasons:

- Ad-hoc method, easy to understand and communicate to managers.
- Formal statistical revision in the last 15 years.
- Implemented in many packages, including automatic identification procedures.

Exponential Smoothing continues to be the most used modeling technique in business and industry, at least in areas ranging from inventory management and scheduling to planning. Several reasons:

- Ad-hoc method, easy to understand and communicate to managers.
- Formal statistical revision in the last 15 years.
- Implemented in many packages, including automatic identification procedures.

ARIMA is the second method most used, because it is well-known to many researchers and also there are many packages implementing even automatic identification procedures.

There is a family of models that has been systematically overlooked, namely structural Unobserved Components models (UC). At least five reasons:

There is a family of models that has been systematically overlooked, namely structural Unobserved Components models (UC). At least five reasons:

- General use for signal extraction and very little for forecasting purposes.

There is a family of models that has been systematically overlooked, namely structural Unobserved Components models (UC). At least five reasons:

- General use for signal extraction and very little for forecasting purposes.
- UC models have been developed in academic environments, with no strategy for their dissemination among practitioners for their everyday use in business and industry.

There is a family of models that has been systematically overlooked, namely structural Unobserved Components models (UC). At least five reasons:

- General use for signal extraction and very little for forecasting purposes.
- UC models have been developed in academic environments, with no strategy for their dissemination among practitioners for their everyday use in business and industry.
- The widely-held but not-scientifically-tested feeling that UC models do not really have anything relevant to add to ETS methods.

There is a family of models that has been systematically overlooked, namely structural Unobserved Components models (UC). At least five reasons:

- General use for signal extraction and very little for forecasting purposes.
- UC models have been developed in academic environments, with no strategy for their dissemination among practitioners for their everyday use in business and industry.
- The widely-held but not-scientifically-tested feeling that UC models do not really have anything relevant to add to ETS methods.
- UC models are usually identified by hand, with automatic identification being very rare.

There is a family of models that has been systematically overlooked, namely structural Unobserved Components models (UC). At least five reasons:

- General use for signal extraction and very little for forecasting purposes.
- UC models have been developed in academic environments, with no strategy for their dissemination among practitioners for their everyday use in business and industry.
- The widely-held but not-scientifically-tested feeling that UC models do not really have anything relevant to add to ETS methods.
- UC models are usually identified by hand, with automatic identification being very rare.
- Software is scarcer.

Unobserved Components models

UC models aim at decomposing a vector of time series into meaningful components explicitly, namely trend, cycle, seasonal, and irregular. Other components may be considered as well, typically cycles and components relating the output variables to inputs modeled as linear regressions, transfer functions or non-linear relationships. A general representation is given by

$$z_t = T_t + C_t + S_t + f(u_t) + I_t \tag{1}$$

where z_t is a vector of time series, T_t, C_t, S_t and I_t stand for vectors of trends, cycles, seasonal and irregular components, respectively. The term $f(u_t)$ models the relation between a vector of inputs u_t and the outputs z_t.

In this structural approach, the key issue is to select appropriate dynamic models for each of the components involved. In general terms, trends should have at least one unit root, seasonal components should show up some kind of stochastic sinusoidal dynamic behaviour, and irregular components should be either white or coloured noise.

In this structural approach, the key issue is to select appropriate dynamic models for each of the components involved. In general terms, trends should have at least one unit root, seasonal components should show up some kind of stochastic sinusoidal dynamic behaviour, and irregular components should be either white or coloured noise.

The UC model is set up in a State Space framework and the Kalman Filter and Smoothing algorithms provide the optimal estimates of the state vector and its covariance matrices. Maximum Likelihood estimation may be done on the general formulation of SS systems, etc.

Forecasting methods (I)

A number of methods has been compared in this work, all based on automatic identification techniques:

- Benchmarks: i) **Naïve**, Seasonal Naïve and ii) **AR**, automatic AR models with order optimised by minimising the BIC.

Forecasting methods (I)

A number of methods has been compared in this work, all based on automatic identification techniques:

- Benchmarks: i) **Naïve**, Seasonal Naïve and ii) **AR**, automatic AR models with order optimised by minimising the BIC.
- Automatic identification of ARIMA models with three implementations:
  - **R-ARIMA**: Hyndman and Khandakar (2008) in R package 'forecast'.

Forecasting methods (I)

A number of methods has been compared in this work, all based on automatic identification techniques:

- Benchmarks: i) **Naïve**, Seasonal Naïve and ii) **AR**, automatic AR models with order optimised by minimising the BIC.
- Automatic identification of ARIMA models with three implementations:
  - **R-ARIMA**: Hyndman and Khandakar (2008) in R package 'forecast'.
  - **ARIMA**: A modified version of Hyndman and Khandakar (2008) to avoid problems with differentation order, implemented in **SSpace**.

Forecasting methods (I)

A number of methods has been compared in this work, all based on automatic identification techniques:

- Benchmarks: i) **Naïve**, Seasonal Naïve and ii) **AR**, automatic AR models with order optimised by minimising the BIC.
- Automatic identification of ARIMA models with three implementations:
  - **R-ARIMA**: Hyndman and Khandakar (2008) in R package 'forecast'.
  - **ARIMA**: A modified version of Hyndman and Khandakar (2008) to avoid problems with differentation order, implemented in **SSpace**.
  - **TRAMO**: Gómez and Maravall (2001) implemented in TRAMO with and without outliers identification.

Forecasting methods (II)

- Exponential smoothing:

Forecasting methods (II)

- Exponential smoothing:
  - **ETS**: Implemented in **SSpace** (by inimising BIC):
    - Level: no level, additive, damped additive.
    - Growth: no growth, growth.
    - Seasonal: no seasonal, seasonal.
    - AR models of order 0 to 3.

Forecasting methods (II)

- Exponential smoothing:
  - **ETS**: Implemented in **SSpace** (by inimising BIC):
    - Level: no level, additive, damped additive.
    - Growth: no growth, growth.
    - Seasonal: no seasonal, seasonal.
    - AR models of order 0 to 3.
  - **R-ETS**: implemented in the R package 'forecast' (Hyndman and Khandakar, 2008).

Forecasting methods (II)

- Exponential smoothing:
  - **ETS**: Implemented in **SSpace** (by inimising BIC):
    - Level: no level, additive, damped additive.
    - Growth: no growth, growth.
    - Seasonal: no seasonal, seasonal.
    - AR models of order 0 to 3.
  - **R-ETS**: implemented in the R package 'forecast' (Hyndman and Khandakar, 2008).
- Combination methods: i) **Mean** and ii) **Median** of AR, ARIMA, UC and ETS methods estimated with **SSpace**.

Forecasting methods (and III)

- **UC**: Unobserved Components models (in **SSpace**, minimising BIC):
  - Trend: Random Walk, Integrated Random Walk, Local Linear Trend, Damped trend.

Forecasting methods (and III)

- **UC**: Unobserved Components models (in **SSpace**, minimising BIC):
  - Trend: Random Walk, Integrated Random Walk, Local Linear Trend, Damped trend.
  - Seasonal: No seasonal, trigonometric seasonal with common variance for all harmonics and trigonometric seasonal without common variance for all harmonics.

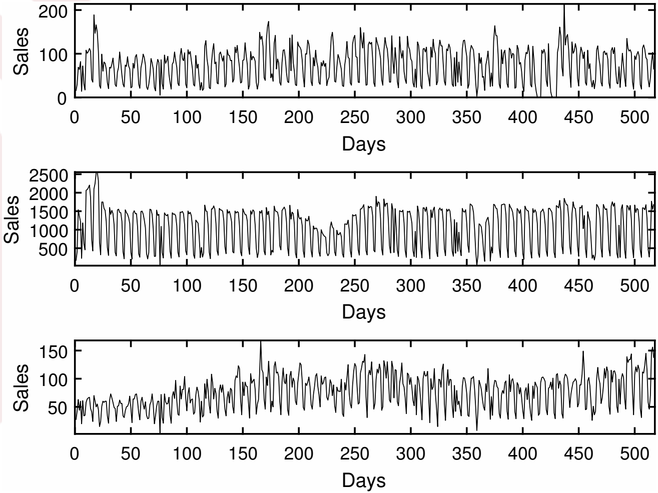
Forecasting methods (and III)

- **UC**: Unobserved Components models (in **SSpace**, minimising BIC):
  - Trend: Random Walk, Integrated Random Walk, Local Linear Trend, Damped trend.
  - Seasonal: No seasonal, trigonometric seasonal with common variance for all harmonics and trigonometric seasonal without common variance for all harmonics.
  - Irregular: White noise or AR coloured noise orders up to 3rd order.

Data

The proposed evaluation of the models was carried out on the 166 products sales from a food franchise in Spain. The company specializes in selling everyday dishes made from natural products at affordable prices in take-away and take-in formats. 517 daily sales observations were made available for each product with 414 observations used for in-sample estimation and the remainder for out-of-sample evaluation. A set of 90, 14 days ahead forecast rounds was carried out for each product.

Data


Sales
200
100
0
0 50 100 150 200 250 300 350 400 450 500
Days
Sales
2500
2000
1500
1000
500
0 50 100 150 200 250 300 350 400 450 500
Days
Sales
150
100
50
0 50 100 150 200 250 300 350 400 450 500
Days

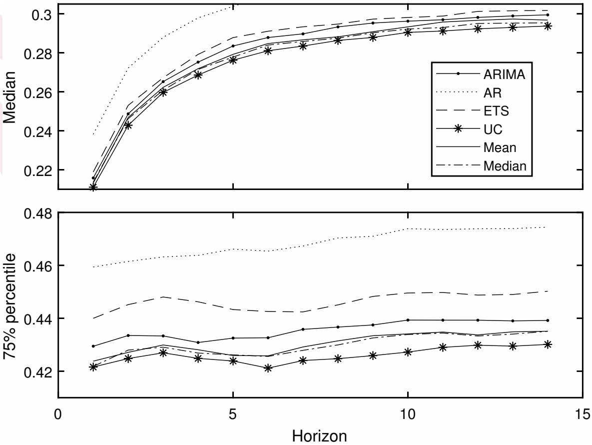
Metrics

Cumulative Absolute Scaled Error:

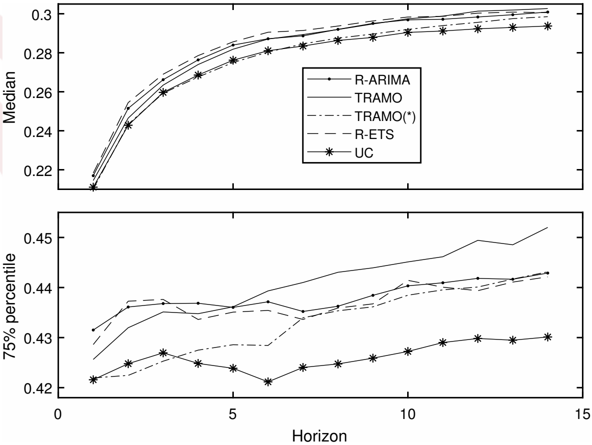
$$\mathrm{CsE}_h = \frac{1}{h} \sum_{l=1}^{h} sE_h$$

where the Absolute Scaled Error is

$$\mathrm{sE}_h = \frac{\left| z_{T+h} - \hat{z}_{T+h} \right|}{\frac{1}{T} \sum_{i=1}^{T} z_i}$$



Median (top panel) and 75 % percentile (bottom panel) of CsE_h for all models implemented in **SSpace**.



Median (top panel) and 75 % percentile (bottom panel) of CsE_h for alternative models.

Conclusions related to all models implemented in **SSpace**:

- As expected, forecasting performance worsens with longer horizons.

Conclusions related to all models implemented in **SSpace**:

- As expected, forecasting performance worsens with longer horizons.
- All methods outperform both **Naïve** and **AR** benchmarks by a wide margin. **AR** systematically outperforms the seasonal **Naïve**.

Conclusions related to all models implemented in **SSpace**:

- As expected, forecasting performance worsens with longer horizons.
- All methods outperform both **Naïve** and **AR** benchmarks by a wide margin. **AR** systematically outperforms the seasonal **Naïve**.
- **UC** is the best model for all forecast horizons, followed by **ARIMA** and **ETS**.

Conclusions related to all models implemented in **SSpace**:

- As expected, forecasting performance worsens with longer horizons.
- All methods outperform both **Naïve** and **AR** benchmarks by a wide margin. **AR** systematically outperforms the seasonal **Naïve**.
- **UC** is the best model for all forecast horizons, followed by **ARIMA** and **ETS**.
- Simple forecast combination methods outperform **ARIMA** and **ETS** but do not improve on **UC**. The performance of **ARIMA** and **ETS** is very similar in this case.

Conclusions related to all models implemented in **SSpace**:

- As expected, forecasting performance worsens with longer horizons.
- All methods outperform both **Naïve** and **AR** benchmarks by a wide margin. **AR** systematically outperforms the seasonal **Naïve**.
- **UC** is the best model for all forecast horizons, followed by **ARIMA** and **ETS**.
- Simple forecast combination methods outperform **ARIMA** and **ETS** but do not improve on **UC**. The performance of **ARIMA** and **ETS** is very similar in this case.
- The median of CsE_h and sE_h shows that most models offer very similar results. However, as the bottom panel shows, differences are more apparent when the 75 % percentile is considered. The advantage of UC models is clearer in these cases.

Conclusions related to other pieces of software:

- **R-AR** performs worse than its **AR** counterpart in **SSpace**.

Conclusions related to other pieces of software:

- **R-AR** performs worse than its **AR** counterpart in **SSpace**.
- On the other hand, **R-ETS** outperforms its **ETS** equivalent in **SSpace**. This may be due mainly to the fact that **R-ETS** allows for general ARMA models for the observational noise, while **ETS** constrain them to AR up to order 3. There are also other minor differences coming from differences in libraries used by MATLAB and R, different initial conditions for parameter searches, different initial conditions for states when starting the Kalman Filter.

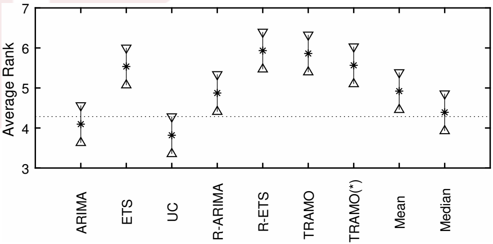
Conclusions related to other pieces of software:

- **R-AR** performs worse than its **AR** counterpart in **SSpace**.
- On the other hand, **R-ETS** outperforms its **ETS** equivalent in **SSpace**. This may be due mainly to the fact that **R-ETS** allows for general ARMA models for the observational noise, while **ETS** constrain them to AR up to order 3. There are also other minor differences coming from differences in libraries used by MATLAB and R, different initial conditions for parameter searches, different initial conditions for states when starting the Kalman Filter.
- Results shown for **R-ARIMA** are obtained constraining the seasonal differences to 1 in the automatic procedure. The errors are considerably larger when this parameter is selected fully automatically with formal unit roots.

- **TRAMO** and **TRAMO(*)** are worse than **R-ARIMA** and **R-ETS** for longer horizons, although better for shorter ones.

- **TRAMO** and **TRAMO(*)** are worse than **R-ARIMA** and **R-ETS** for longer horizons, although better for shorter ones.

Formal statistical tests to evaluate whether differences are statistically significant. This can be done with rank tests (Koning *al*., 2005) that rank methods by forecast error measurement (Mean Absolute Error in this case)



Thank you for your attention!

e-mail: diego.pedregal@uclm.es
blog: www.uclm.es/profesorado/diego/

This work has been supported by the Spanish Ministerio de Economía y Competitividad, under Research Grant no. DPI2015-64133-R (MINECO/FEDER/UE)