Modelling multiple seasonalities across hierarchical aggregation levels

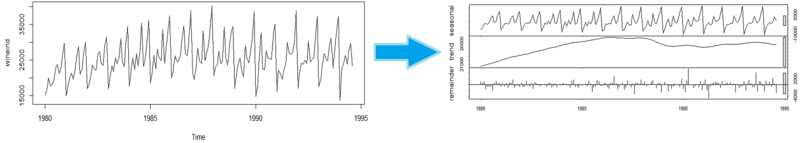
Daniel Waller, John Boylan, Nikolaos Kourentzes

Lancaster University Management School

20/06/2016

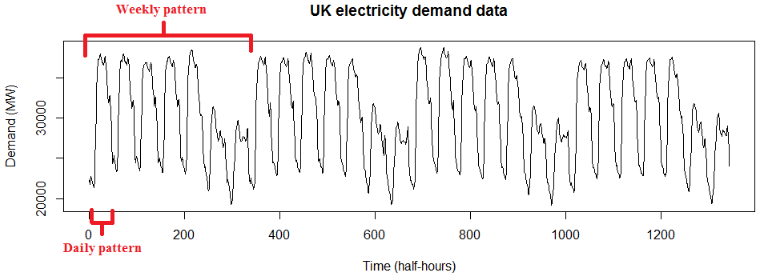
What are multiple seasonalities?

- Time series are often broken down into three components:
  - Trend - the rate of increase/decrease of the series.
  - Seasonality - a pattern which repeats regularly over a fixed period.
  - Error - a random quantity.
- Implicit assumption that there is only one seasonal pattern.
- Holt-Winters exponential smoothing based on this assumption, as are many other base forecasting methods.



What are multiple seasonalities?

- Sometimes there is clearly more than one seasonal influence on the time series.
- For instance, half-hour of day and half-hour of week both have a seasonal effect on the demand of electricity in the series below.



Literature

Exponential-smoothing based approaches in the literature:

- Double/triple seasonal ES (Taylor 2003, 2010).
- Intraday ES (Gould 2008)
- TBATS (De Livera et al. 2011)
- Parsimonious ES (Taylor and Snyder 2012).

Main motivation has been short-term load forecasting for electricity (other utilities as well).

A new motivation - retail

Demand in retail may be subject to multiple seasonal influences:

- Can we use multiple seasonal techniques?
- What adaptations need to be made?

Retail forecasting differs from short-term electricity load forecasting in a few respects:

- Exogenous variables (price, promotions, etc.)
- Substitutable/complementary product effects.
- More hierarchies/levels to forecast.

LCF

Double-seasonal ES

Adaptation of Taylor (2003) to single-seasonal ES.
Additive version:

Level:	$l_t = \alpha(y_t - s_{t-m_1} - d_{t-m_2}) + (1-\alpha)(l_{t-1} + b_{t-1})$
Trend:	$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}$
Seas 1:	$s_t = \gamma(y_t - l_{t-1} - b_{t-1} - d_{t-m_2}) + (1-\gamma)s_{t-m_1}$
Seas 2:	$d_t = \delta(y_t - l_{t-1} - b_{t-1} - s_{t-m_1}) + (1-\delta)d_{t-m_2}$

with forecasting equation:

$$\hat{y}_{t+1} = l_t + b_t + s_{t+1-m_1} + d_{t+1-m_2} + \phi(y_t - l_{t-1} - b_{t-1} - s_{t-m_1} - d_{t-m_2})$$

for a horizon of 1.

Double-seasonal ES

Adaptation of Taylor(2003) to single-seasonal ES.
Multiplicative version:

$$
\begin{aligned}
\text{Level:} \quad & l_t = \alpha \frac{y_t}{s_{t-m_1} d_{t-m_2}} + (1-\alpha)(l_{t-1} + b_{t-1}) \\
\text{Trend:} \quad & b_t = \beta(l_t - l_{t-1}) + (1-\beta) b_{t-1} \\
\text{Seas 1:} \quad & s_t = \gamma \frac{y_t}{l_t d_{t-m_2}} + (1-\gamma) s_{t-m_1} \\
\text{Seas 2:} \quad & d_t = \delta \frac{y_t}{l_t s_{t-m_1}} + (1-\delta) d_{t-m_2}
\end{aligned}
$$

with forecasting equation:

$$\hat{y}_{t+1} = (l_t + b_t) s_{t+1-m_1} d_{t+1-m_2} + \phi(y_t - (l_{t-1} + b_{t-1}) s_{t-m_1} d_{t-m_2})$$

for a horizon of 1.

Parsimonious ES

Proposed by Taylor and Snyder (2012), building on the work of Gould (2008):

$$e_t = y_t - \sum_{i=1}^{M} I_{it} s_{i,t-1}$$
$$s_{it} = s_{i,t-1} + (\alpha + \omega I_{it}) e_t \qquad \text{i = 1,2,...,M}$$
$$I_{it} = \begin{cases} 1 \text{ if period t occurs in season i} \\ 0 \text{ otherwise} \end{cases}$$

with forecasting equation:

$$\hat{y}_{t+1} = \sum_{i=1}^{M} I_{i,(t+1)} s_{i,t} + \phi e_t$$

LCF

Parsimonious ES

Advantages

- Allows unconstrained clustering of periods.
- Fewer number of initial terms to estimate.

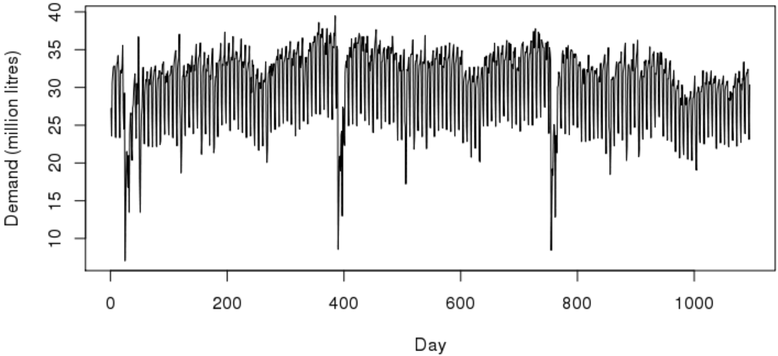
Limitations

- Cannot incorporate exogenous information.
- Clustering of seasons non-automatic/non-scalable.

Empirical testing

We use the example of fuel - below is a plot of demand:

- Daily totals
- Aggregated over a sample of retail sites



Empirical setup

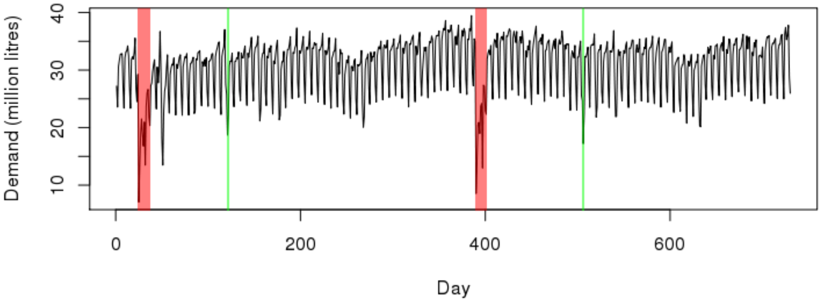
- Comparing three methods:
  - Single-seasonal ES (benchmark)
  - Double-seasonal ES
  - Parsimonious ES
- Estimation: 1st 2 years (730 obs.)
- Holdout: Last year (365 obs.)
- Horizon - Up to 21 days

LCF

PES Model Selection

23 seasons:

- 14 seasons around Christmas
- 2 seasons around Easter
- 7 seasons for ‘normal’ day of week



Results

MAPE for one-step-ahead forecasts:

Table: Excluding Christmas/Easter

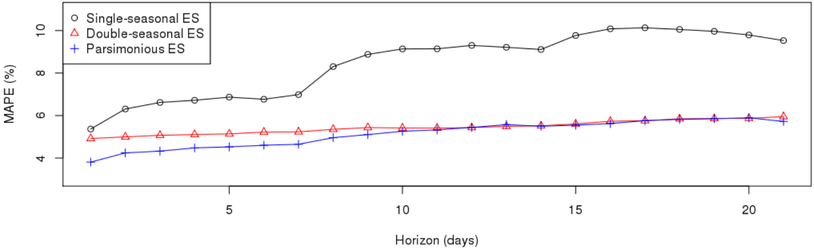
	MAPE	MAE
PES	**3.33%**	**936,422**
DSHW	4.79%	1,388,141
ES	3.95%	1,131,649

Table: Christmas/Easter only

	MAPE	MAE
PES	14.28%	3,286,438
DSHW	**8.80%**	**1,800,825**
ES	36.20%	4,908,546

Accuracy vs. Horizon

Graph shows overall MAPE against horizons of up to 21 observations.

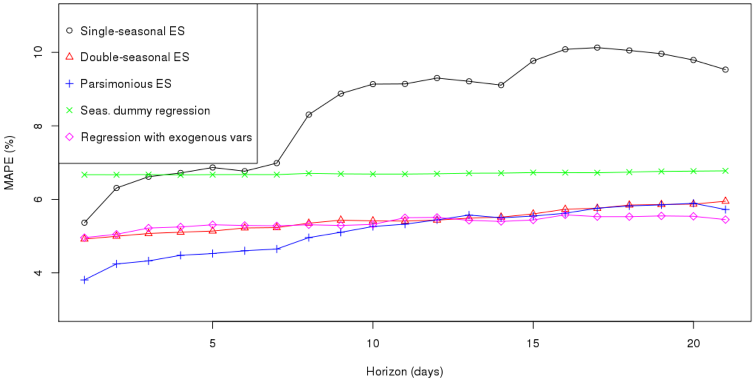


Multivariate testing

Compare univariate results to 2 regression models:

- Seasonal dummies only.
- Inclusion of exogenous information:
  - Price
  - Weather vars x11
- Use naïve for future values of exogenous predictors.

Results



- PES best at short horizons.
- Regression is robust at long horizons.

Summary

Conclusions

- Multi-seasonal methods may hold promise in retail.
- Univariate PES is most accurate at short horizons.
- Longer horizons/short data histories potentially problematic.

Research Plan

- Extension of PES to multivariate case.
- Scalable/automatic approach to season clustering.
- Multiple series/hierarchies.

Any questions?

Thank you for your attention!

Daniel Waller
Lancaster Centre for Forecasting
Lancaster University Management School

d.waller1@lancaster.ac.uk