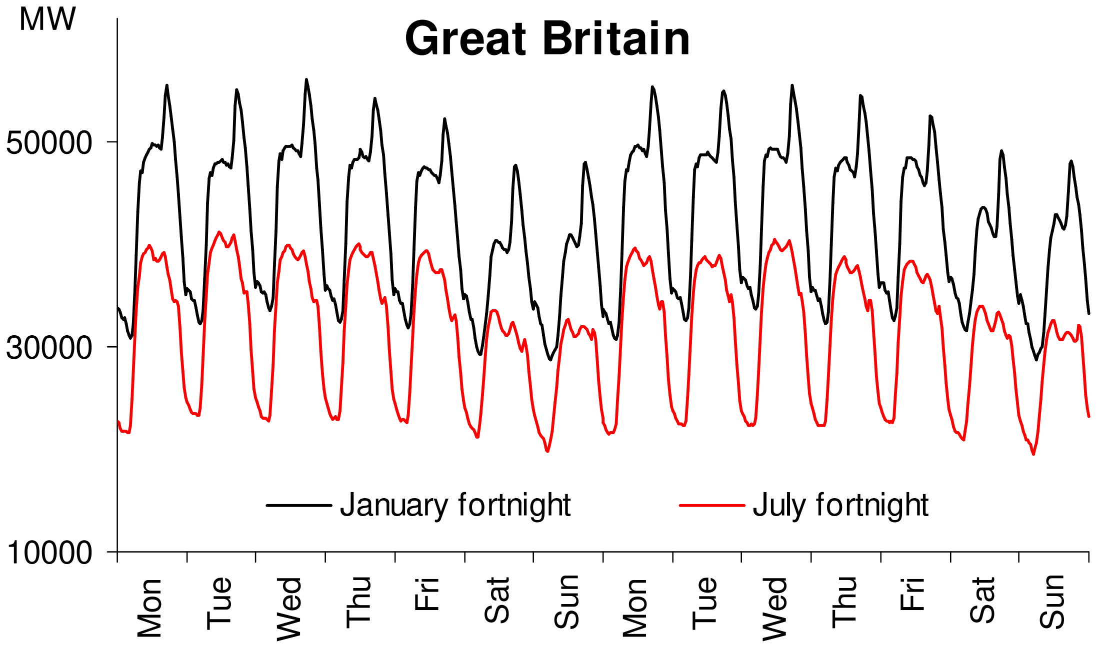
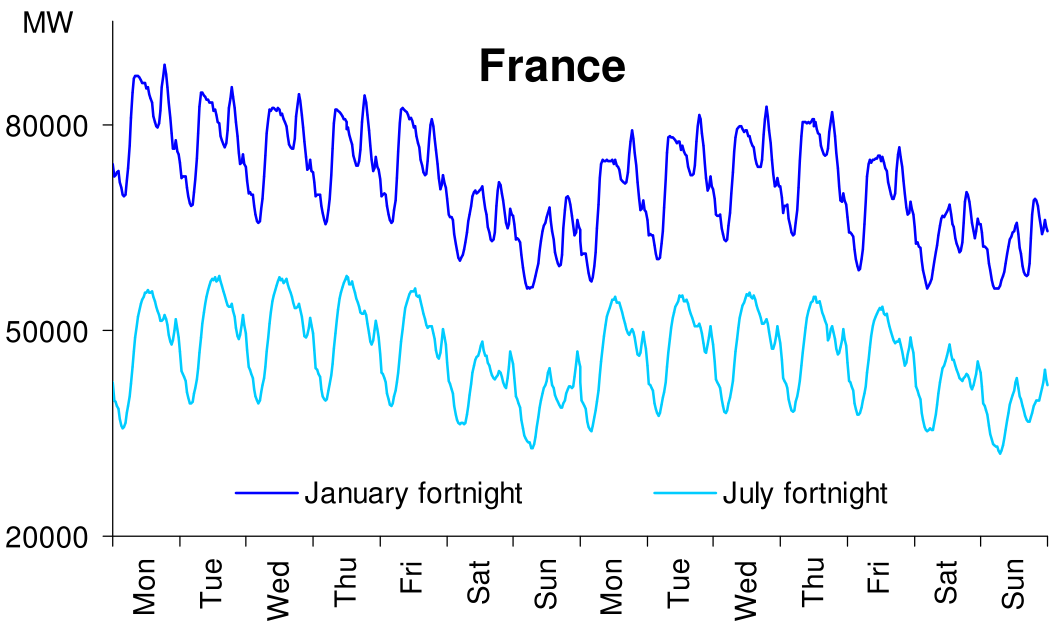
Short-Term Load Forecasting with Exponentially Weighted Methods

James W. Taylor
Saïd Business School
University of Oxford

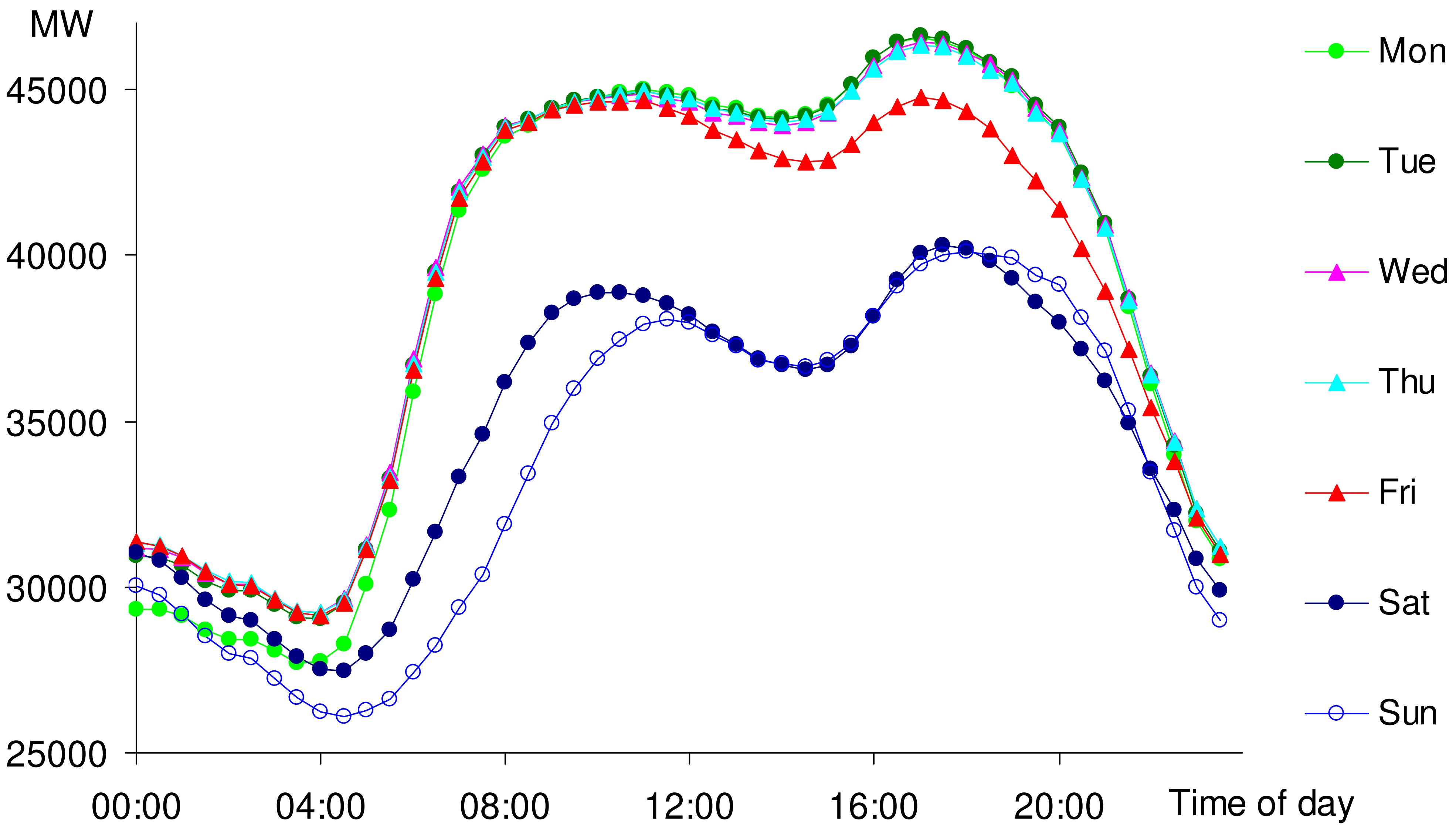
EDF and INRIA Workshop
5 April 2012

Half-hourly Load





Average GB Intraday Cycles

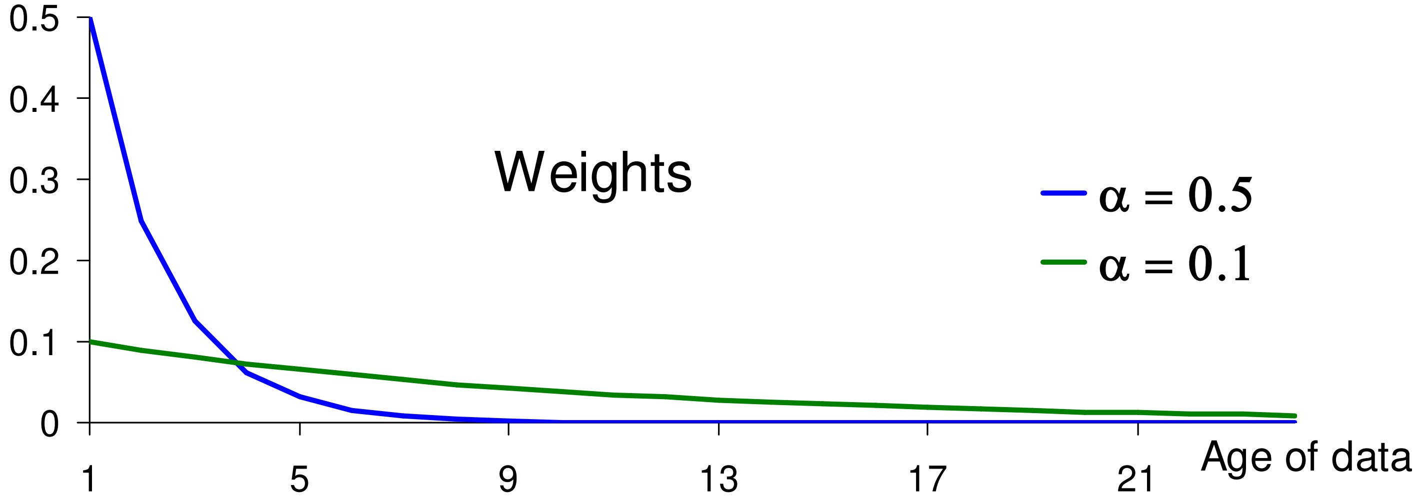


Methods

1. Double seasonal ES
2. Intraday cycle ES
3. DWR with trigonometric terms
4. DWR splines
5. Spline-based ES
6. SVD-based ES

Simple ES

$$y_t = l_{t-1} + e_t$$
$$l_t = l_{t-1} + \alpha\, e_t$$



Note:

$$\begin{aligned} l_t &= \alpha\, y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \alpha(1-\alpha)^3 y_{t-3} + \ldots \\ &= \alpha\, y_t + (1-\alpha)l_{t-1} \\ &= l_{t-1} + \alpha\,(y_t - l_{t-1}) \end{aligned}$$

1. Double Seasonal ES

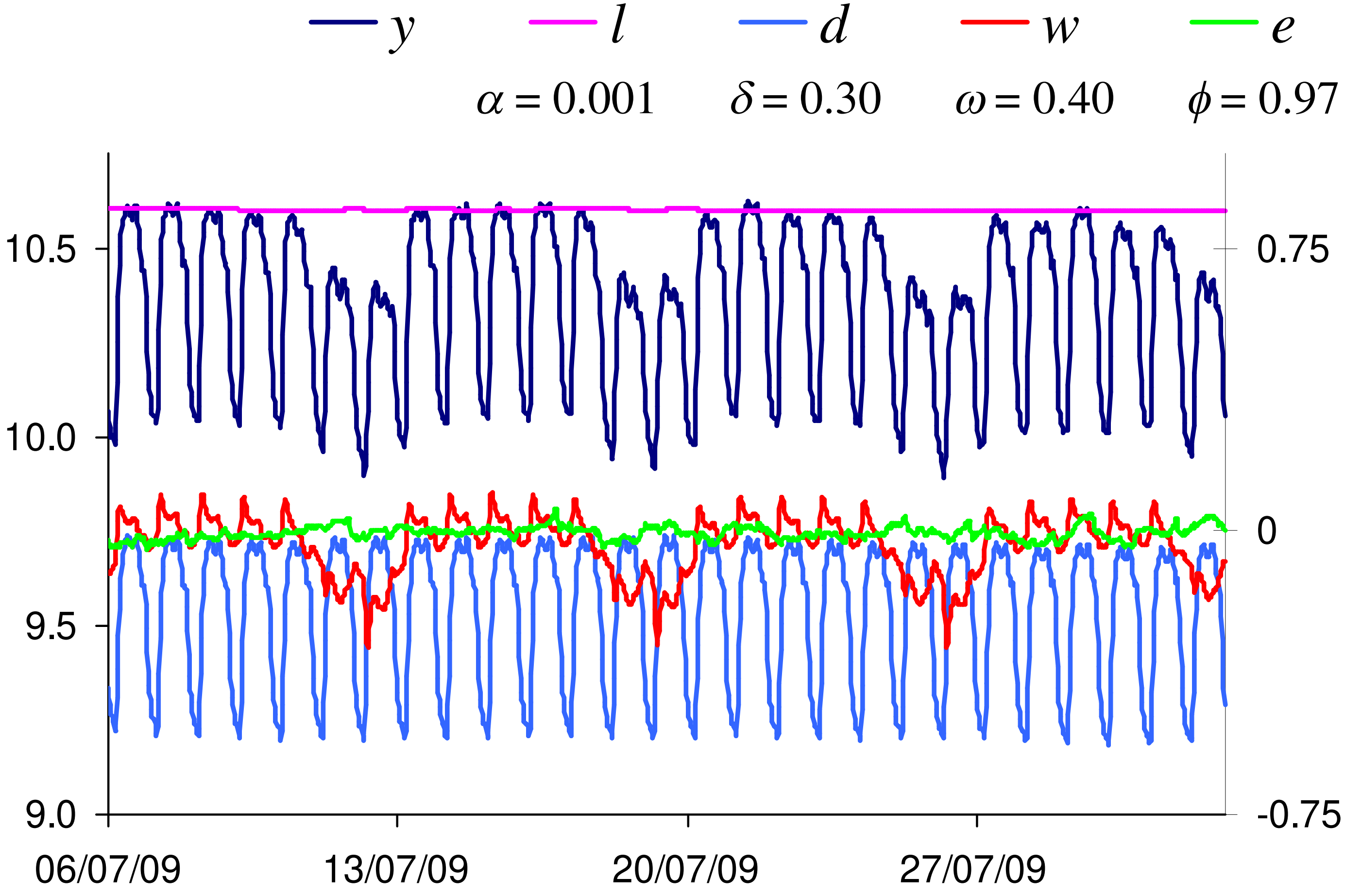
$$y_t = l_{t-1} + d_{t-48} + w_{t-336} + \phi e_{t-1} + \varepsilon_t$$

$$e_t = y_t - \left(l_{t-1} + d_{t-48} + w_{t-336}\right)$$

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

$$d_t = d_{t-48} + \delta e_t$$

$$w_t = w_{t-336} + \omega e_t$$



2. Intraday Cycle ES (Gould et al. 2008)

- Different intraday cycles:
  Mon (d_{1t}), Tue-Thu (d_{2t}), Fri (d_{3t}), Sat (d_{4t}) and Sun (d_{5t}).

$$I_{it} = \begin{cases} 1 & \text{if period } t \text{ occurs in a day of type } i \\ 0 & \text{otherwise} \end{cases}$$

$$y_t = l_{t-1} + \sum_{i=1}^{5} I_{it}\, d_{i,t-48} + \phi\, e_{t-1} + \varepsilon_t$$

$$e_t = y_t - \left(l_{t-1} + \sum_{i=1}^{5} I_{it}\, d_{i,t-48} \right)$$

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

$$d_{it} = d_{i,t-48} + \gamma_{ij} \sum_{j=1}^{5} I_{jt}\, e_t \qquad (i = 1, 2, \ldots, 5)$$

Methods

1. Double seasonal ES
2. Intraday cycle ES
3. DWR with trigonometric terms
4. DWR splines
5. Spline-based ES
6. SVD-based ES

Discount Weighted Regression (Ameen & Harrison 1984)

- EWR: $y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \varepsilon_t$

$$\sum_{i=1}^{t} \lambda^{t-i}\left(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}\right)^2$$

- DWR allows different decay for each parameter:

$$\hat{\boldsymbol{\beta}}_t = \hat{\boldsymbol{\beta}}_{t-1} + \mathbf{Q}_t^{-1}\boldsymbol{x}_t e_t$$

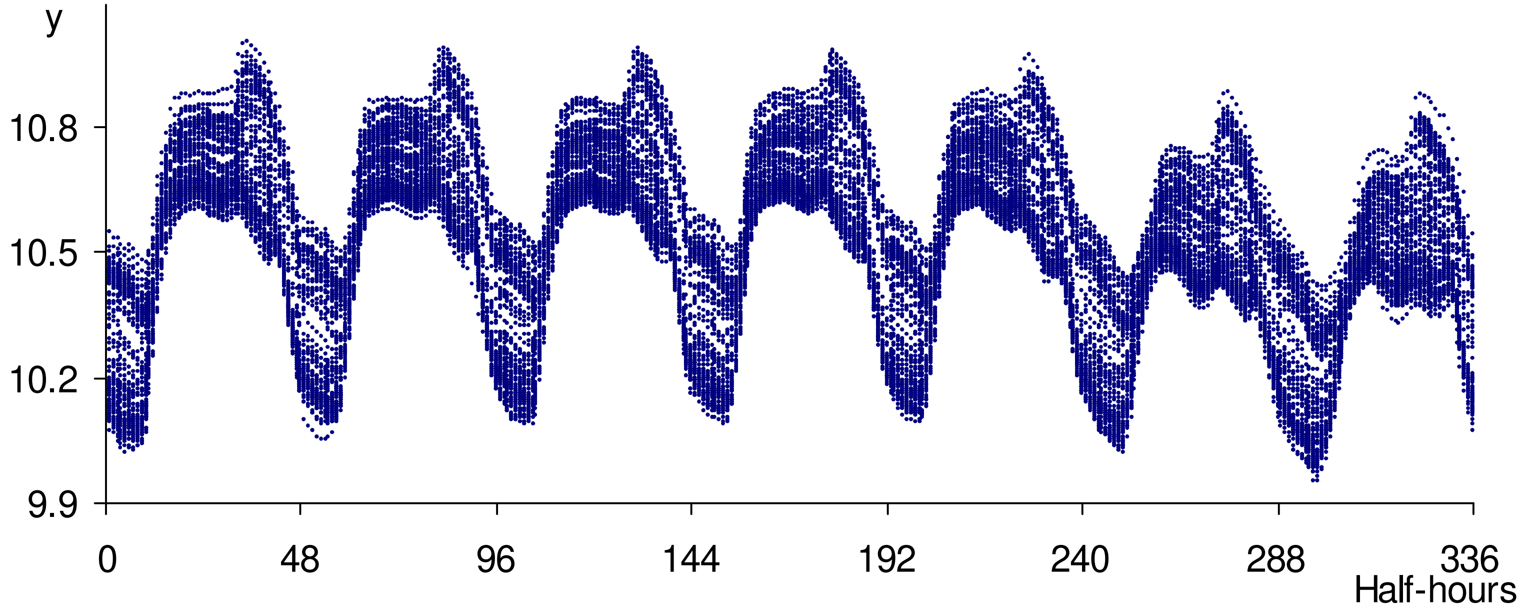
$$e_t = y_t - \boldsymbol{x}_t'\hat{\boldsymbol{\beta}}_t$$

$$\mathbf{Q}_t = \boldsymbol{\lambda}^{\frac{1}{2}}\mathbf{Q}_{t-1}\boldsymbol{\lambda}^{\frac{1}{2}} + \boldsymbol{x}_t\boldsymbol{x}_t'$$

$$\boldsymbol{\lambda}^{\frac{1}{2}} = \mathrm{diag}\left(\lambda_1^{\frac{1}{2}}, \lambda_2^{\frac{1}{2}}, \ldots, \lambda_M^{\frac{1}{2}}\right), \quad 0 \le \lambda_i \le 1 \;\; \forall i$$

3. DWR with Trigonometric Terms

$$y_t = b_0 + \sum_{i=1}^{M_1}\left(b_{1i}\sin\left(\frac{2i\pi t}{48}\right) + b_{2i}\cos\left(\frac{2i\pi t}{48}\right)\right) + \sum_{i=1}^{M_2}\left(b_{3i}\sin\left(\frac{2i\pi t}{336}\right) + b_{4i}\cos\left(\frac{2i\pi t}{336}\right)\right) + \varepsilon_t$$



$\lambda_1 \approx 0.977$ for b_0

$\lambda_2 \approx 0.994$ for b_{1i} and b_{2i}

$\lambda_3 \approx 0.998$ for b_{3i} and b_{4i}

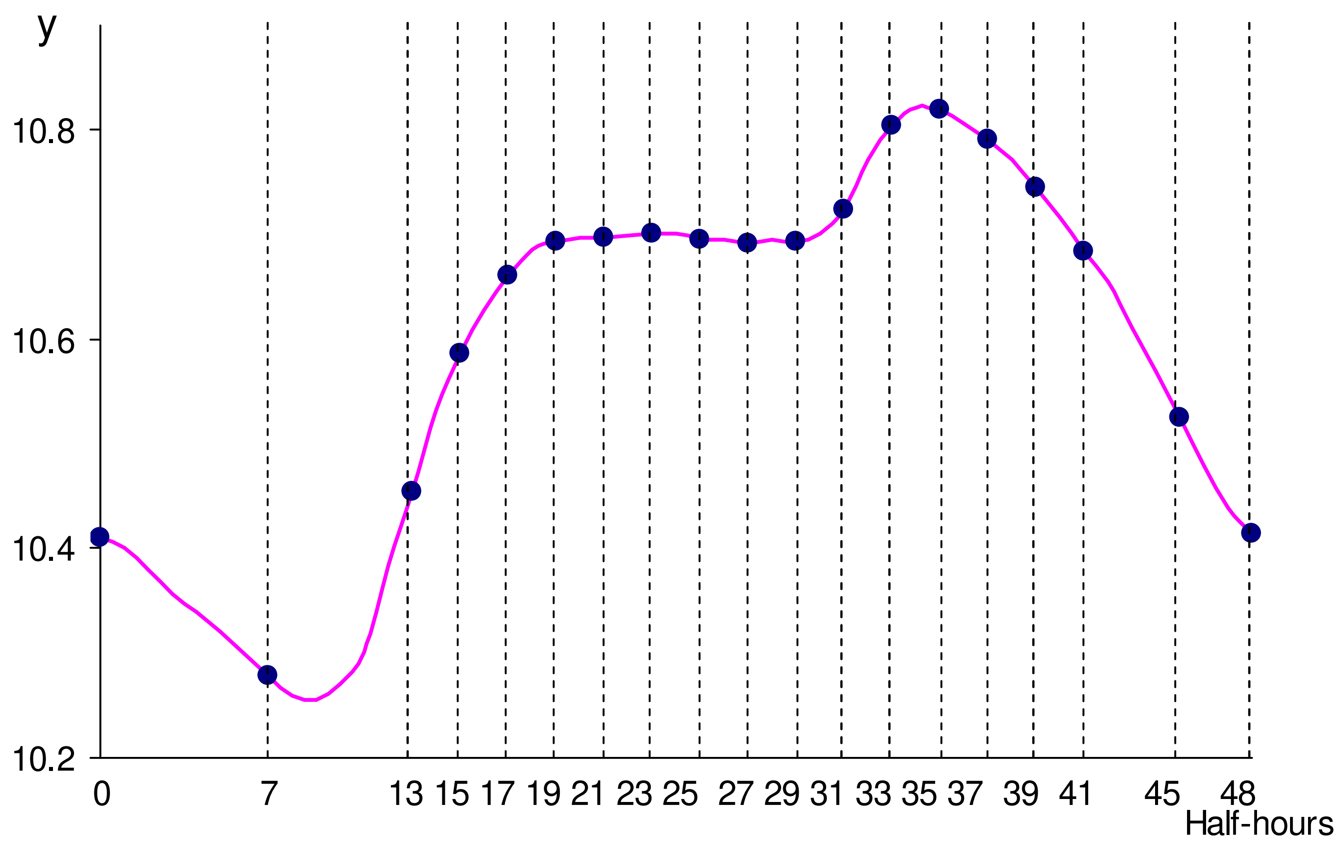
Splines

- Cubic spline joined at (x_i^*, s_i^*).

  Any point on spline is linear combination of values at knots x_i^*:

  $$f(x) = \boldsymbol{w}'\boldsymbol{s}^*$$

  $\boldsymbol{w}$ calculated analytically.



OLS Regression Splines (Poirier 1973)

- Cubic spline joined at (x_i^*, s_i^*).

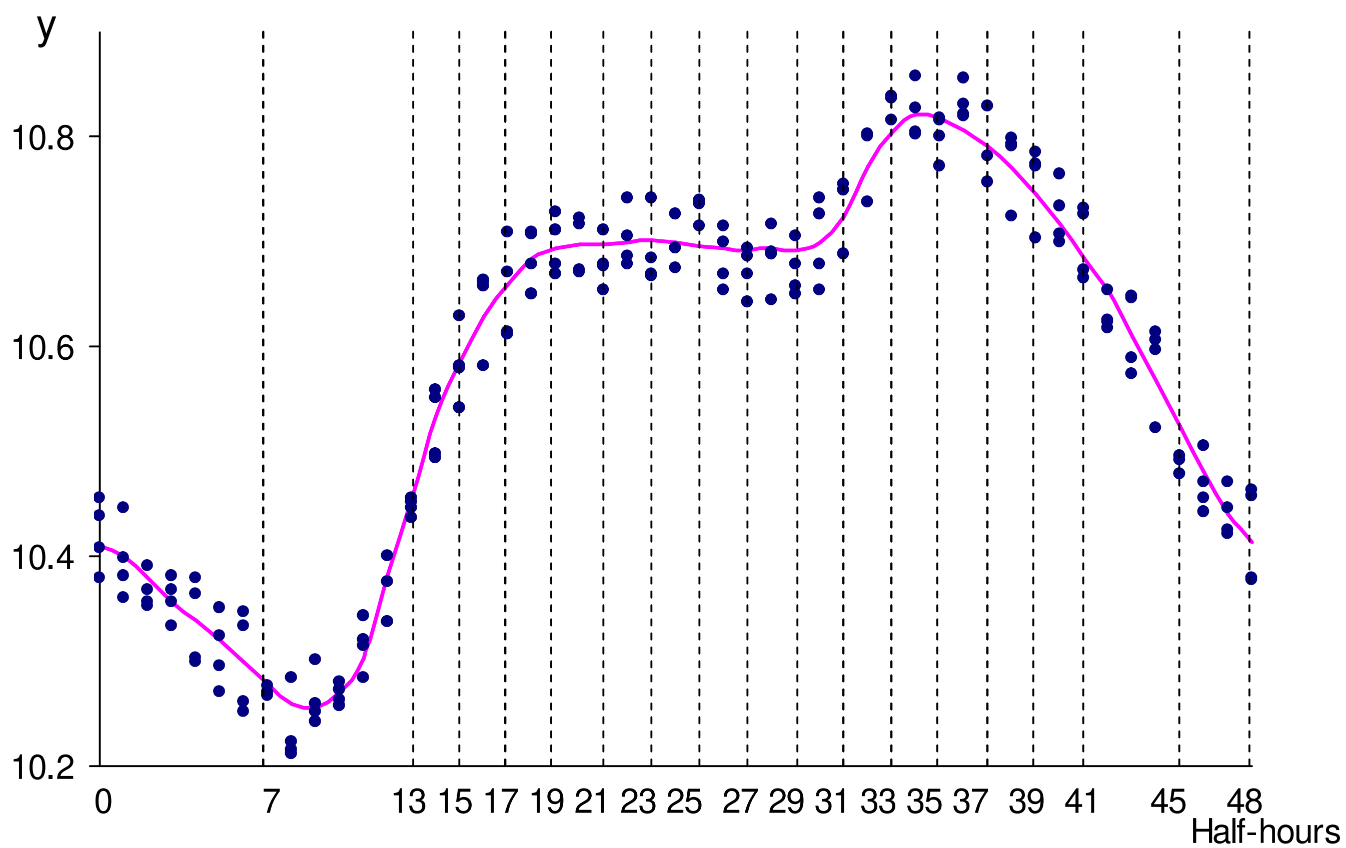
  Any point on spline is linear combination of values at knots x_i^*:

  $$f(x) = \boldsymbol{w}'\boldsymbol{s}^*$$

  $\boldsymbol{w}$ calculated analytically.

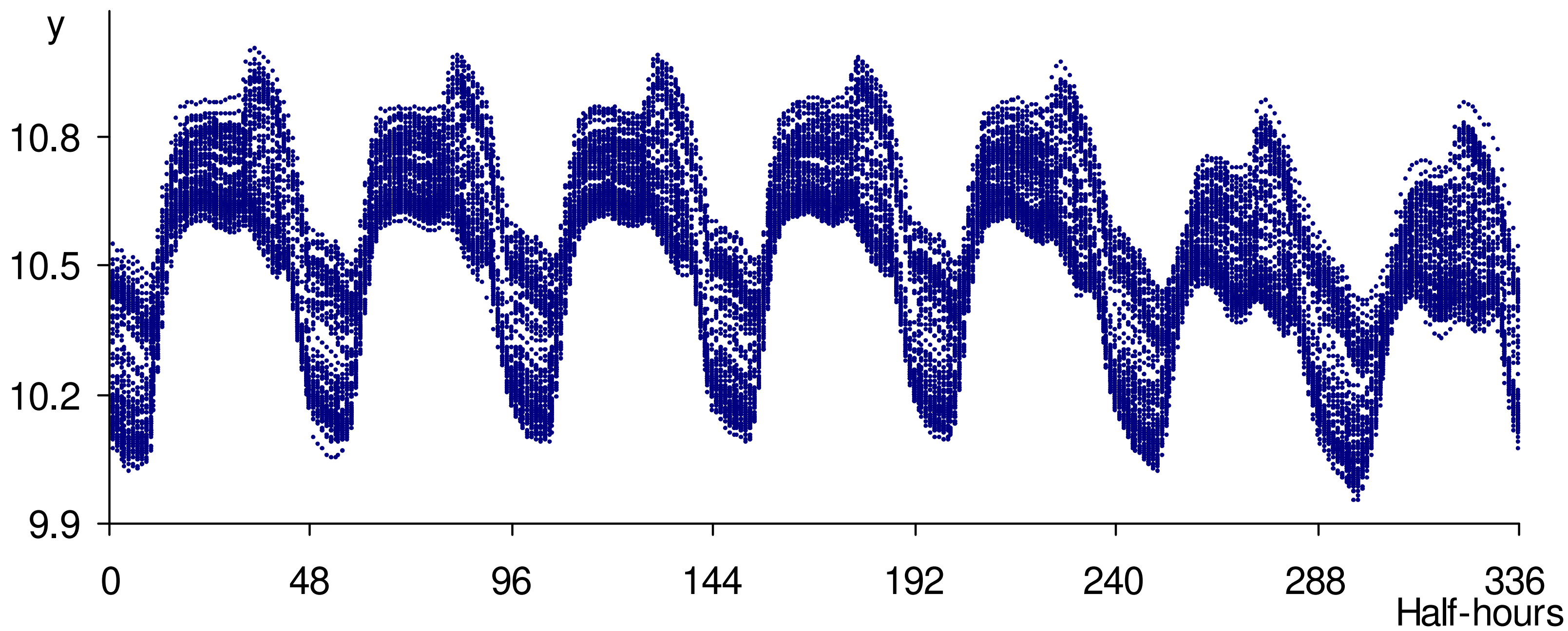
- OLS regression spline:

  $$y_t = \boldsymbol{w}_t'\boldsymbol{s}^* + \varepsilon_t$$



4. DWR Splines

- Use DWR to estimate $\boldsymbol{s}_t^*$ for intraweek cycle
  - select knots
  - constrain spline to be identical at certain knots
  - different λ for night and day knots



5. Spline-Based ES

$$y_t = \boldsymbol{w}_t' \boldsymbol{s}_{t-1}^* + \phi e_{t-1} + \varepsilon_t$$

$$e_t = y_t - \boldsymbol{w}_t' \boldsymbol{s}_{t-1}^*$$

$$s_{j,t}^* = s_{j,t-1}^* + \left(\alpha + \kappa I_{jt}^{knot} + \eta I_{jt}^{nearby}\right) e_t \qquad \left(j = 1, 2, \ldots, M\right)$$

$$I_{it}^{knot} = \begin{cases} 1 & \text{if period } t \text{ is location of knot } i \\ 0 & \text{otherwise} \end{cases}$$

$$I_{it}^{nearby} = \begin{cases} 1 & \text{if period } t \text{ is between knots } (i-1) \text{ and } (i+1) \\ 0 & \text{otherwise} \end{cases}$$

- Similar to Harvey and Koopman 1993.

Methods

- Double seasonal ES
- Intraday cycle ES
- DWR with trigonometric terms
- DWR splines
- Spline-based ES
- SVD-based ES

SVD-Based Forecasting

- Arrange data as $(w \times 336)$ matrix $\boldsymbol{Y}$.

- SVD gives:
  - *intraweek basis vectors* in $\boldsymbol{V}$
  - *weekly feature series* in $\boldsymbol{P}$

$$\underset{(w\times 336)}{\boldsymbol{Y}} = \underset{(w\times 336)}{\boldsymbol{P}}\ \underset{(336\times 336)}{\boldsymbol{V}'}$$

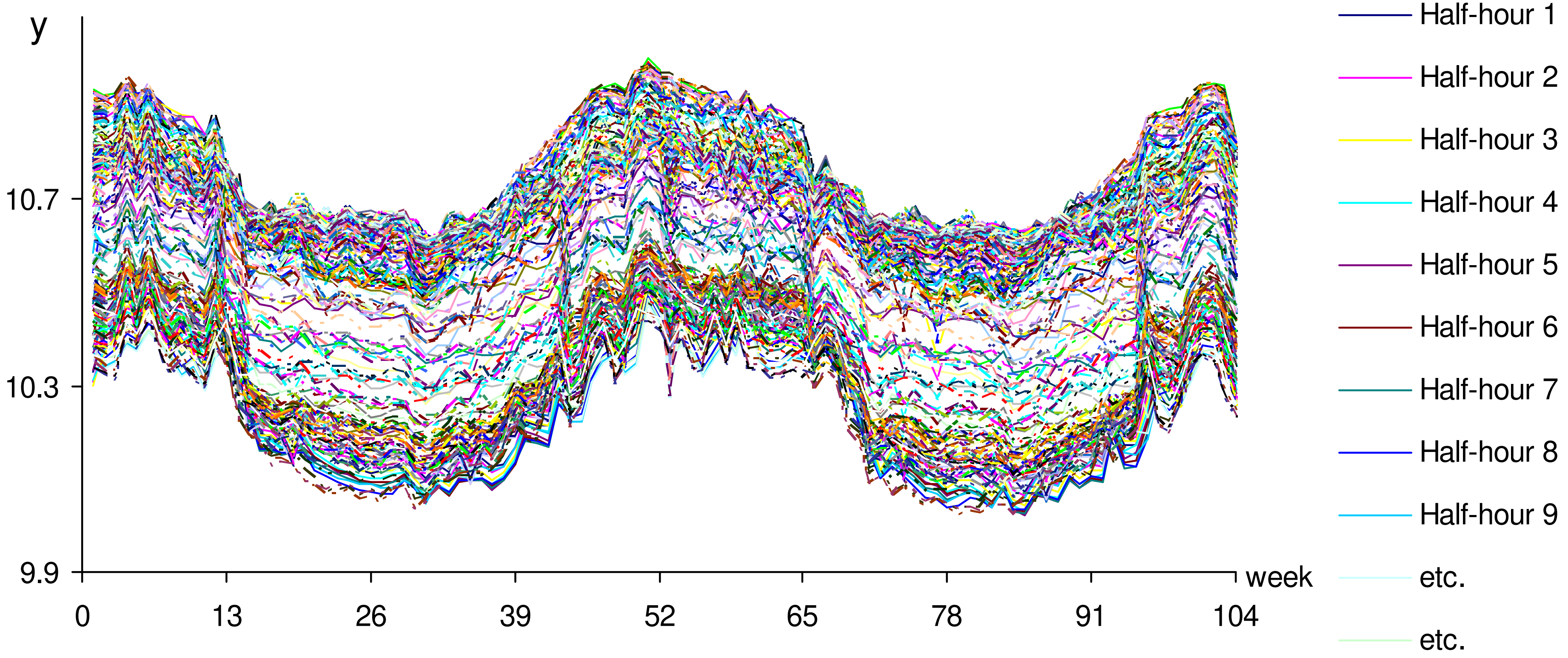
- Select only first k features and bases.

- Forecast each feature and project back onto $\boldsymbol{Y}$ space.

Data Matrix *Y*

Half-hours

	1	2	3	4		335	336
Week 1	10.47	10.46	10.46	10.43	....	10.46	10.43
Week 2	10.44	10.44	10.44	10.42	....	10.47	10.43
Week 3	10.45	10.45	10.45	10.43	....	10.48	10.46
⋮	⋮	⋮	⋮	⋮		⋮	⋮
⋮	⋮	⋮	⋮	⋮		⋮	⋮
Week 104	10.38	10.37	10.36	10.34	....	10.44	10.41



Singular Value Decomposition

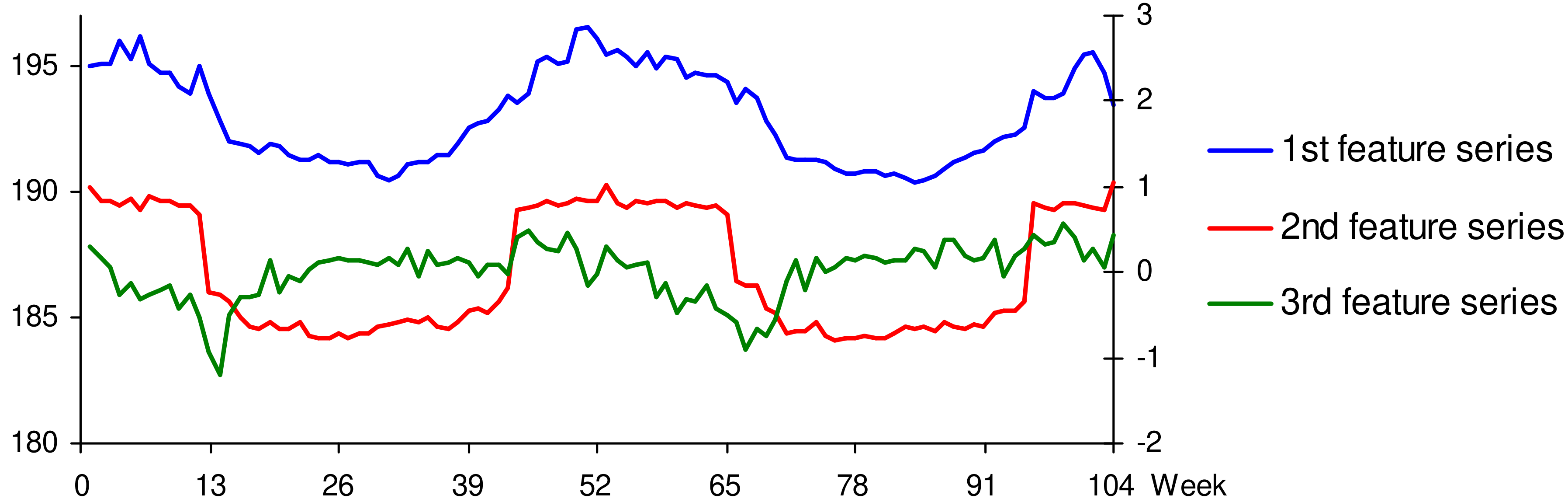
Half-hours

	1	2	3	4		335	336
Week 1	10.47	10.46	10.46	10.43	………	10.46	10.43
Week 2	10.44	10.44	10.44	10.42	………	10.47	10.43
Week 3	10.45	10.45	10.45	10.43	………	10.48	10.46
⋮	⋮	⋮	⋮	⋮		⋮	⋮
⋮	⋮	⋮	⋮	⋮		⋮	⋮
Week 104	10.38	10.37	10.36	10.34	………	10.44	10.41

→

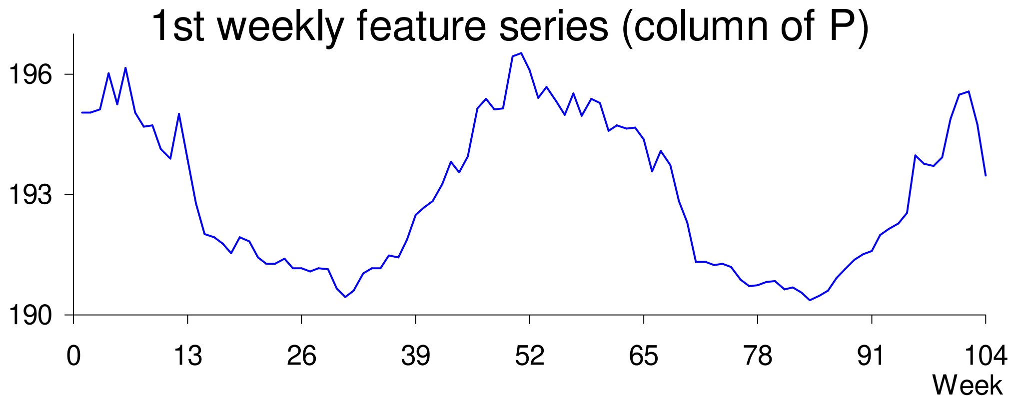
features series

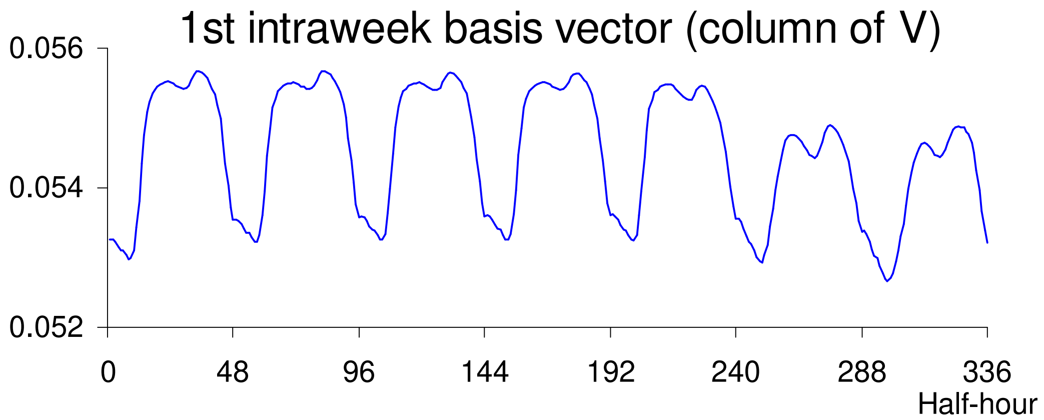
195.0	0.99	0.31	…
195.1	0.83	0.18	…
195.1	0.83	0.06	…
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
193.5	1.04	0.44	…

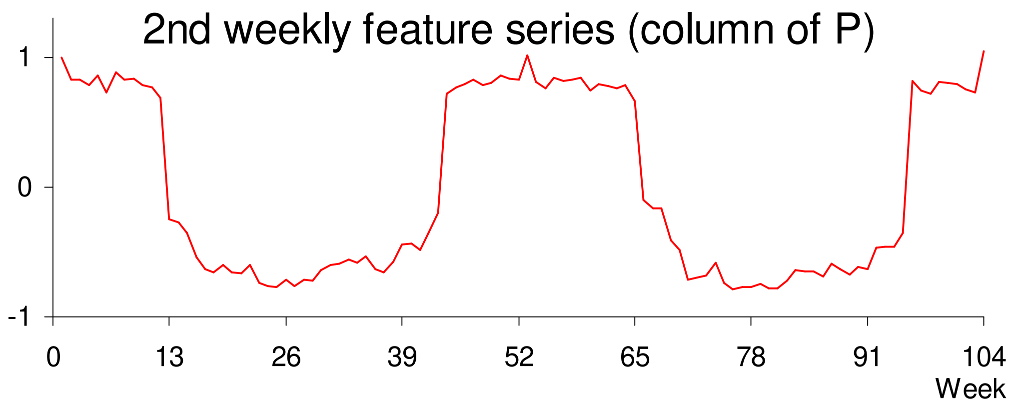


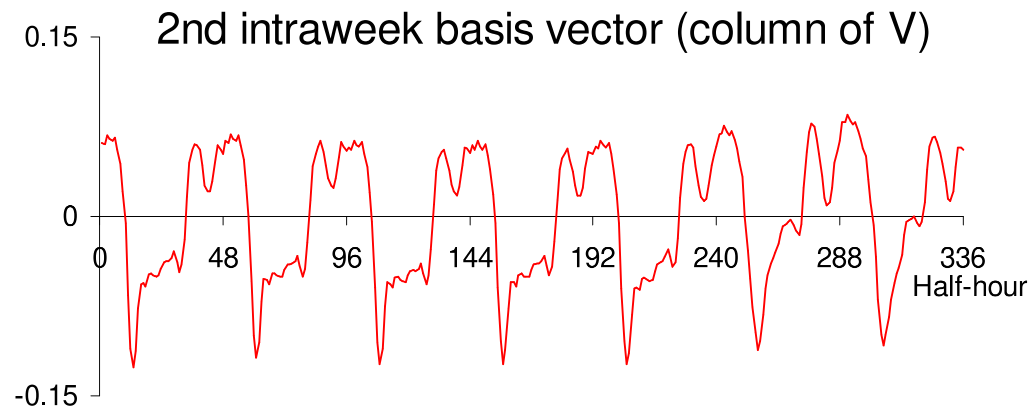
$$\boldsymbol{Y} = \boldsymbol{P}\boldsymbol{V}'$$

$(w \times 336) \quad (w \times 336)\ (336 \times 336)$









6. SVD-Based ES

- Work with just first k features: $$\underset{(w\times 336)}{\boldsymbol{Y}} \approx \underset{(w\times k)}{\tilde{\boldsymbol{P}}}\ \underset{(k\times 336)}{\tilde{\boldsymbol{V}}'}$$

- Need to forecast future week/row of $\tilde{\boldsymbol{P}}$.

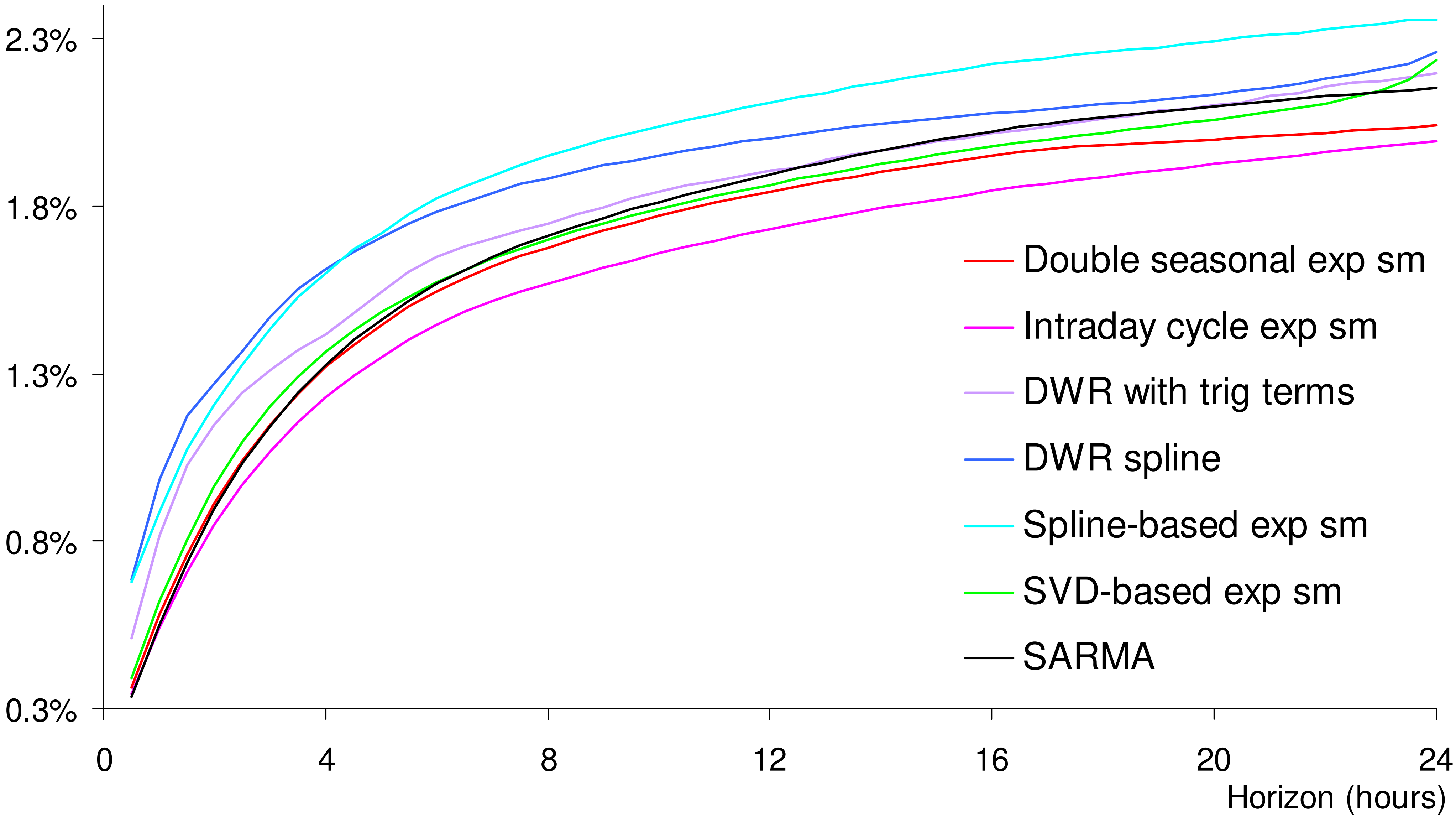
$$y_t = \tilde{\boldsymbol{p}}_{t-1}\tilde{\boldsymbol{V}}_{[t \bmod 336]}{}' + \phi\, e_{t-1} + \varepsilon_t$$

$$e_t = y_t - \tilde{\boldsymbol{p}}_{t-1}\tilde{\boldsymbol{V}}_{[t \bmod 336]}{}'$$

$$\tilde{\boldsymbol{p}}_t = \tilde{\boldsymbol{p}}_{t-1} + \left(\alpha \mathbf{1}_{336}\tilde{\boldsymbol{V}} + \delta\sum_{j=1}^{7}\tilde{\boldsymbol{V}}_{[t \bmod 48]+(j-1)48} + \omega\tilde{\boldsymbol{V}}_{[t \bmod 336]}\right) e_t$$

$\mathbf{1}_{336}$ is (1×336) matrix of 1's.

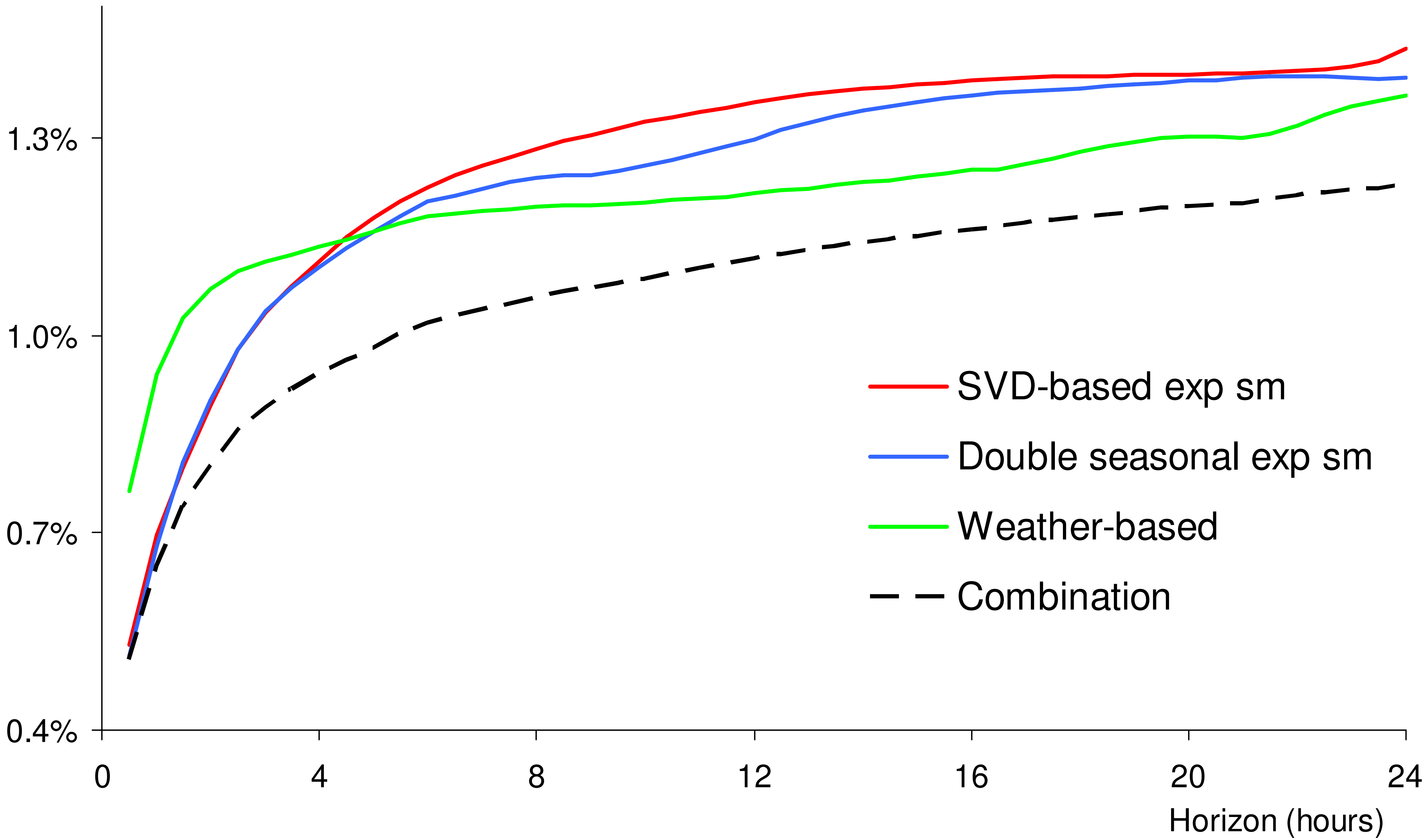
MAPE for France & GB



Broader Comparison

	Conceptual simplicity	Ease of implementation	Judgement required	Point forecast accuracy	Prediction Intervals
Double seasonal ES	3	3	3	3	3
Intraday cycle ES	2	2	2	3	3
DWR with trig. terms	2	1	3	2	1
DWR spline	1	1	1	1	1
Spline-based ES	1	1	1	2	3
SVD-based ES	1	2	2	3	2
SARMA	2	1	1	2	3

MAPE for GB (different study with 10 weeks post-sample)



Other Work

- Triple seasonal models
- Anomalous load
- Minute-by-minute data and very short lead times
- Probabilistic forecasting
- Call centre arrivals