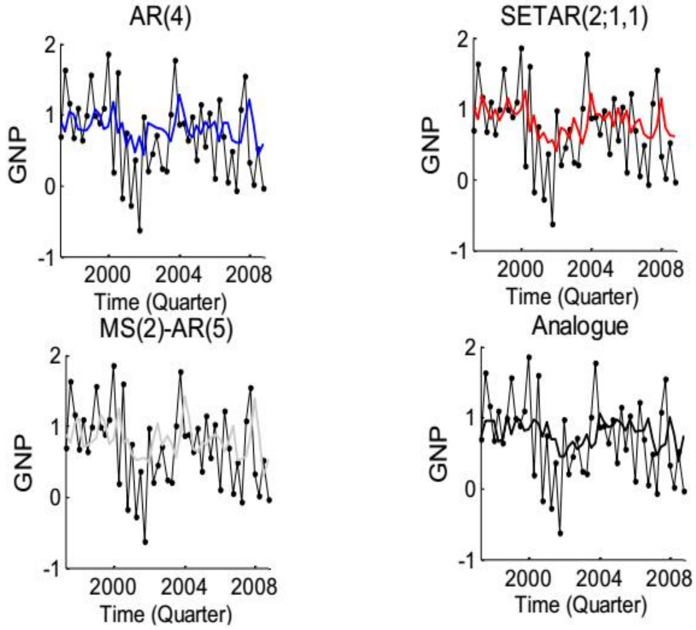
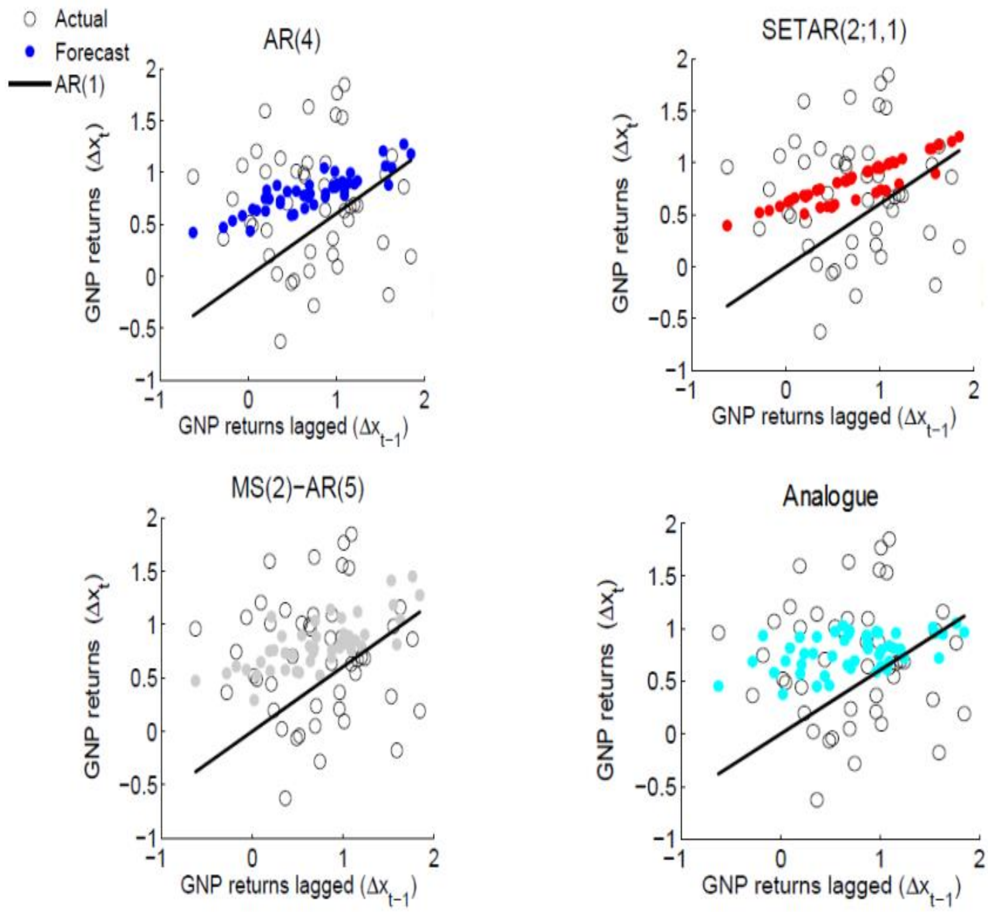
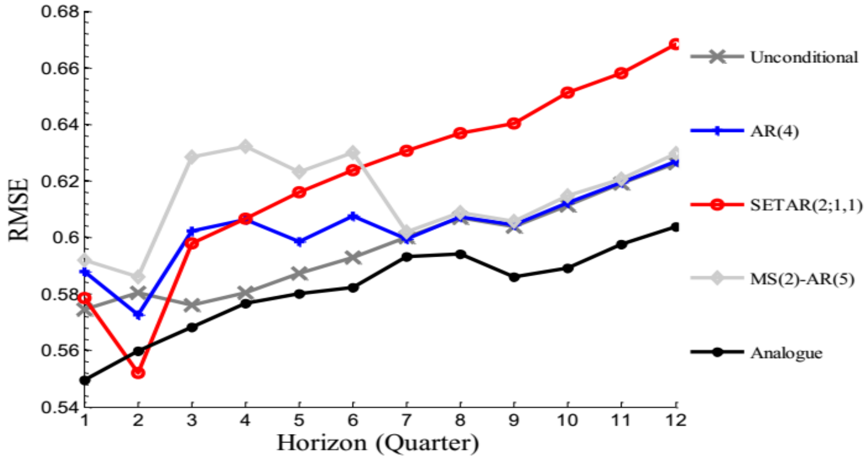


Actual and Forecast: scatter plot



Root mean squared error

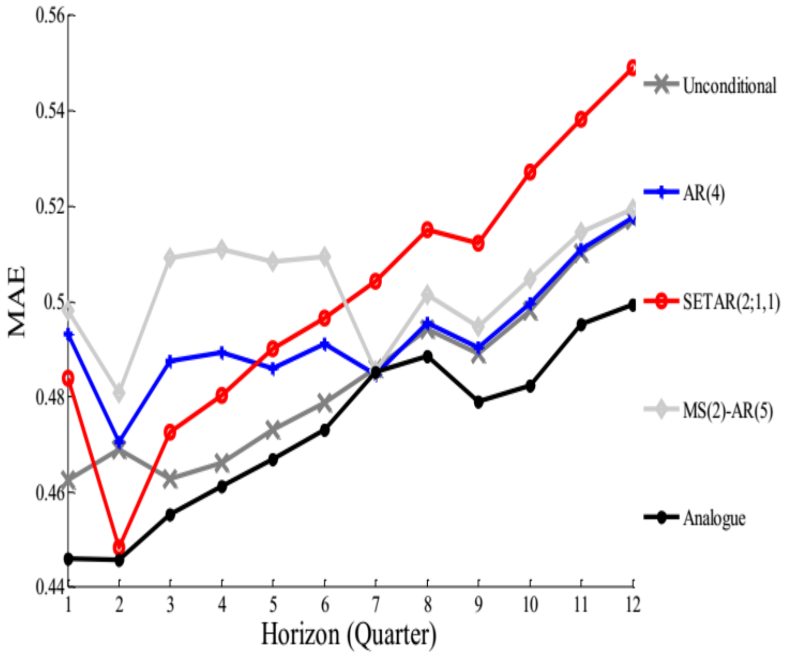
$$RMSE_h = \sqrt{\frac{1}{N-T-h+1}\sum_{i=T+h}^{N}(y_i - \hat{y}_i)^2}$$



Note: lower RMSE corresponds to more accurate point forecasts

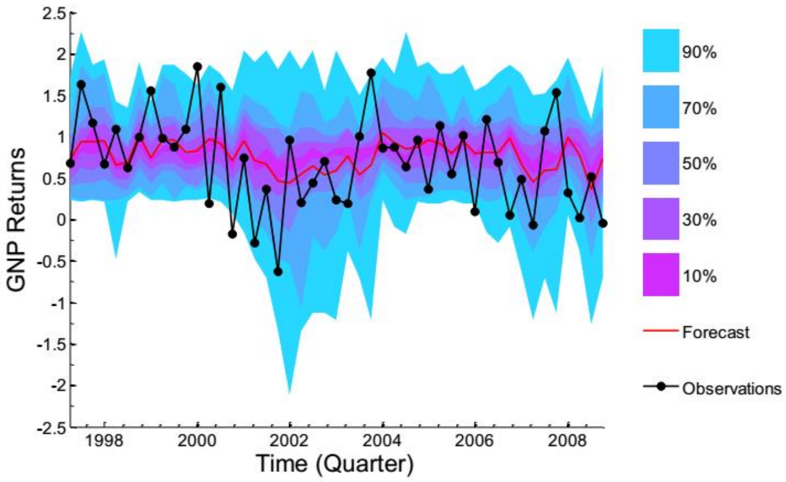
Mean absolute error

$$MAE_h = \frac{1}{N-T-h+1}\sum_{i=T+h}^{N}|y_i - \hat{y}_i|$$

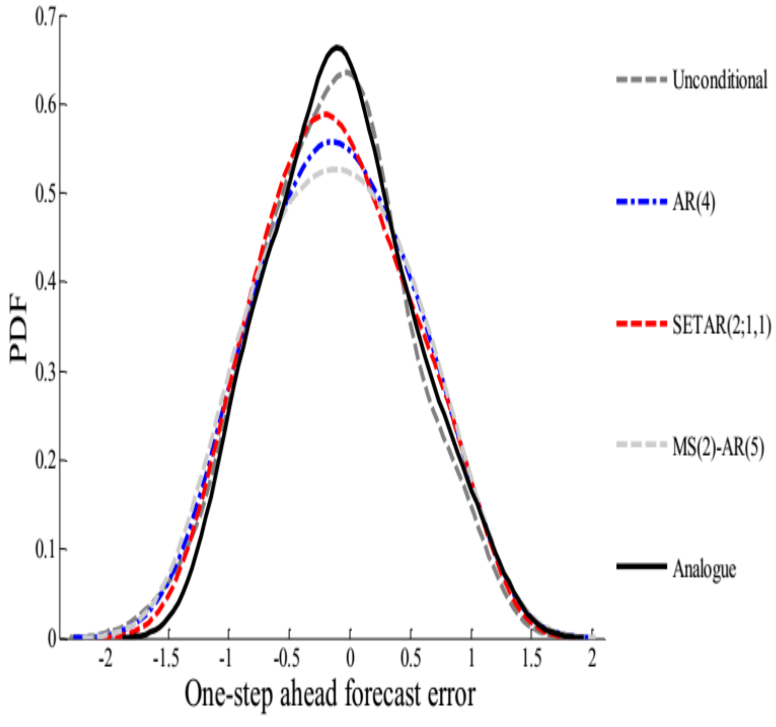


Q: How is your husband? A: Compared to what?

One-step ahead forecast



Error distribution



Multistep ahead forecast using nonlinear models

- Need to use Monte Carlo sampling
- One-step ahead forecast:

$$\hat{y}_{t+1}^{MCi} = (1 - I_t(r))(\alpha_0 + \alpha_1 y_{t-1}) + I_t(r)(\beta_0 + \beta_1 y_{t-1})$$

- Multi-step ahead forecast:

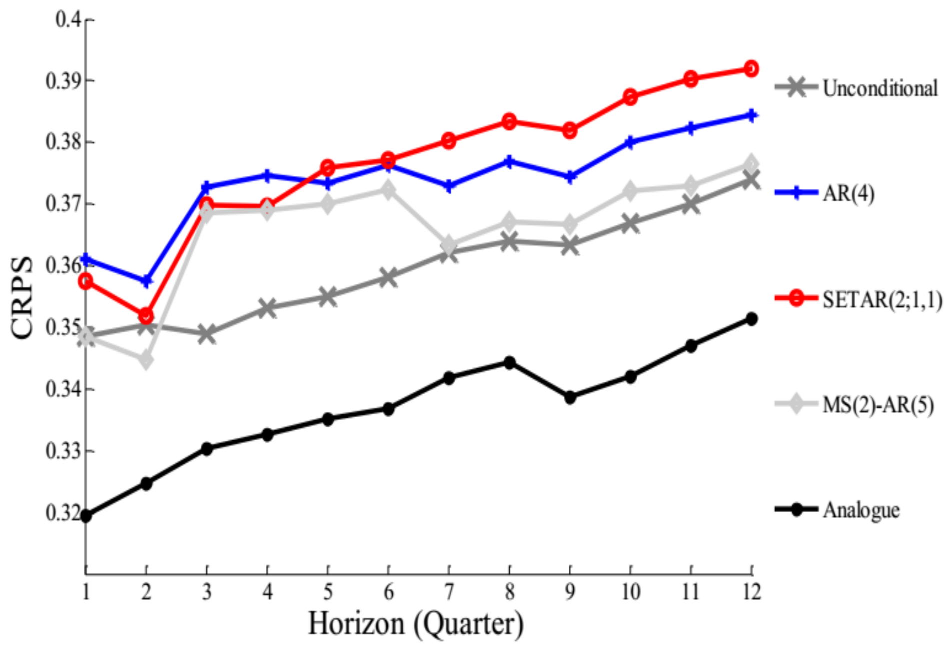
$$\hat{y}_{t+k}^{MCi} = (1 - I_{t+k-1}(r))(\alpha_0 + \alpha_1 \hat{y}_{t+k-1}^{MCi}) + I_{t+k-1}(r)\left(\beta_0 + \beta_1 \hat{y}_{t+1}^{MCi}\right) + \zeta_{k,j}$$
$$I_{t+k-1}(r) = I_{t+k-1}\left(\hat{y}_{t+k-1}^{MCi} > r\right) = 1 \text{ if } \hat{y}_{t+k-1}^{MCi} > r \text{ and 0 otherwise}$$

- Averaging gives the MC point forecast

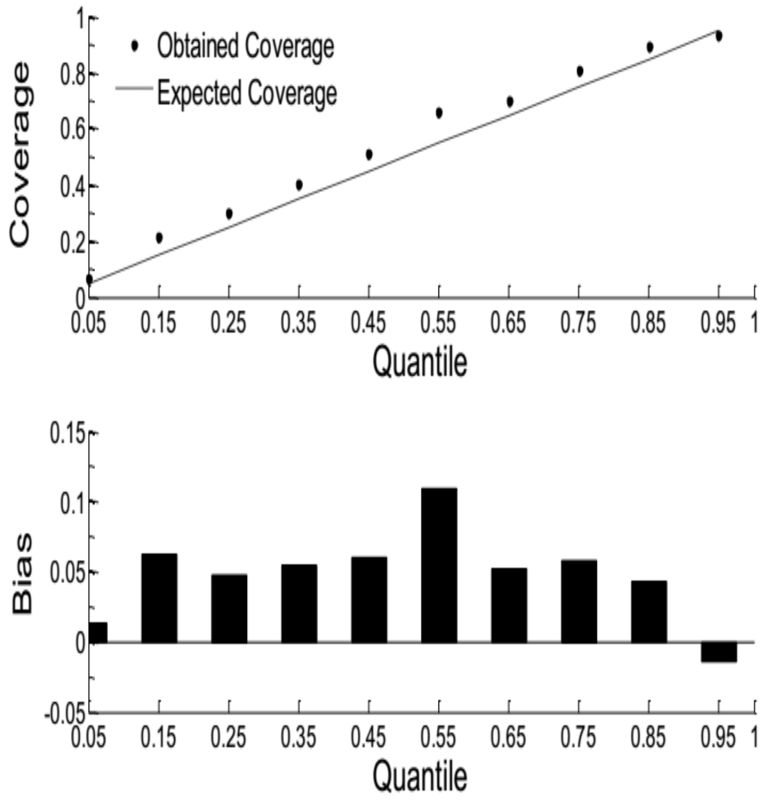
$$\hat{y}_{t+k}^{MC} = \frac{1}{N}\sum_{j=1}^{N} \hat{y}_{t+k}^{MCi}$$

Continuous rank probability score

$$CRPS(F, y) = \int_{z=-\infty}^{z=+\infty} \{F(z) - \mathbf{1}(z \geq y)\}^2 dz$$



Bias and Coverage



Reliability diagram (coverage and bias) for one-step ahead out-of-sample forecasts from 1997Q1 up to 2008Q3 using the analogue prediction model. Note that *Coverage* reflects the degree to which actual observations lie within different quantile ranges of forecast distribution. For a perfect density forecast, the coverage would lie along the diagonal (dashed line), and hence the *Bias* would be zero across all quantiles

144

Worth investigating

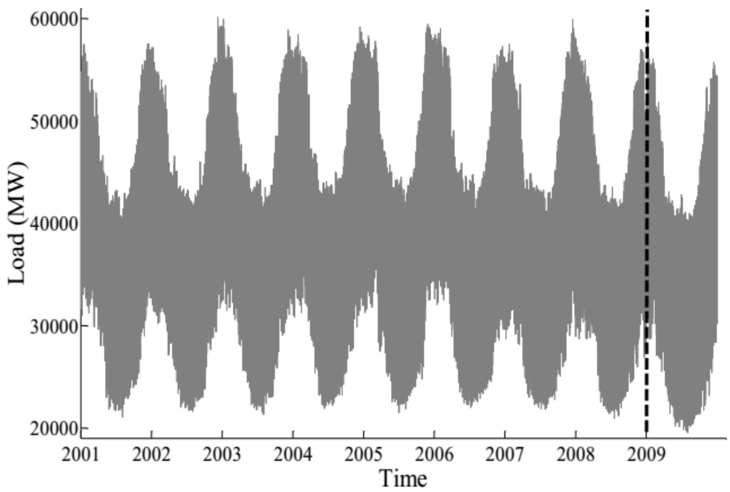
- Comparing forecasts for direct and iterative forecasting schemes
- Re-estimating model parameters with rolling forecast origin (as new information becomes available)
- Combining models based on density forecast performance
- Semi-parametric approaches
- Including additional explanatory variables

Forecasting national electricity demand

- Accurate demand forecasts are important for the safe and reliable operation of power systems, and have huge financial implications for energy markets
- Demand is important for estimating capacity required to avoid black-outs
- Control of generation and distribution
- Ability to make informed decisions
- Reduce risks and minimise costs
- Implementation depends on the forecast horizon:
  – Short-term: ensuring system stability
  – Medium-term: maintenance scheduling
  – Long-term: capital planning
- Required for studying electricity prices
- Design of efficient electricity markets
- Bunn and Farmer estimated that1 % increase in demand forecast error translated into an annual loss of 10 million pounds

Intrayear seasonality

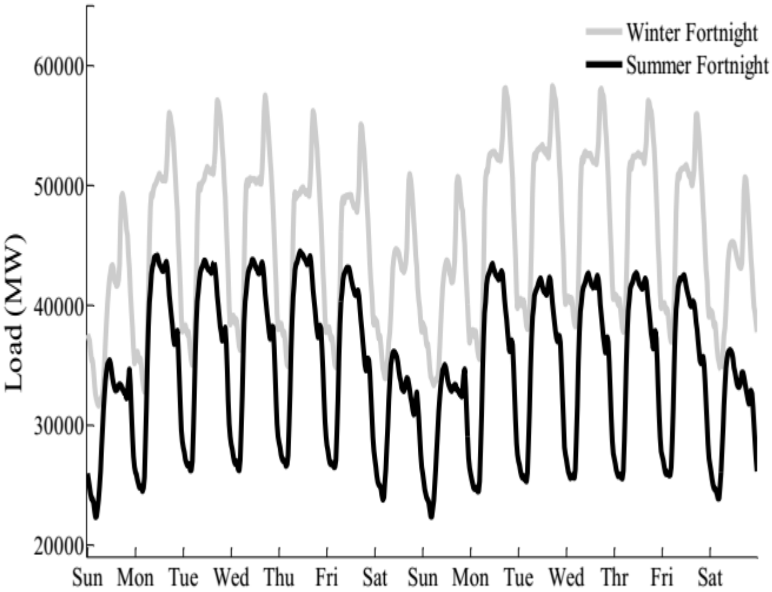
- Electricity demand exhibits strong annual seasonal patterns, whereby demand is higher in winter (due to an increased use of electrical heating equipment in Britain) than in summer



Half-hourly demand (*load*) for Great Britain stretching from 1 January 2001 to 31 December 2009. The vertical dashed line denotes the time index that divides the time series into non-overlapping estimation and evaluation periods

Intraweek seasonality

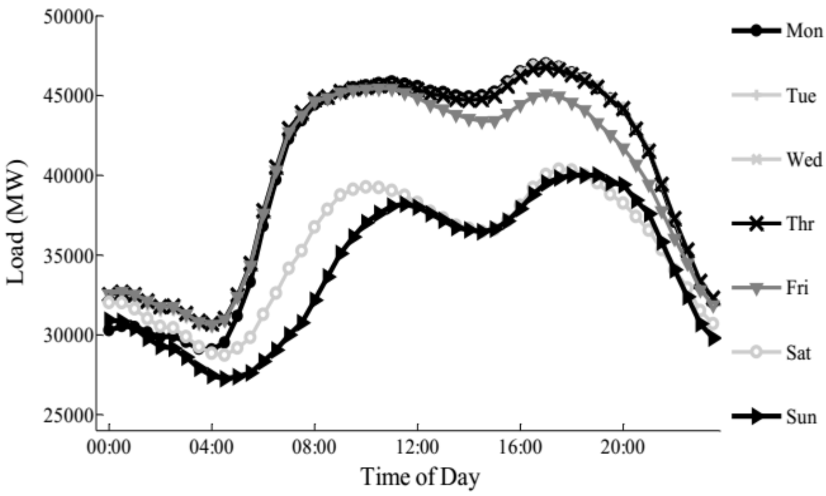
- The weekdays witness a higher demand than the weekends for each fortnight – weekly seasonality



Demand for a winter fortnight (16-29 January, 2005) and a summer fortnight (10-23 July, 2005)

Intraday seasonality

- Demand is consistently higher for certain periods of the day (working hours) than others (night time)



Average intraday profile for each day of the week, computed using only the estimation sample (2001-2008)

Triple seasonal exponential smoothing

- Triple seasonal Holt-Winters-Taylor exponential smoothing (Taylor, 2010)

$$y_t = l_{t-1} + d_{t-m_1} + w_{t-m_2} + \mathrm{a}_{t-m_3} + \phi e_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim \mathrm{NID}(0, \sigma^2)$$

$$e_t = y_t - (l_{t-1} + d_{t-m_1} + w_{t-m_2} + \mathrm{a}_{t-m_3})$$

$l_t = l_{t-1} + \lambda e_t$ → Smoothed Level : (smoothing para. λ)

$d_t = d_{t-m_1} + \delta e_t$ → Seasonal Index : Intraday (smoothing para. δ)

$w_t = w_{t-m_2} + \omega e_t$ → Seasonal Index : Intraweek (smoothing para, ω)

$\mathrm{a}_t = \mathrm{a}_{t-m_3} + \alpha e_t$ → Seasonal Index : Intrayear (smoothing para, α)

$$m_1 = 48\ (2 \times 24), m_2 = 336\ (7 \times m_1), m_3 = 17472\ (52 \times m_2)$$

SARMA

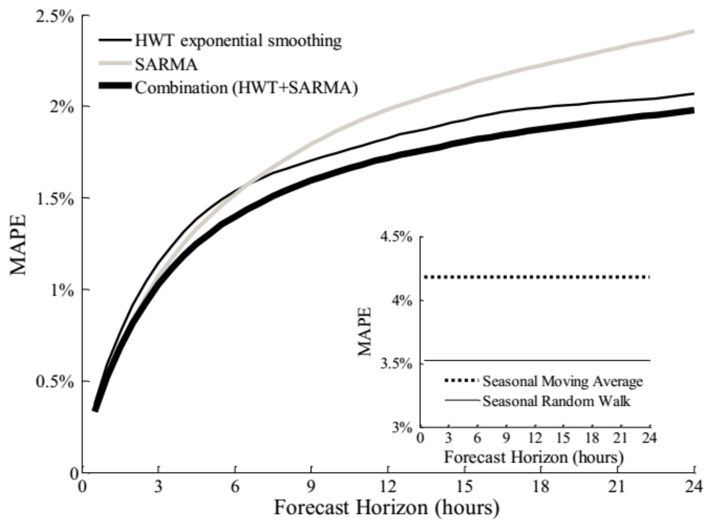
- Triple seasonal SARMA:

$$\Upsilon_p(L)\Phi_{P_1}(L^{m_1})X_{P_2}(L^{m_2})\Psi_{P_3}(L^{m_3})(y_t - c) = \Omega_q(L)\Theta_{Q_1}(L^{m_1})\Gamma_{Q_2}(L^{m_2})\Lambda_{Q_3}(L^{m_3})\varepsilon_t$$

Υ_p, Φ_{P_1}, X_{P_2}, Ψ_{P_3} are polynomial functions for the AR terms having order p, P_1, P_2 and P_3, respectively. The polynomial functions for the MA terms are denoted by Ω_q, Θ_{Q_1}, Γ_{Q_2}, and Λ_{Q_3}, having corresponding model order given by q, Q_1, Q_2 and Q_3

- The model orders can be selected using the AIC

Forecast evaluation



Out-of-sample MAPE for the HWT exponential smoothing (ES), SARMA, and their combination. Benchmarks: Seasonal moving average and seasonal random walk. Forecast horizon ranging from ½ hour to 1 day.

Note: Model combination based on taking the average of two forecasts (ES and SARMA) results in more accurate forecasts!

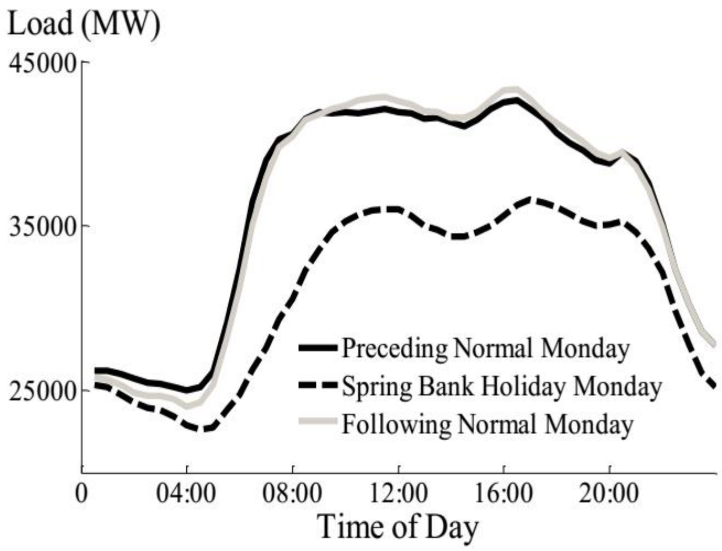
Rule-based forecasting

- Incorporating subjective experience/domain knowledge into the modelling framework
- Useful in cases where time series has a consistent structure and domain knowledge is available
- Rules need to be inferred based on analysis of the time series
- A rule essentially helps identify a historical observation (lag) that will potentially be most useful in enhancing the model's estimate for the future observation.
- For ex: to forecast a future anomalous observation, it may be useful to refer to the corresponding most recently observed anomalous observation (example in following slide)
- Missing observations – fill gap using interpolation, and ignore the interpolated value during model estimation and evaluation
- Irregularly sampled observations – methods like Gaussian processes may be useful

Artificial Neural Networks (ANNs) for Forecasting

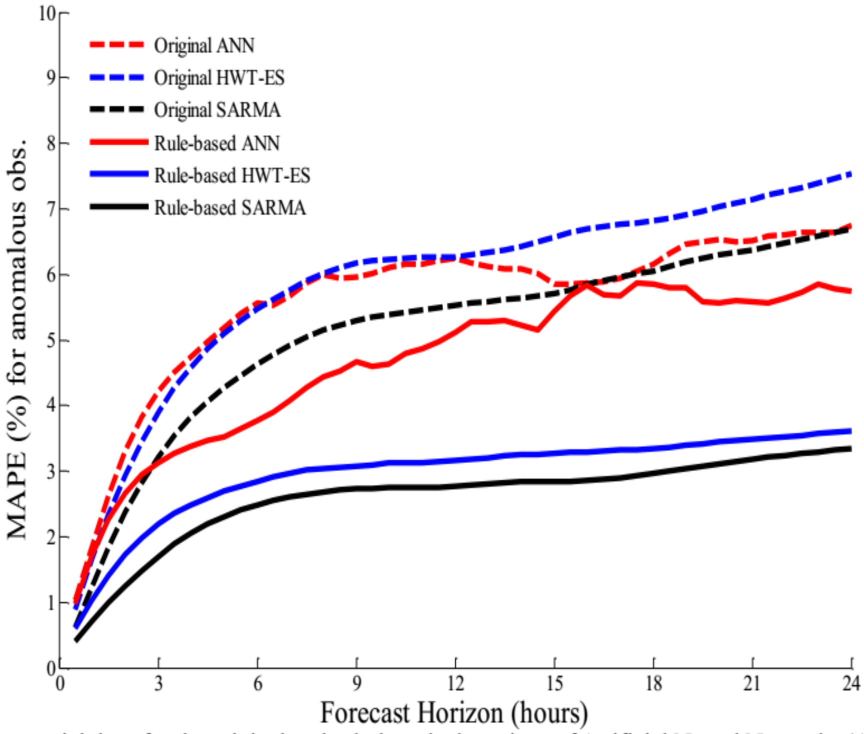
- ANNs:
  - Provide a nonlinear and nonparametric approach for time series modelling
  - Are able to model the complex nonlinear relationships between the dependent variable (for ex. load) and independent variables (for ex. weather variables such as temperature, humidity, cloud cover etc.)
  - No well-established systematic approach, or a consensus among researchers, for choosing a suitable ANN architecture, namely, the number of inputs, hidden layers, and units within each hidden layer, for a given dataset
  - ANNs have been shown to be unsuitable for generating multi-step ahead forecasts (Atiya *et al*., 1999), so it might be more suitable to build a separate ANN model for each forecast horizon
- Other machine learning techniques such as support vector machines and random forests have also been explored for time series forecasting. It is crucial to remember that while generating out-of-sample forecasts, one should use forecasts of the independent variables (preferably ensembles) and not actuals observations.

Modelling anomalous load



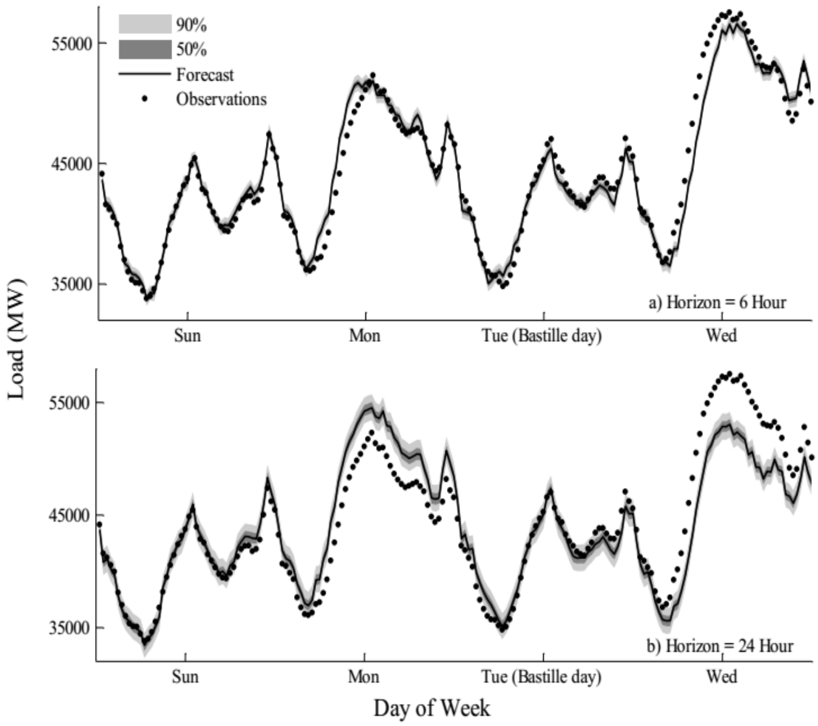
Load profile (GB) for a Spring Bank Holiday Monday (26 May 2008) and a normal Monday from the preceding and following week (in-sample data).

Rule-based forecasting



MAPE across special days for the original and rule-based adaptations of Artificial Neural Networks (ANNs), Holt-Winters-Taylor Exponential Smoothing (HWT-ES) and Seasonal Autoregressive Moving Average (SARMA) models

Rule-based SARMA (French load)



Out-of-sample probabilistic forecasts generated using RB-SARMA, for horizon = 6 hours, and 24 hours. Note that the load profile on weekend (Sun) and special day (Bastille day) is noticeably different from load observed on normal working days.