Forecasting electricity demand distributions using a semiparametric additive model

Rob J Hyndman
Joint work with Shu Fan
The problem in 2007

- We want to forecast the peak electricity demand in a half-hour period in ten years time.
- We have twelve years of half-hourly electricity data, temperature data and some economic and demographic data.
- The location is South Australia: home to the most volatile electricity demand in the world.

Sounds impossible?
We want to forecast the peak electricity demand in a half-hour period in ten years time.

We have twelve years of half-hourly electricity data, temperature data and some economic and demographic data.

The location is South Australia: home to the most volatile electricity demand in the world.
The problem in 2007

- We want to forecast the peak electricity demand in a half-hour period in ten years time.
- We have twelve years of half-hourly electricity data, temperature data and some economic and demographic data.
- The location is South Australia: home to the most volatile electricity demand in the world.

Sounds impossible?
We want to forecast the peak electricity demand in a half-hour period in ten years time.

We have twelve years of half-hourly electricity data, temperature data and some economic and demographic data.

The location is South Australia: home to the most volatile electricity demand in the world.

Sounds impossible?
The problem in 2007

We want to forecast the peak electricity demand in a half-hour period in ten years time.

We have twelve years of half-hourly electricity data, temperature data and some economic and demographic data.

The location is South Australia: home to the most volatile electricity demand in the world.

Sounds impossible?
South Australian demand data

South Australia state-wide demand (winters only)
South Australian demand data

South Australia state wide demand (winter 09/10)

Forecasting electricity demand distributions

The problem
South Australian demand data

South Australia state wide demand (summers only)
South Australian demand data

South Australia state wide demand (summers only)

Black Saturday →
South Australian demand data

South Australia state wide demand (January 2011)

Date in January
South Australian demand (GW)
1.5 2.0 2.5 3.0 3.5
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31
Demand boxplots (Sth Aust)

The problem

Time: 12 midnight

Forecasting electricity demand distributions

Day of week

[Graph showing boxplots for different days of the week, with a y-axis labeled 'Demand (GW)' ranging from 1.0 to 3.5]
Forecasting electricity demand distributions

The problem
Density of demand: 12 midnight

South Australian half-hourly demand (GW)
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework

Forecasting electricity demand distributions  The model
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework

Forecasting electricity demand distributions
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework

- Semi-parametric additive models with correlated errors.
- Each half-hour period modelled separately for each season.
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework

- Semi-parametric additive models with correlated errors.
- Each half-hour period modelled separately for each season.
- Variables selected to provide best out-of-sample predictions using cross-validation on each summer.
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework

- Semi-parametric additive models with correlated errors.
- Each half-hour period modelled separately for each season.
- Variables selected to provide best out-of-sample predictions using cross-validation on each summer.
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework

- **Semi-parametric additive models** with correlated errors.
- Each half-hour period modelled separately for each season.
- Variables selected to provide best out-of-sample predictions using cross-validation on each summer.
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework

- **Semi-parametric additive models** with correlated errors.
- Each half-hour period modelled separately for each season.
- Variables selected to provide best out-of-sample predictions using cross-validation on each summer.
Predictors

- calendar effects
- prevailing and recent weather conditions
- climate changes
- economic and demographic changes
- changing technology

Modelling framework

- **Semi-parametric additive models** with correlated errors.
- Each half-hour period modelled separately for each season.
- Variables selected to provide best out-of-sample predictions using cross-validation on each summer.
\[ \log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- \( y_t \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) and \( p \) denotes the time of day \( p = 1, \ldots, 48 \);
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \);
- \( n_t \) denotes the model error at time \( t \).
\[ \log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- \( y_t \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) and \( p \) denotes the time of day \( p = 1, \ldots, 48; \)
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \)
- \( n_t \) denotes the model error at time \( t \).
\[ \log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- \(y_t\) denotes per capita demand (minus offset) at time \(t\) (measured in half-hourly intervals) and \(p\) denotes the time of day \(p = 1, \ldots, 48\);
- \(h_p(t)\) models all calendar effects;
- \(f_p(w_{1,t}, w_{2,t})\) models all temperature effects where \(w_{1,t}\) is a vector of recent temperatures at location 1 and \(w_{2,t}\) is a vector of recent temperatures at location 2;
- \(z_{j,t}\) is a demographic or economic variable at time \(t\);
- \(n_t\) denotes the model error at time \(t\).
\[ \log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- \( y_t \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) and \( p \) denotes the time of day \( p = 1, \ldots, 48 \);
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \);
- \( n_t \) denotes the model error at time \( t \).
\[
\log(y_t) = h_p(t) + f_p(w_1,t, w_2,t) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

- \( y_t \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) and \( p \) denotes the time of day \( p = 1, \ldots, 48; \)
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_1,t, w_2,t) \) models all temperature effects where \( w_1,t \) is a vector of recent temperatures at location 1 and \( w_2,t \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \)
- \( n_t \) denotes the model error at time \( t \).
\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\(h_p(t)\) includes handle annual, weekly and daily seasonal patterns as well as public holidays:

\[
h_p(t) = \ell_p(t) + \alpha_{t,p} + \beta_{t,p} + \gamma_{t,p} + \delta_{t,p}
\]

- \(\ell_p(t)\) is “time of summer” effect (a regression spline);
- \(\alpha_{t,p}\) is day of week effect;
- \(\beta_{t,p}\) is “holiday” effect;
- \(\gamma_{t,p}\) New Year’s Eve effect;
- \(\delta_{t,p}\) is millennium effect;
\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\(h_p(t)\) includes handle annual, weekly and daily seasonal patterns as well as public holidays:

\[
h_p(t) = \ell_p(t) + \alpha_{t,p} + \beta_{t,p} + \gamma_{t,p} + \delta_{t,p}
\]

- \(\ell_p(t)\) is “time of summer” effect (a regression spline);
- \(\alpha_{t,p}\) is day of week effect;
- \(\beta_{t,p}\) is “holiday” effect;
- \(\gamma_{t,p}\) New Year’s Eve effect;
- \(\delta_{t,p}\) is millennium effect;
Monash Electricity Forecasting Model

\[ \log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

\( h_p(t) \) includes handle annual, weekly and daily seasonal patterns as well as public holidays:

\[ h_p(t) = \ell_p(t) + \alpha_{t,p} + \beta_{t,p} + \gamma_{t,p} + \delta_{t,p} \]

- \( \ell_p(t) \) is “time of summer” effect (a regression spline);
- \( \alpha_{t,p} \) is day of week effect;
- \( \beta_{t,p} \) is “holiday” effect;
- \( \gamma_{t,p} \) New Year’s Eve effect;
- \( \delta_{t,p} \) is millennium effect;
\[ \log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

\( h_p(t) \) includes handle annual, weekly and daily seasonal patterns as well as public holidays:

\[ h_p(t) = \ell_p(t) + \alpha_{t,p} + \beta_{t,p} + \gamma_{t,p} + \delta_{t,p} \]

- \( \ell_p(t) \) is “time of summer” effect (a regression spline);
- \( \alpha_{t,p} \) is day of week effect;
- \( \beta_{t,p} \) is “holiday” effect;
- \( \gamma_{t,p} \) New Year’s Eve effect;
- \( \delta_{t,p} \) is millennium effect;
Monash Electricity Forecasting Model

\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\(h_p(t)\) includes handle annual, weekly and daily seasonal patterns as well as public holidays:

\[
h_p(t) = \ell_p(t) + \alpha_{t,p} + \beta_{t,p} + \gamma_{t,p} + \delta_{t,p}
\]

- \(\ell_p(t)\) is “time of summer” effect (a regression spline);
- \(\alpha_{t,p}\) is day of week effect;
- \(\beta_{t,p}\) is “holiday” effect;
- \(\gamma_{t,p}\) New Year’s Eve effect;
- \(\delta_{t,p}\) is millennium effect;
Monash Electricity Forecasting Model

\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\(h_p(t)\) includes handle annual, weekly and daily seasonal patterns as well as public holidays:

\[
h_p(t) = \ell_p(t) + \alpha_{t,p} + \beta_{t,p} + \gamma_{t,p} + \delta_{t,p}
\]

- \(\ell_p(t)\) is “time of summer” effect (a regression spline);
- \(\alpha_{t,p}\) is day of week effect;
- \(\beta_{t,p}\) is “holiday” effect;
- \(\gamma_{t,p}\) New Year’s Eve effect;
- \(\delta_{t,p}\) is millennium effect;
Forecasting electricity demand distributions

The model

Time: 3:00 pm

Day of summer

Day of week

Holiday

Effect on demand

Normal  Day before  Holiday  Day after
Monash Electricity Forecasting Model

\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\[
f_p(w_{1,t}, w_{2,t}) = \sum_{k=0}^{6} \left[ f_{k,p}(x_{t-k}) + g_{k,p}(d_{t-k}) \right] + q_p(x_t^+) + r_p(x_t^-) + s_p(\bar{x}_t)
\]

\[
+ \sum_{j=1}^{6} \left[ F_{j,p}(x_{t-48j}) + G_{j,p}(d_{t-48j}) \right]
\]

- \(x_t\) is ave temp across two sites (Kent Town and Adelaide Airport) at time \(t\);
- \(d_t\) is the temp difference between two sites at time \(t\);
- \(x_t^+\) is max of \(x_t\) values in past 24 hours;
- \(x_t^-\) is min of \(x_t\) values in past 24 hours;
- \(\bar{x}_t\) is ave temp in past seven days.

Each function is smooth & estimated using regression splines.
\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\[
f_p(w_{1,t}, w_{2,t}) = \sum_{k=0}^{6} \left[ f_{k,p}(x_{t-k}) + g_{k,p}(d_{t-k}) \right] + q_p(x_t^+) + r_p(x_t^-) + s_p(\bar{x}_t)
\]

\[
+ \sum_{j=1}^{6} \left[ F_{j,p}(x_{t-48j}) + G_{j,p}(d_{t-48j}) \right]
\]

- \(x_t\) is ave temp across two sites (Kent Town and Adelaide Airport) at time \(t\);
- \(d_t\) is the temp difference between two sites at time \(t\);
- \(x_t^+\) is max of \(x_t\) values in past 24 hours;
- \(x_t^-\) is min of \(x_t\) values in past 24 hours;
- \(\bar{x}_t\) is ave temp in past seven days.

Each function is smooth & estimated using regression splines.
Monash Electricity Forecasting Model

\[ \log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

\[ f_p(w_{1,t}, w_{2,t}) = \sum_{k=0}^{6} \left[ f_{k,p}(x_{t-k}) + g_{k,p}(d_{t-k}) \right] + q_p(x_t^+) + r_p(x_t^-) + s_p(\bar{x}_t) \]

\[ + \sum_{j=1}^{6} \left[ F_{j,p}(x_{t-48j}) + G_{j,p}(d_{t-48j}) \right] \]

- \( x_t \) is ave temp across two sites (Kent Town and Adelaide Airport) at time \( t \);
- \( d_t \) is the temp difference between two sites at time \( t \);
- \( x_t^+ \) is max of \( x_t \) values in past 24 hours;
- \( x_t^- \) is min of \( x_t \) values in past 24 hours;
- \( \bar{x}_t \) is ave temp in past seven days.

Each function is smooth & estimated using regression splines.
Monash Electricity Forecasting Model

\[ \log(y_t) = h_p(t) + f_p(w_1,t, w_2,t) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

\[ f_p(w_1,t, w_2,t) = \sum_{k=0}^{6} \left[ f_{k,p}(x_{t-k}) + g_{k,p}(d_{t-k}) \right] + q_p(x_t^+) + r_p(x_t^-) + s_p(\bar{x}_t) \]

\[ + \sum_{j=1}^{6} \left[ F_{j,p}(x_{t-48j}) + G_{j,p}(d_{t-48j}) \right] \]

- \( x_t \) is ave temp across two sites (Kent Town and Adelaide Airport) at time \( t \);
- \( d_t \) is the temp difference between two sites at time \( t \);
- \( x_t^+ \) is max of \( x_t \) values in past 24 hours;
- \( x_t^- \) is min of \( x_t \) values in past 24 hours;
- \( \bar{x}_t \) is ave temp in past seven days.

Each function is smooth & estimated using regression splines.

Forecasting electricity demand distributions

The model
\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\[
f_p(w_{1,t}, w_{2,t}) = \sum_{k=0}^{6} \left[ f_{k,p}(x_{t-k}) + g_{k,p}(d_{t-k}) \right] + q_p(x_t^+) + r_p(x_t^-) + s_p(\bar{x}_t) + \sum_{j=1}^{6} \left[ F_{j,p}(x_{t-48j}) + G_{j,p}(d_{t-48j}) \right]
\]

- \(x_t\) is ave temp across two sites (Kent Town and Adelaide Airport) at time \(t\);
- \(d_t\) is the temp difference between two sites at time \(t\);
- \(x_t^+\) is max of \(x_t\) values in past 24 hours;
- \(x_t^-\) is min of \(x_t\) values in past 24 hours;
- \(\bar{x}_t\) is ave temp in past seven days.

Each function is smooth & estimated using regression splines.
\[ \log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j Z_{j,t} + n_t \]

\[ f_p(w_{1,t}, w_{2,t}) = \sum_{k=0}^{6} \left[ f_{k,p}(x_{t-k}) + g_{k,p}(d_{t-k}) \right] + q_p(x_t^+) + r_p(x_t^-) + s_p(\bar{x}_t) \]

\[ + \sum_{j=1}^{6} \left[ F_{j,p}(x_{t-48j}) + G_{j,p}(d_{t-48j}) \right] \]

- \( x_t \) is ave temp across two sites (Kent Town and Adelaide Airport) at time \( t \);
- \( d_t \) is the temp difference between two sites at time \( t \);
- \( x_t^+ \) is max of \( x_t \) values in past 24 hours;
- \( x_t^- \) is min of \( x_t \) values in past 24 hours;
- \( \bar{x}_t \) is ave temp in past seven days.

Each function is smooth & estimated using regression splines.
\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\[
f_p(w_{1,t}, w_{2,t}) = \sum_{k=0}^{6} \left[ f_{k,p}(x_{t-k}) + g_{k,p}(d_{t-k}) \right] + q_p(x_t^+) + r_p(x_t^-) + s_p(\bar{x}_t) \\
+ \sum_{j=1}^{6} \left[ F_{j,p}(x_{t-48j}) + G_{j,p}(d_{t-48j}) \right]
\]

- \(x_t\) is ave temp across two sites (Kent Town and Adelaide Airport) at time \(t\);
- \(d_t\) is the temp difference between two sites at time \(t\);
- \(x_t^+\) is max of \(x_t\) values in past 24 hours;
- \(x_t^-\) is min of \(x_t\) values in past 24 hours;
- \(\bar{x}_t\) is ave temp in past seven days.

Each function is smooth & estimated using regression splines.
Fitted results (Summer 3pm)

Forecasting electricity demand distributions

The model
log\(y_t\) = \(h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t\)

- Other variables described by linear relationships with coefficients \(c_1, \ldots, c_J\).
- Estimation based on annual data.
Monash Electricity Forecasting Model

\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

- Other variables described by linear relationships with coefficients \(c_1, \ldots, c_J\).
- Estimation based on annual data.
Split model

\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\[
\log(y_t) = \log(y^*_t) + \log(\bar{y}_i)
\]

\[
\log(y^*_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + e_t
\]

\[
\log(\bar{y}_i) = \sum_{j=1}^{J} c_j z_{j,i} + \varepsilon_i
\]

- \(\bar{y}_i\) is the average demand for year \(i\) where \(t\) is in year \(i\).
- \(y^*_t\) is the standardized demand for time \(t\).
Split model

South Australia state wide demand (summers only)
Split model

South Australia state wide demand (summers only)

Adjusted demand (GW)

96/97  98/99  00/01  02/03  04/05  06/07  08/09  10/11
Annual model

\[ \log(\bar{y}_i) = \sum_j c_jz_{j,i} + \varepsilon_i \]

\[ \log(\bar{y}_i) - \log(\bar{y}_{i-1}) = \sum_j c_j(z_{j,i} - z_{j,i-1}) + \varepsilon_i^* \]

- First differences modelled to avoid non-stationary variables.
- Predictors: Per-capita GSP, Price, Summer CDD, Winter HDD.
Annual model

\[
\log(\bar{y}_i) = \sum_j c_j z_{j,i} + \varepsilon_i
\]

\[
\log(\bar{y}_i) - \log(\bar{y}_{i-1}) = \sum_j c_j (z_{j,i} - z_{j,i-1}) + \varepsilon^*_i
\]

- First differences modelled to avoid non-stationary variables.
- Predictors: Per-capita GSP, Price, Summer CDD, Winter HDD.
Annual model

\[
\log(\bar{y}_i) = \sum_j c_j z_{j,i} + \varepsilon_i
\]

\[
\log(\bar{y}_i) - \log(\bar{y}_{i-1}) = \sum_j c_j (z_{j,i} - z_{j,i-1}) + \varepsilon_i^*
\]

- First differences modelled to avoid non-stationary variables.
- Predictors: Per-capita GSP, Price, Summer CDD, Winter HDD.

\[
z_{\text{CDD}} = \sum_{\text{summer}} \max(0, \bar{T} - 18.5)
\]

\[\bar{T} = \text{daily mean}\]
Annual model

\[ \log(\bar{y}_i) = \sum_j c_j Z_{j,i} + \varepsilon_i \]

\[ \log(\bar{y}_i) - \log(\bar{y}_{i-1}) = \sum_j c_j (Z_{j,i} - Z_{j,i-1}) + \varepsilon_i^* \]

- First differences modelled to avoid non-stationary variables.
- Predictors: Per-capita GSP, Price, Summer CDD, Winter HDD.

\[ Z_{\text{HDD}} = \sum_{\text{winter}} \max(0, 18.5 - \bar{T}) \]

\[ \bar{T} = \text{daily mean} \]
Annual model

Cooling and Heating Degree Days

Forecasting electricity demand distributions

The model
### Annual model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta gsp.pc$</td>
<td>$2.02 \times 10^{-6}$</td>
<td>$5.05 \times 10^{-6}$</td>
<td>0.38</td>
<td>0.711</td>
</tr>
<tr>
<td>$\Delta price$</td>
<td>$-1.67 \times 10^{-8}$</td>
<td>$6.76 \times 10^{-9}$</td>
<td>$-2.46$</td>
<td>0.026</td>
</tr>
<tr>
<td>$\Delta scdd$</td>
<td>$1.11 \times 10^{-10}$</td>
<td>$2.48 \times 10^{-11}$</td>
<td>4.49</td>
<td>0.000</td>
</tr>
<tr>
<td>$\Delta whdd$</td>
<td>$2.07 \times 10^{-11}$</td>
<td>$3.28 \times 10^{-11}$</td>
<td>0.63</td>
<td>0.537</td>
</tr>
</tbody>
</table>

- GSP needed to stay in the model to allow scenario forecasting.
- All other variables led to improved $\text{AIC}_C$. 

Forecasting electricity demand distributions
### Annual model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δgsp.pc</td>
<td>$2.02 \times 10^{-6}$</td>
<td>$5.05 \times 10^{-6}$</td>
<td>0.38</td>
<td>0.711</td>
</tr>
<tr>
<td>Δprice</td>
<td>$-1.67 \times 10^{-8}$</td>
<td>$6.76 \times 10^{-9}$</td>
<td>-2.46</td>
<td>0.026</td>
</tr>
<tr>
<td>Δscdd</td>
<td>$1.11 \times 10^{-10}$</td>
<td>$2.48 \times 10^{-11}$</td>
<td>4.49</td>
<td>0.000</td>
</tr>
<tr>
<td>Δwhdd</td>
<td>$2.07 \times 10^{-11}$</td>
<td>$3.28 \times 10^{-11}$</td>
<td>0.63</td>
<td>0.537</td>
</tr>
</tbody>
</table>

- GSP needed to stay in the model to allow scenario forecasting.
- All other variables led to improved AIC<sub>C</sub>.
### Annual model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta gsp.pc)</td>
<td>2.02 \times 10^{-6}</td>
<td>5.05 \times 10^{-6}</td>
<td>0.38</td>
<td>0.711</td>
</tr>
<tr>
<td>(\Delta price)</td>
<td>-1.67 \times 10^{-8}</td>
<td>6.76 \times 10^{-9}</td>
<td>-2.46</td>
<td>0.026</td>
</tr>
<tr>
<td>(\Delta scdd)</td>
<td>1.11 \times 10^{-10}</td>
<td>2.48 \times 10^{-11}</td>
<td>4.49</td>
<td>0.000</td>
</tr>
<tr>
<td>(\Delta whdd)</td>
<td>2.07 \times 10^{-11}</td>
<td>3.28 \times 10^{-11}</td>
<td>0.63</td>
<td>0.537</td>
</tr>
</tbody>
</table>

- GSP needed to stay in the model to allow scenario forecasting.
- All other variables led to improved AIC\(_C\).
Annual model

Forecasting electricity demand distributions

The model
Half-hourly models

\[
\log(y_t) = \log(y_t^*) + \log(\bar{y}_i)
\]

\[
\log(y_t^*) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + e_t
\]

- Separate model for each half-hour.
- Same predictors used for all models.

Each model is fitted to the data twice: first excluding the summer of 2009/2010 and then excluding the summer of 2010/2011. The average out-of-sample MSE is calculated from the omitted data for the time periods 12noon–8.30pm.
Half-hourly models

\[
\log(y_t) = \log(y^*_t) + \log(\bar{y}_i)
\]
\[
\log(y^*_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + e_t
\]

- Separate model for each half-hour.
- Same predictors used for all models.
- Each model is fitted to the data twice, first excluding the summer of 2009/2010 and then excluding the summer of 2010/2011. The average out-of-sample MSE is calculated from the omitted data for the time periods 12noon–8.30pm.
Half-hourly models

\[
\begin{align*}
\log(y_t) &= \log(y^*_t) + \log(\bar{y}_i) \\
\log(y^*_t) &= h_p(t) + f_p(w_{1,t}, w_{2,t}) + e_t
\end{align*}
\]

- Separate model for each half-hour.
- Same predictors used for all models.
- Each model is fitted to the data twice, first excluding the summer of 2009/2010 and then excluding the summer of 2010/2011. The average out-of-sample MSE is calculated from the omitted data for the time periods 12noon–8.30pm.
\[
\log(y_t) = \log(y_t^*) + \log(\bar{y}_i)
\]
\[
\log(y_t^*) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + e_t
\]

- Separate model for each half-hour.
- Same predictors used for all models.
- Each model is fitted to the data twice, first excluding the summer of 2009/2010 and then excluding the summer of 2010/2011. The average out-of-sample MSE is calculated from the omitted data for the time periods 12noon–8:30pm.
Half-hourly models

\[ \log(y_t) = \log(y^*_t) + \log(\bar{y}_i) \]
\[ \log(y^*_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + e_t \]

- Separate model for each half-hour.
- Same predictors used for all models.
- Each model is fitted to the data twice, first excluding the summer of 2009/2010 and then excluding the summer of 2010/2011. The average out-of-sample MSE is calculated from the omitted data for the time periods 12noon–8.30pm.
## Half-hourly models

<table>
<thead>
<tr>
<th></th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
<th>$x_6$</th>
<th>$x_{48}$</th>
<th>$x_{96}$</th>
<th>$x_{144}$</th>
<th>$x_{192}$</th>
<th>$x_{240}$</th>
<th>$x_{288}$</th>
<th>$d_1$</th>
<th>$d_2$</th>
<th>$d_3$</th>
<th>$d_4$</th>
<th>$d_5$</th>
<th>$d_6$</th>
<th>$d_{48}$</th>
<th>$d_{96}$</th>
<th>$d_{144}$</th>
<th>$d_{192}$</th>
<th>$d_{240}$</th>
<th>$d_{288}$</th>
<th>$x^+$</th>
<th>$x^-$</th>
<th>$\bar{x}$</th>
<th>dow</th>
<th>hol</th>
<th>dos</th>
<th>MSE</th>
</tr>
</thead>
</table>
| 1 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.037 |}
| 2 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.034 |}
| 3 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.031 |}
| 4 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.027 |}
| 5 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.025 |}
| 6 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.020 |}
| 7 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.025 |}
| 8 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.026 |}
| 9 |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.035 |}
| 10|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.044 |}
| 11|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.057 |}
| 12|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.076 |}
| 13|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.102 |}
| 14|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.018 |}
| 15|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.021 |}
| 16|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.037 |}
| 17|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.074 |}
| 18|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.152 |}
| 19|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.180 |}
| 20|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.021 |}
| 21|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.027 |}
| 22|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.038 |}
| 23|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.056 |}
| 24|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.086 |}
| 25|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.135 |}
| 26|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.009 |}
| 27|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.063 |}
| 28|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.028 |}
| 29|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 3.523 |}
| 30|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 2.143 |}
| 31|       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |       |       |       |         |         |           |           |           |           |       |       |       |     |     |     | 1.523 |}

Forecasting electricity demand distributions

The model
Half-hourly models

R-squared

R-squared (%)

Time of day

12 midnight 3:00 am 6:00 am 9:00 am 12 noon 3:00 pm 6:00 pm 9:00 pm 3:00 am 12 midnight
Half-hourly models

South Australian demand (January 2011)

- Actual
- Fitted

South Australian demand (GW)

Date in January

Forecasting electricity demand distributions
Half-hourly models

The model
Half-hourly models

Forecasting electricity demand distributions

The model
**Adjusted model**

**Original model**

\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j Z_{j,t} + n_t
\]

**Model allowing saturated usage**

\[
q_t = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j Z_{j,t} + n_t
\]

\[
\log(y_t) = \begin{cases} 
q_t & \text{if } q_t \leq \tau; \\
\tau + k(q_t - \tau) & \text{if } q_t > \tau.
\end{cases}
\]
Adjusted model

Original model

\[
\log(y_t) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

Model allowing saturated usage

\[
q_t = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

\[
\log(y_t) = \begin{cases} 
q_t & \text{if } q_t \leq \tau; \\
\tau + k(q_t - \tau) & \text{if } q_t > \tau.
\end{cases}
\]
Outline

1 The problem

2 The model

3 Long-term forecasts

4 Short term forecasts
Peak demand forecasting

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

Multiple alternative futures created:

- \( h_p(t) \) known;
- simulate future temperatures using double seasonal block bootstrap with variable blocks (with adjustment for climate change);
- use assumed values for GSP, population and price;
- resample residuals using double seasonal block bootstrap with variable blocks.
Peak demand forecasting

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j Z_{j,t} + n_t \]

Multiple alternative futures created:
- \( h_p(t) \) known;
- simulate future temperatures using double seasonal block bootstrap with variable blocks (with adjustment for climate change);
- use assumed values for GSP, population and price;
- resample residuals using double seasonal block bootstrap with variable blocks.
Peak demand forecasting

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

Multiple alternative futures created:

- \( h_p(t) \) known;
- simulate future temperatures using double seasonal block bootstrap with variable blocks (with adjustment for climate change);
- use assumed values for GSP, population and price;
- resample residuals using double seasonal block bootstrap with variable blocks.
Peak demand forecasting

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

Multiple alternative futures created:
- \( h_p(t) \) known;
- simulate future temperatures using double seasonal block bootstrap with variable blocks (with adjustment for climate change);
- use assumed values for GSP, population and price;
- resample residuals using double seasonal block bootstrap with variable blocks.
Seasonal block bootstrapping

Conventional seasonal block bootstrap

- Same as block bootstrap but with whole years as the blocks to preserve seasonality.
- But we only have about 10–15 years of data, so there is a limited number of possible bootstrap samples.

Double seasonal block bootstrap

Suitable when there are two seasonal periods (here we have years of 151 days and days of 48 half-hours).
Conventional seasonal block bootstrap

- Same as block bootstrap but with whole years as the blocks to preserve seasonality.
- But we only have about 10–15 years of data, so there is a limited number of possible bootstrap samples.

Double seasonal block bootstrap

- Suitable when there are two seasonal periods (here we have years of 151 days and days of 48 half-hours).
- Divide each year into blocks of length 48m.
- Block 1 consists of the first m days of the year, block 2 consists of the next m days, and so on.
- Bootstrap sample consists of a sample of blocks where each block may come from a different randomly selected year but must be at the correct time of year.
Seasonal block bootstrapping

Conventional seasonal block bootstrap

- Same as block bootstrap but with whole years as the blocks to preserve seasonality.
- But we only have about 10–15 years of data, so there is a limited number of possible bootstrap samples.

Double seasonal block bootstrap

- Suitable when there are two seasonal periods (here we have years of 151 days and days of 48 half-hours).
- Divide each year into blocks of length $48m$.
- Block 1 consists of the first $m$ days of the year, block 2 consists of the next $m$ days, and so on.
- Bootstrap sample consists of a sample of blocks where each block may come from a different randomly selected year but must be at the correct time of year.
**Seasonal block bootstrapping**

**Conventional seasonal block bootstrap**
- Same as block bootstrap but with whole years as the blocks to preserve seasonality.
- But we only have about 10–15 years of data, so there is a limited number of possible bootstrap samples.

**Double seasonal block bootstrap**
- Suitable when there are two seasonal periods (here we have years of 151 days and days of 48 half-hours).
- Divide each year into blocks of length $48m$.
- Block 1 consists of the first $m$ days of the year, block 2 consists of the next $m$ days, and so on.
- Bootstrap sample consists of a sample of blocks where each block may come from a different randomly selected year but must be at the correct time of year.
Seasonal block bootstrapping

Conventional seasonal block bootstrap
- Same as block bootstrap but with whole years as the blocks to preserve seasonality.
- But we only have about 10–15 years of data, so there is a limited number of possible bootstrap samples.

Double seasonal block bootstrap
- Suitable when there are two seasonal periods (here we have years of 151 days and days of 48 half-hours).
- **Divide each year into blocks of length 48m.**
  - Block 1 consists of the first \( m \) days of the year, block 2 consists of the next \( m \) days, and so on.
  - Bootstrap sample consists of a sample of blocks where each block may come from a different randomly selected year but must be at the correct time of year.
Seasonal block bootstrapping

Conventional seasonal block bootstrap

- Same as block bootstrap but with whole years as the blocks to preserve seasonality.
- But we only have about 10–15 years of data, so there is a limited number of possible bootstrap samples.

Double seasonal block bootstrap

- Suitable when there are two seasonal periods (here we have years of 151 days and days of 48 half-hours).
- Divide each year into blocks of length $48m$.
- Block 1 consists of the first $m$ days of the year, block 2 consists of the next $m$ days, and so on.
- Bootstrap sample consists of a sample of blocks where each block may come from a different randomly selected year but must be at the correct time of year.
Seasonal block bootstrapping

**Conventional seasonal block bootstrap**

- Same as block bootstrap but with whole years as the blocks to preserve seasonality.
- But we only have about 10–15 years of data, so there is a limited number of possible bootstrap samples.

**Double seasonal block bootstrap**

- Suitable when there are two seasonal periods (here we have years of 151 days and days of 48 half-hours).
- Divide each year into blocks of length $48m$.
- Block 1 consists of the first $m$ days of the year, block 2 consists of the next $m$ days, and so on.
- Bootstrap sample consists of a sample of blocks where each block may come from a different randomly selected year but must be at the correct time of year.
Seasonal block bootstrapping

Forecasting electricity demand distributions

Actual temperatures

Bootstrap temperatures (fixed blocks)

Bootstrap temperatures (variable blocks)
Seasonal block bootstrapping

Problems with the double seasonal bootstrap

- Boundaries between blocks can introduce large jumps. However, only at midnight.
- Number of values that any given time in year is still limited to the number of years in the data set.
Seasonal block bootstrapping

Problems with the double seasonal bootstrap

- Boundaries between blocks can introduce large jumps. However, only at midnight.
- Number of values that any given time in year is still limited to the number of years in the data set.
Seasonal block bootstrapping

Variable length double seasonal block bootstrap

- Blocks allowed to vary in length between $m - \Delta$ and $m + \Delta$ days where $0 \leq \Delta < m$.
- Blocks allowed to move up to $\Delta$ days from their original position.
- Has little effect on the overall time series patterns provided $\Delta$ is relatively small.
- Use uniform distribution on $(m - \Delta, m + \Delta)$ to select block length, and independent uniform distribution on $(-\Delta, \Delta)$ to select variation on starting position for each block.
Seasonal block bootstrapping

Variable length double seasonal block bootstrap

- Blocks allowed to vary in length between $m - \Delta$ and $m + \Delta$ days where $0 \leq \Delta < m$.
- Blocks allowed to move up to $\Delta$ days from their original position.
- Has little effect on the overall time series patterns provided $\Delta$ is relatively small.
- Use uniform distribution on $(m - \Delta, m + \Delta)$ to select block length, and independent uniform distribution on $(-\Delta, \Delta)$ to select variation on starting position for each block.
Seasonal block bootstrapping

Variable length double seasonal block bootstrap

- Blocks allowed to vary in length between $m - \Delta$ and $m + \Delta$ days where $0 \leq \Delta < m$.
- Blocks allowed to move up to $\Delta$ days from their original position.
- Has little effect on the overall time series patterns provided $\Delta$ is relatively small.
- Use uniform distribution on $(m - \Delta, m + \Delta)$ to select block length, and independent uniform distribution on $(-\Delta, \Delta)$ to select variation on starting position for each block.
Seasonal block bootstrapping

Variable length double seasonal block bootstrap

- Blocks allowed to vary in length between \( m - \Delta \) and \( m + \Delta \) days where \( 0 \leq \Delta < m \).
- Blocks allowed to move up to \( \Delta \) days from their original position.
- Has little effect on the overall time series patterns provided \( \Delta \) is relatively small.
- Use uniform distribution on \((m - \Delta, m + \Delta)\) to select block length, and independent uniform distribution on \((-\Delta, \Delta)\) to select variation on starting position for each block.
Seasonal block bootstrapping

Forecasting electricity demand distributions

Long-term forecasts
Seasonal block bootstrapping

Forecasting electricity demand distributions

Long-term forecasts
Peak demand forecasting

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_jz_{j,t} + n_t \]

Multiple alternative futures created:

- \( h_p(t) \) known;
- simulate future temperatures using double seasonal block bootstrap with variable blocks (with adjustment for climate change);
- use assumed values for GSP, population and price;
- resample residuals using double seasonal block bootstrap with variable blocks.
Peak demand backcasting

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

Multiple alternative pasts created:

- \( h_p(t) \) known;
- simulate past temperatures using double seasonal block bootstrap with variable blocks;
- use actual values for GSP, population and price;
- resample residuals using double seasonal block bootstrap with variable blocks.
Estimated historical quantiles

PoE (annual interpretation)

Year
PoE Demand
98/99 00/01 02/03 04/05 06/07 08/09 10/11
10 %
50 %
90 %
●
●
●
●
●
●
●
●
●
●
●

Forecasting electricity demand distributions
Long-term forecasts
Peak demand forecasting

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j Z_{j,t} + n_t \]

Multiple alternative futures created:

- \( h_p(t) \) known;
- simulate future temperatures using double seasonal block bootstrap with variable blocks (with adjustment for climate change);
- use assumed values for GSP, population and price;
- resample residuals using double seasonal block bootstrap with variable blocks.
Peak demand forecasting

South Australia GSP

South Australia population

Average electricity prices
Peak demand distribution

PoE (annual interpretation)

<table>
<thead>
<tr>
<th>Year</th>
<th>PoE Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>98/99</td>
<td>2.0</td>
</tr>
<tr>
<td>00/01</td>
<td>2.5</td>
</tr>
<tr>
<td>02/03</td>
<td>3.0</td>
</tr>
<tr>
<td>04/05</td>
<td>3.5</td>
</tr>
<tr>
<td>06/07</td>
<td>4.0</td>
</tr>
<tr>
<td>08/09</td>
<td>10 %</td>
</tr>
<tr>
<td>10/11</td>
<td>50 %</td>
</tr>
</tbody>
</table>

Forecasting electricity demand distributions

Long-term forecasts
Peak demand distribution

Annual POE levels

Year

PoE Demand

1 % POE
5 % POE
10 % POE
50 % POE
90 % POE

Actual annual maximum
Peak demand forecasting

Forecasting electricity demand distributions

Long-term forecasts
Outline

1 The problem

2 The model

3 Long-term forecasts

4 Short term forecasts

Forecasting electricity demand distributions
Short term forecasts

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- Bootstrapping temperatures and residuals is ok for long-term forecasts because short-term dynamics wash out after a few weeks.
- But short-term forecasts need to take account of recent temperatures and recent residuals due to serial correlation.
- Short-term temperature forecasts are available.
- Building a separate model for \( n_t \) is possible, but there is a simpler approach.
Short term forecasts

\[ q_{t,p} = h_p(t) + f_p(\mathbf{w}_{1,t}, \mathbf{w}_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- Bootstrapping temperatures and residuals is ok for long-term forecasts because short-term dynamics wash out after a few weeks.
- But short-term forecasts need to take account of recent temperatures and recent residuals due to serial correlation.
- Short-term temperature forecasts are available.
- Building a separate model for \( n_t \) is possible, but there is a simpler approach.
Short term forecasts

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- Bootstrapping temperatures and residuals is ok for long-term forecasts because short-term dynamics wash out after a few weeks.
- But short-term forecasts need to take account of recent temperatures and recent residuals due to serial correlation.
- Short-term temperature forecasts are available.
- Building a separate model for \( n_t \) is possible, but there is a simpler approach.
Short term forecasts

\[ q_{t,p} = h_p(t) + f_p(w_{1,t}, w_{2,t}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- Bootstrapping temperatures and residuals is ok for long-term forecasts because short-term dynamics wash out after a few weeks.
- But short-term forecasts need to take account of recent temperatures and recent residuals due to serial correlation.
- **Short-term temperature forecasts are available.**
- Building a separate model for \( n_t \) is possible, but there is a simpler approach.
Short term forecasts

\[ q_{t,p} = h_p(t) + f_p(\mathbf{w}_{1,t}, \mathbf{w}_{2,t}) + \sum_{j=1}^{J} c_j Z_{j,t} + n_t \]

- Bootstrapping temperatures and residuals is ok for long-term forecasts because short-term dynamics wash out after a few weeks.
- But short-term forecasts need to take account of recent temperatures and recent residuals due to serial correlation.
- Short-term temperature forecasts are available.
- Building a separate model for \( n_t \) is possible, but there is a simpler approach.
Short-term forecasting model

\[
\log(y_{t,p}) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + a_p(y_{t-1}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

- \( y_{t,p} \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) during period \( p \), \( p = 1, \ldots, 48 \);
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \);
- \( n_t \) denotes the model error at time \( t \);
- \( y_t = [y_t, y_{t-1}, y_{t-2}, \ldots] \);
- \( a_p(y_{t-1}) \) models effects of recent demands.
Short-term forecasting model

\[
\log(y_{t,p}) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + a_p(y_{t-1}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

- \( y_{t,p} \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) during period \( p \), \( p = 1, \ldots, 48 \);
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \);
- \( n_t \) denotes the model error at time \( t \);
- \( y_t = [y_t, y_{t-1}, y_{t-2}, \ldots] \);
- \( a_p(y_{t-1}) \) models effects of recent demands.
Short-term forecasting model

\[
\log(y_{t,p}) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + a_p(y_{t-1}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

- \( y_{t,p} \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) during period \( p \), \( p = 1, \ldots, 48; \)
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \)
- \( n_t \) denotes the model error at time \( t \)
- \( y_t = [y_t, y_{t-1}, y_{t-2}, \ldots] \)
- \( a_p(y_{t-1}) \) models effects of recent demands
Short-term forecasting model

\[
\log(y_{t,p}) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + a_p(y_{t-1}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

- \(y_{t,p}\) denotes per capita demand (minus offset) at time \(t\) (measured in half-hourly intervals) during period \(p\), \(p = 1, \ldots, 48\);
- \(h_p(t)\) models all calendar effects;
- \(f_p(w_{1,t}, w_{2,t})\) models all temperature effects where \(w_{1,t}\) is a vector of recent temperatures at location 1 and \(w_{2,t}\) is a vector of recent temperatures at location 2;
- \(z_{j,t}\) is a demographic or economic variable at time \(t\);
- \(n_t\) denotes the model error at time \(t\);
- \(y_t = [y_t, y_{t-1}, y_{t-2}, \ldots]\);
- \(a_p(y_{t-1})\) models effects of recent demands.
Short-term forecasting model

\[ \log(y_{t,p}) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + a_p(y_{t-1}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t \]

- \( y_{t,p} \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) during period \( p \), \( p = 1, \ldots, 48 \);
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \);
- \( n_t \) denotes the model error at time \( t \);
- \( y_t = [y_{t}, y_{t-1}, y_{t-2}, \ldots] \);
- \( a_p(y_{t-1}) \) models effects of recent demands.
Short-term forecasting model

\[
\log(y_{t,p}) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + a_p(y_{t-1}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

- \( y_{t,p} \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) during period \( p \), \( p = 1, \ldots, 48 \);
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \);
- \( n_t \) denotes the model error at time \( t \);
- \( y_t = [y_t, y_{t-1}, y_{t-2}, \ldots] \);
- \( a_p(y_{t-1}) \) models effects of recent demands.
Short-term forecasting model

\[
\log(y_{t,p}) = h_p(t) + f_p(w_{1,t}, w_{2,t}) + a_p(y_{t-1}) + \sum_{j=1}^{J} c_j z_{j,t} + n_t
\]

- \( y_{t,p} \) denotes per capita demand (minus offset) at time \( t \) (measured in half-hourly intervals) during period \( p \), \( p = 1, \ldots, 48 \);
- \( h_p(t) \) models all calendar effects;
- \( f_p(w_{1,t}, w_{2,t}) \) models all temperature effects where \( w_{1,t} \) is a vector of recent temperatures at location 1 and \( w_{2,t} \) is a vector of recent temperatures at location 2;
- \( z_{j,t} \) is a demographic or economic variable at time \( t \);
- \( n_t \) denotes the model error at time \( t \);
- \( y_t = [y_t, y_{t-1}, y_{t-2}, \ldots] \);
- \( a_p(y_{t-1}) \) models effects of recent demands.
Short-term forecasting model

\[ a_p(y_{t-1}) = \sum_{k=1}^{n} b_{k,p}(y_{t-k}) + \sum_{j=1}^{m} B_{j,p}(y_{t-48j}) \]
\[ + Q_p(y_t^+) + R_p(y_t^-) + S_p(\bar{y}_t) \]

where

- \( y_t^+ \) is maximum of \( y_t \) values in past 24 hours;
- \( y_t^- \) is minimum of \( y_t \) values in past 24 hours;
- \( \bar{y}_t \) is average demand in past 7 days
- \( b_{k,p}, B_{j,p}, Q_p, R_p \) and \( S_p \) are estimated using cubic splines.
Short-term forecasting model

\[ a_p(y_{t-1}) = \sum_{k=1}^{n} b_{k,p}(y_{t-k}) + \sum_{j=1}^{m} B_{j,p}(y_{t-48j}) \]
\[ + Q_p(y_t^+) + R_p(y_t^-) + S_p(\bar{y}_t) \]

where

- \( y_t^+ \) is maximum of \( y_t \) values in past 24 hours;
- \( y_t^- \) is minimum of \( y_t \) values in past 24 hours;
- \( \bar{y}_t \) is average demand in past 7 days
- \( b_{k,p}, B_{j,p}, Q_p, R_p \) and \( S_p \) are estimated using cubic splines.
Short-term forecasting model

\[
a_p(y_{t-1}) = \sum_{k=1}^{n} b_{k,p}(y_{t-k}) + \sum_{j=1}^{m} B_{j,p}(y_{t-48j}) \\
+ Q_p(y^+_t) + R_p(y^-_t) + S_p(\bar{y}_t)
\]

where

- \(y^+_t\) is maximum of \(y_t\) values in past 24 hours;
- \(y^-_t\) is minimum of \(y_t\) values in past 24 hours;
- \(\bar{y}_t\) is average demand in past 7 days

\(b_{k,p}, B_{j,p}, Q_p, R_p\) and \(S_p\) are estimated using cubic splines.
Short-term forecasting model

\[ a_p(y_{t-1}) = \sum_{k=1}^{n} b_{k,p}(y_{t-k}) + \sum_{j=1}^{m} B_{j,p}(y_{t-48j}) \]
\[ + Q_p(y_t^+) + R_p(y_t^-) + S_p(\bar{y}_t) \]

where

- \( y_t^+ \) is maximum of \( y_t \) values in past 24 hours;
- \( y_t^- \) is minimum of \( y_t \) values in past 24 hours;
- \( \bar{y}_t \) is average demand in past 7 days
- \( b_{k,p}, B_{j,p}, Q_p, R_p \) and \( S_p \) are estimated using cubic splines.