Forecasting the residual demand function in electricity auctions

Matteo Pelagatti

Università degli Studi di Milano-Bicocca

SAS Campus 2011

Università degli Studi di Milano-Bicocca
22 November 2011
Models for the optimal bidding behaviour of suppliers in electricity auctions assume that the aggregate supply curve of the competitors is known (Wolak, 2003; Hortaçsu and Puller, 2008; Bosco et al., 2010).

Competitors’ cost functions may be approximately known, since technologies of power plants are generally standard, but their supply functions are *ex-ante* unknown.
Models for the optimal bidding behaviour of suppliers in electricity auctions assume that the aggregate supply curve of the competitors is known (Wolak, 2003; Hortaçsu and Puller, 2008; Bosco et al., 2010).

Competitors’ cost functions may be approximately known, since technologies of power plants are generally standard, but their supply functions are *ex-ante* unknown.
The framework

- Firm’s $i$ profit function:

$$
\pi(p) = (D - S_{-i}(p))p - C(D - S_{-i}(p)),
$$

with $D$ inelastic demand for electricity, $C(q)$ cost function and $S_{-i}(p)$ supply function of all firms but $i$.

- In equilibrium firm $i$ should bid as

$$
p^e = C'(D - S_{-i}(p^e)) + \frac{D - S_{-i}(p^e)}{S'_{-i}(p^e)}.
$$

- Firm $i$ needs inference on $S_{-i}(p)$.
Firm’s $i$ profit function:

$$\pi(p) = (D - S_{-i}(p))p - C(D - S_{-i}(p)), $$

with $D$ inelastic demand for electricity, $C(q)$ cost function and $S_{-i}(p)$ supply function of all firms but $i$.

In equilibrium firm $i$ should bid as

$$p^e = C'(D - S_{-i}(p^e)) + \frac{D - S_{-i}(p^e)}{S'_{-i}(p^e)}. $$

Firm $i$ needs inference on $S_{-i}(p)$. 

Pelagatti (Bicocca U.) Residual demand forecast 22 November 2011 3 / 26
The framework

- Firm’s $i$ profit function:

$$\pi(p) = (D - S_{-i}(p))p - C(D - S_{-i}(p)),$$

with $D$ inelastic demand for electricity, $C(q)$ cost function and $S_{-i}(p)$ supply function of all firms but $i$.

- In equilibrium firm $i$ should bid as

$$p^e = C'(D - S_{-i}(p^e)) + \frac{D - S_{-i}(p^e)}{S'_{-i}(p^e)}.$$

- Firm $i$ needs inference on $S_{-i}(p)$.
Something about $S_i(p)$

- The rules of electricity auctions allow to submit a small number of price-quantity pairs for every plant.
- Thus, supply schedules are step functions.
- Since in models for bidding strategies supply functions are assumed differentiable, real supply function can be approximated by kernel-smoothed versions of them

$$\hat{S}(p) := \sum_{i=1}^{n} q_i \Phi \left( \frac{p - p_i}{\gamma} \right),$$

where $\gamma$ is a bandwidth parameter and $\Phi$ is the Gaussian cumulative distribution function.

- The derivative of $\hat{S}$ is given by

$$\hat{S}'(p) := \frac{1}{\gamma} \sum_{i=1}^{n} q_i \phi \left( \frac{p - p_i}{\gamma} \right).$$
The rules of electricity auctions allow to submit a small number of price-quantity pairs for every plant.

Thus, supply schedules are step functions.

Since in models for bidding strategies supply functions are assumed differentiable, real supply function can be approximated by kernel-smoothed versions of them

\[ \hat{S}(p) := \sum_{i=1}^{n} q_i \Phi \left( \frac{p - p_i}{\gamma} \right), \]

where \( \gamma \) is a bandwidth parameter and \( \Phi \) is the Gaussian cumulative distribution function.

The derivative of \( \hat{S} \) is given by

\[ \hat{S}'(p) := \frac{1}{\gamma} \sum_{i=1}^{n} q_i \phi \left( \frac{p - p_i}{\gamma} \right). \]
The rules of electricity auctions allow to submit a small number of price-quantity pairs for every plant.

Thus, supply schedules are step functions.

Since in models for bidding strategies supply functions are assumed differentiable, real supply function can be approximated by kernel-smoothed versions of them

\[ \hat{S}(p) := \sum_{i=1}^{n} q_i \Phi \left( \frac{p - p_i}{\gamma} \right), \]

where \( \gamma \) is a bandwidth parameter and \( \Phi \) is the Gaussian cumulative distribution function.

The derivative of \( \hat{S} \) is given by

\[ \hat{S}'(p) := \frac{1}{\gamma} \sum_{i=1}^{n} q_i \phi \left( \frac{p - p_i}{\gamma} \right). \]
Something about $S_i(p)$

- The rules of electricity auctions allow to submit a small number of price-quantity pairs for every plant.
- Thus, supply schedules are step functions.
- Since in models for bidding strategies supply functions are assumed differentiable, real supply function can be approximated by kernel-smoothed versions of them

$$\hat{S}(p) := \sum_{i=1}^{n} q_i \Phi \left( \frac{p - p_i}{\gamma} \right),$$

where $\gamma$ is a bandwidth parameter and $\Phi$ is the Gaussian cumulative distribution function.

- The derivative of $\hat{S}$ is given by

$$\hat{S}'(p) := \frac{1}{\gamma} \sum_{i=1}^{n} q_i \phi \left( \frac{p - p_i}{\gamma} \right).$$
Examples of $S_i(p)$ and $\hat{S}_i(p)$
We have one function for every auction/hour.

The function is not evaluated at the same price values.

We sample the function at a constant grid of \( n \) price points obtaining \( n \) time series of offered quantities.

We sample more frequently where bid-prices are more dense (we use quantiles).

Should we sample the actual function or the smooth version of it? Smooth. See why in the next slide.
Sampling the function

- We have one function for every auction/hour.
- **The function is not evaluated at the same price values.**
- We sample the function at a constant grid of $n$ price points obtaining $n$ time series of offered quantities.
- We sample more frequently where bid-prices are more dense (we use quantiles).
- Should we sample the actual function or the smooth version of it? Smooth. See why in the next slide.
Sampling the function

- We have one function for every auction/hour.
- The function is not evaluated at the same price values.
- We sample the function at a constant grid of $n$ price points obtaining $n$ time series of offered quantities.
- We sample more frequently where bid-prices are more dense (we use quantiles).
- Should we sample the actual function or the smooth version of it? Smooth. See why in the next slide.
We have one function for every auction/hour.
The function is not evaluated at the same price values.
We sample the function at a constant grid of $n$ price points obtaining $n$ time series of offered quantities.
We sample more frequently where bid-prices are more dense (we use quantiles).
Should we sample the actual function or the smooth version of it? Smooth. See why in the next slide.
Sampling the function

- We have one function for every auction/hour.
- The function is not evaluated at the same price values.
- We sample the function at a constant grid of $n$ price points obtaining $n$ time series of offered quantities.
- We sample more frequently where bid-prices are more dense (we use quantiles).
- Should we sample the actual function or the smooth version of it? Smooth. See why in the next slide.
Price grid for sampling the supply function

Aggregate supply function and sampling grid

Price (Euro)
MWh
18000
20000
22000
24000
26000

0 100 200 300 400 500
The data

- We collected the bids (price and quantity) of every plant of all the firms operating in the Italian market from 2005 to 2008 (now also 2009 and 2010).
- The models we show here are fitted on 2007-2008 and evaluated on Jan 2009.
- We sample Enel’s competitors’ supply function at 51 price-points (percentiles 0% (min), 2%, 4%, ..., 98%, 100% (max), min = 0, max = 500).
The data

- We collected the bids (price and quantity) of every plant of all the firms operating in the Italian market from 2005 to 2008 (now also 2009 and 2010).
- The models we show here are fitted on 2007-2008 and evaluated on Jan 2009.
- We sample Enel’s competitors’ supply function at 51 price-points (percentiles 0% (min), 2%, 4%, …, 98%, 100% (max), min = 0, max = 500).
We collected the bids (price and quantity) of every plant of all the firms operating in the Italian market from 2005 to 2008 (now also 2009 and 2010).

The models we show here are fitted on 2007-2008 and evaluated on Jan 2009.

We sample Enel’s competitors’ supply function at 51 price-points (percentiles 0% (min), 2%, 4%, …, 98%, 100% (max), min = 0, max = 500).
### One week of functions

<table>
<thead>
<tr>
<th>Time</th>
<th>Price</th>
<th>Supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>10000</td>
</tr>
<tr>
<td>24</td>
<td>100</td>
<td>10000</td>
</tr>
<tr>
<td>48</td>
<td>150</td>
<td>15000</td>
</tr>
<tr>
<td>72</td>
<td>200</td>
<td>20000</td>
</tr>
<tr>
<td>96</td>
<td>250</td>
<td>20000</td>
</tr>
<tr>
<td>120</td>
<td>120</td>
<td>15000</td>
</tr>
<tr>
<td>144</td>
<td>144</td>
<td>10000</td>
</tr>
<tr>
<td>168</td>
<td>168</td>
<td>5000</td>
</tr>
</tbody>
</table>

Residual demand forecast

22 November 2011
Supply functions are positive non-decreasing and it is certainly a desirable property that predictions obey the same rule.

In order to guarantee this, instead of modelling the vector of sampled supply function points $S_t(p)$, we model

$$q_{i,t} := \begin{cases} 
\log S_t(p_i), & \text{for } i = 0; \\
\log (S_t(p_i) - S_t(p_{i-1}) + c), & \text{for } i = 1, \ldots, 50,
\end{cases}$$

with $c$ small constant.
Dealing with positive increments

- Supply functions are positive non-decreasing and it is certainly a desirable property that predictions obey the same rule.
- In order to guarantee this, instead of modelling the vector of sampled supply function points $S_t(p)$, we model

$$q_{i,t} := \begin{cases} \log S_t(p_i), & \text{for } i = 0; \\ \log (S_t(p_i) - S_t(p_{i-1}) + c), & \text{for } i = 1, \ldots, 50, \end{cases}$$

with $c$ small constant.
Inverting the transform

If we assume that the predictor of $q_{i,t}$, say $\hat{q}_{i,t}$, is unbiased, and the prediction error is approximately normal with standard error $s_{i,t}$, unbiased forecasts of the original function can be recovered as

$$\hat{S}_t(p_i) = \begin{cases} 
\exp(\hat{q}_{i,t} + s_{i,t}^2/2), & \text{for } i = 0; \\
\exp(\hat{q}_{i,t} + s_{i,t}^2/2) + \hat{S}_t(p_{i-1}) - c, & \text{for } i = 1 \ldots, 50.
\end{cases}$$
Issues with the data

- The ordered (sampled) function points could be considered as part of just one time series, but this would add even more complexity:
  - Many interacting seasonal patterns: within-year, -week, -day and -auction.
  - High frequency data: 51 obs. per auction, 24 auctions per day, 7 days per week, 52 weeks per year = 446760 data points per year.
  - Within-auction data (sampled supply functions) must be non-decreasing.

- Some form of data reduction may be necessary. We begin with the simplest, that is, linear maps:
  - Principal components.
  - Reduced rank regression.
Issues with the data

- The ordered (sampled) function points could be considered as part of just one time series, but this would add even more complexity:
  - Many interacting seasonal patterns: within-year, -week, -day and -auction.
  - High frequency data: 51 obs. per auction, 24 auctions per day, 7 days per week, 52 weeks per year = 446760 data points per year.
  - Within-auction data (sampled supply functions) must be non-decreasing.

- Some form of data reduction may be necessary. We begin with the simplest, that is, linear maps:
  - Principal components.
  - Reduced rank regression.
The ordered (sampled) function points could be considered as part of just one time series, but this would add even more complexity:

- Many interacting seasonal patterns: within-year, -week, -day and -auction.
- High frequency data: 51 obs. per auction, 24 auctions per day, 7 days per week, 52 weeks per year = 446760 data points per year.
- Within-auction data (sampled supply functions) must be non-decreasing.

Some form of data reduction may be necessary. We begin with the simplest, that is, linear maps:

- Principal components.
- Reduced rank regression.
Issues with the data

- The ordered (sampled) function points could be considered as part of just one time series, but this would add even more complexity:
  - Many interacting seasonal patterns: within-year, -week, -day and -auction.
  - High frequency data: 51 obs. per auction, 24 auctions per day, 7 days per week, 52 weeks per year = 446760 data points per year.
  - Within-auction data (sampled supply functions) must be non-decreasing.

- Some form of data reduction may be necessary. We begin with the simplest, that is, linear maps:
  - Principal components.
  - Reduced rank regression.
How we proceed

- We fit the models to 2007-2008 data and use Jan-2009 for out-of-sample prediction evaluation. We also try cross validation with 24 observation blocks.
- We forecast 1-step (one hour) and 24-steps (one day) ahead.
- The rank (dimensionality) of each model is determined through out-of-sample fit (MSE).
- In each model we will denote with $y_t$ the vector of time series to forecast, $x_t$ its lags and $z_t$ deterministic regressors.
How we proceed

- We fit the models to 2007-2008 data and use Jan-2009 for out-of-sample prediction evaluation. We also try cross validation with 24 observation blocks.

- We forecast 1-step (one hour) and 24-steps (one day) ahead.

- The rank (dimensionality) of each model is determined through out-of-sample fit (MSE).

- In each model we will denote with $y_t$ the vector of time series to forecast, $x_t$ its lags and $z_t$ deterministic regressors.
How we proceed

- We fit the models to 2007-2008 data and use Jan-2009 for out-of-sample prediction evaluation. We also try cross validation with 24 observation blocks.
- We forecast 1-step (one hour) and 24-steps (one day) ahead.
- The rank (dimensionality) of each model is determined through out-of-sample fit (MSE).
- In each model we will denote with $y_t$ the vector of time series to forecast, $x_t$ its lags and $z_t$ deterministic regressors.
How we proceed

- We fit the models to 2007-2008 data and use Jan-2009 for out-of-sample prediction evaluation. We also try cross validation with 24 observation blocks.
- We forecast 1-step (one hour) and 24-steps (one day) ahead.
- The rank (dimensionality) of each model is determined through out-of-sample fit (MSE).
- In each model we will denote with $y_t$ the vector of time series to forecast, $x_t$ its lags and $z_t$ deterministic regressors.
for $r = 51$ to $1$

1. Take the first $r$ PCs of $q_t$ (supply function log increments) based on its in-sample covariance matrix, and name the scores $y_t$.

2. Regress each score $y_{i,t}$ on its lags $x_{i,t}$ and deterministic regressors $z_t$ and compute predictions $\hat{y}_t$.

3. Regress the vector $q_t$ on the predicted scores $\hat{y}_t$ (and a constant).

4. Compute the out-of-sample predictions of the supply function $S_t$ and the relative MSE.

}  
Pick the rank $r$ that minimize the out-of-sample MSE.
Principal components approach
(PROC PLS METHOD=PCR)

for $r = 51$ to 1
{

1. Take the first $r$ PCs of $q_t$ (supply function log increments) based on its in-sample covariance matrix, and name the scores $y_t$.
2. Regress each score $y_{i,t}$ on its lags $x_{i,t}$ and deterministic regressors $z_t$ and compute predictions $\hat{y}_t$.
3. Regress the vector $q_t$ on the predicted scores $\hat{y}_t$ (and a constant).
4. Compute the out-of-sample predictions of the supply function $S_t$ and the relative MSE.
}

Pick the rank $r$ that minimize the out-of-sample MSE.
for \( r = 51 \) to 1

1. Take the first \( r \) PCs of \( q_t \) (supply function log increments) based on its in-sample covariance matrix, and name the scores \( y_t \).

2. Regress each score \( y_{i,t} \) on its lags \( x_{i,t} \) and deterministic regressors \( z_t \) and compute predictions \( \hat{y}_t \).

3. Regress the vector \( q_t \) on the predicted scores \( \hat{y}_t \) (and a constant).

4. Compute the out-of-sample predictions of the supply function \( S_t \) and the relative MSE.

} 

Pick the rank \( r \) that minimize the out-of-sample MSE.
Principal components approach
(PROC PLS METHOD=PCR)

for \( r = 51 \) to 1
{
1. Take the first \( r \) PCs of \( q_t \) (supply function log increments) based on its in-sample covariance matrix, and name the scores \( y_t \).
2. Regress each score \( y_{i,t} \) on its lags \( x_{i,t} \) and deterministic regressors \( z_t \) and compute predictions \( \hat{y}_t \).
3. Regress the vector \( q_t \) on the predicted scores \( \hat{y}_t \) (and a constant).
4. Compute the out-of-sample predictions of the supply function \( S_t \) and the relative MSE.
}
Pick the rank \( r \) that minimize the out-of-sample MSE.
Reduced rank regression

\[ y_t = \Pi x_t + \Gamma z_t + \epsilon_t, \]

where \( \Pi \) has reduced rank \( r < \min(n, m) \). The reduced rank is imposed by the restriction

\[ \Pi = A \cdot B. \]

The LS estimates of the matrix \( B \) are given by the first \( r \) eigenvectors of the following generalized eigenvalue problem

\[ S_{xx|z} V \Lambda = S_{xy|z} S_{yy|z}^{-1} S_{yx|z} V. \]

where \( S_{ab|c} \) represent a (sample) partial second moment matrix. The matrices \( A \) and \( \Gamma \) are derived by regressing \( y_t \) on \( Bx_t \) and \( z_t \).
for \( r = 51 \) to 1

1. Regress the vector \( y_t = q_t \) on its lags \( x_t \) imposing rank \( r \) on the matrix \( \Pi \) and on the deterministic components \( z_t \) without restricting the rank.

2. Compute the out-of-sample predictions of the supply function \( S_t \) and the relative MSE.

} \( r \) that minimize the out-of-sample MSE.
Reduced rank regression approach  
(PROC PLS METHOD=RRR)

for \( r = 51 \) to 1

\{ 

1. Regress the vector \( y_t = q_t \) on its lags \( x_t \) imposing rank \( r \) on the matrix \( \Pi \) and on the deterministic components \( z_t \) without restricting the rank.

2. Compute the out-of-sample predictions of the supply function \( S_t \) and the relative MSE.

\}

Pick the rank \( r \) that minimize the out-of-sample MSE.
**Comparison of the two approaches**

**Main advantage of PC over RRR** Any univariate time series model can be used on the PCs. For example, we could predict the PCs using UC models, non-linear models, neural networks, ... 

**Main advantage of RRR over RRR** The linear combinations that reduce the dimensions of the problem are optimized for forecasting.
The simple models we used

Deterministic regressors $z_t$:  
1. Linear trend $+ \cos(\omega_j) + \sin(\omega_j)$, with $\omega_j = j \cdot 2\pi/(24 \times 365.25)$ and $j = 1, \ldots, 20$.  
2. $1 +$ dummies for Saturday, Sunday and Monday.  
3. $2 + \cos(\omega_j) + \sin(\omega_j)$, with $\omega_j = j \cdot 2\pi/24$ and $j = 1, \ldots, 6$.  

Dynamic models:  
Level 1-step $y_t$ on $y_{t-1}, y_{t-24}, y_{t-168}, z_t$  
Diff 1-step $\Delta y_t$ on $y_{t-1}, \Delta y_{t-1}, \Delta y_{t-24}, \Delta y_{t-168}, \Delta z_t$  
Level 24-step $y_t$ on $y_{t-24}, y_{t-168}, z_t$  
Diff 24-step $\Delta_{24} y_t$ on $y_{t-24}, \Delta_{24} y_{t-24}, \Delta_{24} y_{t-168}, \Delta_{24} z_t$
Example Cross Validation: RRR 1h-step prediction

Cross-Validation Analysis

Selected = 51 (T Square test)

Root Mean PRESS
Model Term R Square
T Square
Dependent Variable R Square
Significantly inferior

Root Mean PRESS
Model Term R Square
T Square
Dependent Variable R Square
Significantly inferior
Example Cross Validation: RRR 24h-step prediction

Cross-Validation Analysis

Root Mean PRESS

R Square

Number of Factors

Selected = 51 (T Square test)

Root Mean PRESS
Model Term R Square
T Square
Dependent Variable R Square
Significantly inferior
## Out-of-sample results

<table>
<thead>
<tr>
<th></th>
<th>Det. Reg 1</th>
<th></th>
<th>Det. Reg 2</th>
<th></th>
<th>Det. Reg 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rank</td>
<td>RMSE</td>
<td>rank</td>
<td>RMSE</td>
<td>rank</td>
<td>RMSE</td>
</tr>
<tr>
<td><strong>1-step</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRR-Level</td>
<td>50</td>
<td>76.5</td>
<td>50</td>
<td>76.5</td>
<td>50</td>
<td><strong>67.1</strong></td>
</tr>
<tr>
<td>RRR-Diff</td>
<td>51</td>
<td><strong>73.7</strong></td>
<td>51</td>
<td><strong>73.7</strong></td>
<td>51</td>
<td>73.7</td>
</tr>
<tr>
<td>PC-level</td>
<td>50</td>
<td>80.9</td>
<td>50</td>
<td>80.9</td>
<td>50</td>
<td>72.7</td>
</tr>
<tr>
<td>PC-Diff</td>
<td>37</td>
<td>73.7</td>
<td>47</td>
<td>73.7</td>
<td>37</td>
<td>73.7</td>
</tr>
<tr>
<td><strong>24-step</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRR-Level</td>
<td>44</td>
<td>216.1</td>
<td>44</td>
<td>216.1</td>
<td>44</td>
<td>215.5</td>
</tr>
<tr>
<td>RRR-SDiff</td>
<td>51</td>
<td>198.7</td>
<td>51</td>
<td>198.7</td>
<td>51</td>
<td>198.7</td>
</tr>
<tr>
<td>PC-level</td>
<td>41</td>
<td><strong>191.1</strong></td>
<td>41</td>
<td><strong>191.1</strong></td>
<td>41</td>
<td><strong>188.8</strong></td>
</tr>
<tr>
<td>PC-Diff</td>
<td>51</td>
<td>198.8</td>
<td>51</td>
<td>198.8</td>
<td>51</td>
<td>198.8</td>
</tr>
</tbody>
</table>
RMSE w/r to Time and Price

RMSE × Time

RRR−Level 1−step  PC−Level 24−step

RMSE × Price

RRR−Level 1−step × Price  PC−Level 24−step × Price

Pelagatti (Bicocca U.)
Prediction example

Supply function on Wed 14.01.2009 at 10

Actual Level 24-step × Price
RRR–Level 1-step

Pelagatti (Bicocca U.)
Residual demand forecast

22 November 2011
### RMSE by day of the week

<table>
<thead>
<tr>
<th></th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-step</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRR-level</td>
<td>75.9</td>
<td>62.8</td>
<td>71.6</td>
<td>72.7</td>
<td>68.5</td>
<td>61.3</td>
<td>54.0</td>
</tr>
<tr>
<td>RRR-diff</td>
<td>80.1</td>
<td>64.8</td>
<td>83.9</td>
<td>81.6</td>
<td>72.7</td>
<td>72.9</td>
<td><strong>53.9</strong></td>
</tr>
<tr>
<td>PC-level</td>
<td>90.9</td>
<td>77.8</td>
<td>74.2</td>
<td>78.0</td>
<td>71.8</td>
<td>65.4</td>
<td>56.3</td>
</tr>
<tr>
<td>PC-diff</td>
<td>80.1</td>
<td>64.8</td>
<td>83.9</td>
<td>81.6</td>
<td>72.7</td>
<td>72.9</td>
<td><strong>53.9</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>24-step</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 24-step</td>
<td>241.4</td>
<td>250.6</td>
<td><strong>177.7</strong></td>
<td>243.2</td>
<td>249.6</td>
<td>158.2</td>
<td>155.4</td>
</tr>
<tr>
<td>Diff 24-step</td>
<td>275.2</td>
<td><strong>146.3</strong></td>
<td>198.3</td>
<td><strong>171.4</strong></td>
<td><strong>144.0</strong></td>
<td>240.1</td>
<td>186.8</td>
</tr>
<tr>
<td>PC-level</td>
<td><strong>235.9</strong></td>
<td>219.3</td>
<td>179.0</td>
<td>199.5</td>
<td>181.9</td>
<td><strong>151.3</strong></td>
<td><strong>151.4</strong></td>
</tr>
<tr>
<td>PC-diff</td>
<td>275.3</td>
<td>146.4</td>
<td>198.3</td>
<td><strong>171.4</strong></td>
<td>144.0</td>
<td>240.1</td>
<td>186.9</td>
</tr>
</tbody>
</table>
Simple, robust and effective: can be easily automatized (no numerical optimisation needed).

Further regressors as gas/oil price, temperature (fcst), weather (fcst) etc. may improve forecasts (we have done this, slightly better fcst).

Data are published with one week delay. We tried also 8-days-ahead forecasts: RMSE are some 70% higher.

Different time series models can be tried in the PC approach.

The optimal bidding strategy should be updated to include uncertainty in the residual demand.

Profit effectiveness should be checked by comparing actual profits of some firms with their potential profits using residual demand forecast.


