Forecasts of actual imbalance unit costs and simulated 5 minute prices for the two Danish Nordpool Spot price areas

April 17, 2013

ENFOR

Forecasting and Optimization for the Energy Sector

ENFOR A/S
Lyngsø Allé 3
DK-2970 Hørsholm
www.enfor.dk

Document ID: 08EKS0004A004-A

Contents

1 Introduction 6

2 Motivating imbalance unit cost forecasting 6

3 Data 10

4 Forecast method 10

5 Evaluation of results 10

5.1 Evaluation based on original imbalance unit costs 11

5.2 Evaluation of price forecasts based on simulated 5 minute price signals 16

6 Artificial forecast data for system simulation purposes 26

7 Data sets containing forecasts of 5 minute prices 26

8 Operational setup 28

9 Conclusion and discussion 29

A Alternative derivation of the optimal bid 30

B Evaluation of results individually for each year 31

C Evaluation of day ahead horizons 37

D Example R-script for exporting data 43

References **45**

Summary

This report considers price forecasts for system simulation in the FlexPower project. The report contains a general motivation for imbalance unit cost forecasting, evaluates such forecasts, and address evaluation of forecasts of simulated 5 minute prices. The price forecasts are evaluated separately for each year in the 11-year period 2001-2011 for the two Danish Nordpool spot price areas separately. When using the price forecasts in system simulations the forecast performance results presented here can be used when selecting the simulation period. Furthermore, for system simulation, the influence of forecast quality on simulation results might be of interest. A simple approach whereby this can be obtained is suggested. Furthermore, the operational setup used for demonstration purposes is briefly described.

1 Introduction

This report is prepared as part of the ForskEl-project *FlexPower* (2009-1-10486) which focus on providing flexibility to the power system by utilizing one-way price signals updated every 5 minutes [1]. Since such 5 minute price signals does not exists they have been simulated so that the actual one hour prices are followed. See ENFOR/08EKS0004A003-A[3] for further details.

Section 2 contains a motivation for price forecasting given the current market structure and considers aspects of the future electricity price which are important w.r.t. decision making. Section 3 briefly describes the data used and the forecast method is outlined in Section 4. Evaluation of forecast performance is presented in Section 5, both for hourly values and for the simulated 5 minute prices. Section 6 describes a method by which the forecasts can be artificially improved for system simulation purposes. The actual data sets containing the forecasts of simulated 5 minute prices are described in Section 7. Section 8 gives a brief outline of the operational setup which is linked to the operational setup described in ENFOR/08EKS0004A003-A[3]. Finally, Section 9 contain concluding remarks.

2 Motivating imbalance unit cost forecasting

When trading energy production in the Nordic day ahead market (Nordpool Spot) the revenue R_t for a specific hour t can be expressed as

$$R_t = P_{S,t}B_t + P_{D,t}(A_t - B_t)I(A_t > B_t) + P_{U,t}(A_t - B_t)I(A_t < B_t)\,, \tag{1}$$

where $P_{S,t}$, $P_{D,t}$, $P_{U,t}$ are the spot price and down- and up-regulation prices, B_t is the bid and A_t is the actual production, $I(\cdot)$ is the indicator function taken the value 1 if its argument is true and 0 otherwise.

The above neglects the intra day market *Elbas* because it historically has been consider rather illiquid and hence bidding on the spot market should not consider this intra day market. However, after bidding on the spot market Elbas should be considered for further reducing the imbalance costs.

E.g. for a wind power producer, at the time where the value of the bid B_t is decided none of the remaining quantities on the right hand side of (1) is know. Instead they are random variables which, to some extend, can be forecast. For this reason R_t is also a random variable for which the properties depends on the decided bid, which is a decision variable. Relevant properties of R_t includes the expected value and the variance, i.e. the expected revenue for hour t and the financial risk for hour t.

However, the situation as outlined above is considering a single hour only and relevant expected revenues and risks are related to markedly longer periods, probably quarters or years. Also, the financial risk of the individual hours in such a period will to a large extend cancel. Furthermore, e.g. for wind power the financial risk might even be dominated by external factors such as annual fluctuations in wind speeds and long term variations in the overall level of energy prices, none of which can be influenced by the bid decisions.

For this reason, and because revenues are fully separated between hours, it is argued that the bid B_t should be selected such that the expected revenue given information available at decision time $\mathcal{X}_{t_0}$, i.e. $E[R_t|\mathcal{X}_{t_0}]$, is maximized. Before implementing a particular bid the decision maker should then ensure that the risk associated with the particular hour is not so large that it can seriously affect the risk for longer periods.

Using (1) the expected revenue for hour t given information $\mathcal{X}_{t_0}$ available at decision time t_0 can be expressed as

$$\begin{aligned} E[R_t|\mathcal{X}_{t_0}] =& E[P_{S,t}|\mathcal{X}_{t_0}]B_t \\ &+E[P_{D,t}(A_t - B_t)I(A_t > B_t)|\mathcal{X}_{t_0}] \\ &+E[P_{U,t}(A_t - B_t)I(A_t < B_t)|\mathcal{X}_{t_0}] \end{aligned} \tag{2}$$

Using a first order approximation for the last two terms the expected revenue can be written

$$\begin{aligned} E[R_t|\mathcal{X}_{t_0}] =& E[P_{S,t}|\mathcal{X}_{t_0}]B_t \\ &+E[P_{D,t}|\mathcal{X}_{t_0}]E[(A_t - B_t)I(A_t > B_t)|\mathcal{X}_{t_0}] \\ &+E[P_{U,t}|\mathcal{X}_{t_0}]E[(A_t - B_t)I(A_t < B_t)|\mathcal{X}_{t_0}] \end{aligned} \tag{3}$$

As a further refinement a 2nd order approximation could be used. This would result in the terms $Cov[P_{D,t}|\mathcal{X}_{t_0}, (A_t - B_t)I(A_t > B_t)|\mathcal{X}_{t_0}]$ and $Cov[P_{U,t}|\mathcal{X}_{t_0}, (A_t - B_t)I(A_t < B_t)|\mathcal{X}_{t_0}]$ being added to the right hand side of (3).

The actual production A_t takes values between 0 and $\bar{A}$ and it follows that

$$E[(A_t - B_t)I(A_t > B_t)|\mathcal{X}_{t_0}] = \int_{-\infty}^{\infty} (x - B_t)I(x > B_t)f_{A_t|\mathcal{X}_{t_0}}(x)dx = \int_{B_t}^{\bar{A}} (x - B_t)f_{A_t|\mathcal{X}_{t_0}}(x)dx \tag{4}$$

and

$$E[(A_t - B_t)I(A_t < B_t)|\mathcal{X}_{t_0}] = \int_{-\infty}^{\infty} (x - B_t)I(x < B_t)f_{A_t|\mathcal{X}_{t_0}}(x)dx = \int_{0}^{B_t} (x - B_t)f_{A_t|\mathcal{X}_{t_0}}(x)dx \tag{5}$$

where $f_{A_t|\mathcal{X}_{t_0}}$ is the probability density function of $A_t|\mathcal{X}_{t_0}$, i.e. the actual production given information available at decision time.

In order to find the bid maximizing the expected revenue we seek the value of B_t for which $\partial E[R_t|\mathcal{X}_{t_0}]/\partial B_t = 0$. Hence, the partial derivatives of (4) and (5) w.r.t. B_t are needed. Using Leibniz rule for differentiation under the integral sign [4] the following is obtained:

$$\frac{\partial}{\partial B_t} \int_{B_t}^{\bar{A}} (x - B_t)f_{A_t|\mathcal{X}_{t_0}}(x)dx = -\int_{B_t}^{\bar{A}} f_{A_t|\mathcal{X}_{t_0}}(x)dx = -(1 - F_{A_t|\mathcal{X}_{t_0}}(B_t)) \tag{6}$$

and

$$\frac{\partial}{\partial B_t}\int_0^{B_t}(x-B_t)f_{A_t|\mathcal{X}_{t_0}}(x)dx = -\int_0^{B_t} f_{A_t|\mathcal{X}_{t_0}}(x)dx = -F_{A_t|\mathcal{X}_{t_0}}(B_t)\,, \qquad (7)$$

where $F_{A_t|\mathcal{X}_{t_0}}$ is the cumulative density function of $A_t|\mathcal{X}_{t_0}$. This may be obtained as quantile forecasts of the production.

Furthermore, if the bid does not affect the expected spot price

$$\frac{\partial}{\partial B_t}E[P_{S,t}|\mathcal{X}_{t_0}]B_t = E[P_{S,t}|\mathcal{X}_{t_0}] \qquad (8)$$

whereas in the general case

$$\frac{\partial}{\partial B_t}E[P_{S,t}|\mathcal{X}_{t_0}]B_t = E[P_{S,t}|\mathcal{X}_{t_0}] + \frac{\partial E[P_{S,t}|\mathcal{X}_{t_0}]}{\partial B_t}B_t \qquad (9)$$

In the general case the producer must know how the bids are influencing the spot price. In principle, both $E[P_{S,t}|\mathcal{X}_{t_0}]$ and $\partial E[P_{S,t}|\mathcal{X}_{t_0}]/\partial B_t$ are functions of B_t. As a simple solution it is assumed that for realistic values of the bid the partial derivative is constant (non-positive) and hence

$$E[P_{S,t}|\mathcal{X}_{t_0}] = E[P_{S,t}^{(0)}|\mathcal{X}_{t_0}] - a_p(B_t - E[A_t|\mathcal{X}_{t_0}])\,,$$

where $a_p = -\partial E[P_{S,t}|\mathcal{X}_{t_0}]/\partial B_t$ and $P_{S,t}^{(0)}$ is the spot price when bidding the expected production. Inserting into (9) yields

$$\frac{\partial}{\partial B_t}E[P_{S,t}|\mathcal{X}_{t_0}]B_t = E[P_{S,t}^{(0)}|\mathcal{X}_{t_0}] + a_pE[A_t|\mathcal{X}_{t_0}] - 2a_pB_t \qquad (10)$$

Inserting (6), (7), and (10) into the partial derivative of (3) w.r.t. B_t yields

$$\begin{aligned}\frac{\partial}{\partial B_t}E[R_t|\mathcal{X}_{t_0}] =& a_pE[A_t|\mathcal{X}_{t_0}] + E[P_{S,t}^{(0)}|\mathcal{X}_{t_0}] - E[P_{D,t}|\mathcal{X}_{t_0}] \\ & -2a_pB_t - (E[P_{U,t}|\mathcal{X}_{t_0}] - E[P_{D,t}|\mathcal{X}_{t_0}])F_{A_t|\mathcal{X}_{t_0}}(B_t)\end{aligned} \qquad (11)$$

where it is further assumed that the expected regulating prices ($E[P_{D,t}|\mathcal{X}_{t_0}]$ and $E[P_{U,t}|\mathcal{X}_{t_0}]$) are not influenced by the bid.

The overall price level can be eliminated from (11) by defining the imbalance unit costs

$$C_{D,t} = P_{S,t}^{(0)} - P_{D,t} \quad \text{and} \quad C_{U,t} = P_{U,t} - P_{S,t}^{(0)} \qquad (12)$$

whereby (11) can be written

$$\frac{\partial}{\partial B_t}E[R_t|\mathcal{X}_{t_0}] = a_pE[A_t|\mathcal{X}_{t_0}] + E[C_{D,t}|\mathcal{X}_{t_0}] - 2a_pB_t - (E[C_{D,t}|\mathcal{X}_{t_0}] + E[C_{U,t}|\mathcal{X}_{t_0}])F_{A_t|\mathcal{X}_{t_0}}(B_t) \quad (13)$$

Hence, the value of B_t for which the partial derivative is zero, i.e. the optimal bid, fulfills

$$2a_pB_t + (E[C_{D,t}|\mathcal{X}_{t_0}] + E[C_{U,t}|\mathcal{X}_{t_0}])F_{A_t|\mathcal{X}_{t_0}}(B_t) = a_pE[A_t|\mathcal{X}_{t_0}] + E[C_{D,t}|\mathcal{X}_{t_0}] \qquad (14)$$

Given knowledge of a_p, forecasts of the imbalance unit costs, and quantile forecasts of the actual production (14) can be solved w.r.t. B_t in order to find the optimal bid. For the special case where $a_p = 0$ (price taker) the solution can be written

$$F_{A_t|\mathcal{X}_{t_0}}(B_t) = \frac{E[C_{D,t}|\mathcal{X}_{t_0}]}{E[C_{D,t}|\mathcal{X}_{t_0}] + E[C_{U,t}|\mathcal{X}_{t_0}]} \tag{15}$$

i.e. the optimal bid is a quantile in the conditional distribution of the actual production. E.g. if the right hand side of (15) is 0.4 we look up the 40% quantile forecast and bid this on the spot marked.

Above it is assumed that the regulating prices are not affected by the bid. In Appendix A a deviation based on the assumption that the imbalance unit costs are not affected by the bid is presented and it is argued that this assumption might actually be more appropriate. This leads to a slightly different (simpler) solution for the general case, but for a price taker the solutions are identical and in the general case, for realistic values of the bid, the solutions are approximately equal.

In order to solve either (14) or (15) it is seen that we need the expected values of the differences between each of the regulating prices and the spot price, here called the imbalance unit costs. Even though in principle the "unaffected" spot price should be used, presumably the imbalance unit costs are dominated by the difference between the spot- and regulating-prices so that when modelling the conditionally expected values of the imbalance unit costs these can be calculated based on actually observed spot prices. The solution of the equation leads to a probability $F_{A_t|\mathcal{X}_{t_0}}(B_t)$, converting this to a bid requires a quantile forecast of the production.

A price taker is characterized by $a_p = 0$. In the general case the producer must know how much increasing the bid over the expected production reduce the spot price. This is expressed in the value of a_p which has the unit DKK/MWh or similar. A linear approximation to (9) is used which result in (10). A non-linear function can be handled without further complications. However, the producer must be able to specify these nonlinearities.

It is noted that the expected spot price can be eliminated from the equation defining the optimal bid on the spot market. Basically, this is because the overall level of the prices is beyond the control of the bidding process which therefore can only strive to select the bid such that the imbalance penalty is balanced optimally against the expected influence of the bid on the spot price. Hence, expected imbalance unit costs, and not expected spot prices, are of main interest when seeking optimal bids. It should be noted that since the imbalance unit costs (12) are non-negative the conditionally expected values of these are both positive for most time points, despite the fact that for the actual imbalance costs at most one can be positive for any given time point. In conclusion, the conditionally expected values, i.e. the forecasts, should not be expected to behave as the actual costs.

3 Data

The data used here consists of actual hourly data for the two Danish Nordpool price areas covering the period 2000-10-01 to 2011-12-31. Furthermore, simulated 5 minute data is used. See ENFOR/08EKS0004A003-A[3] for further details.

4 Forecast method

For practical reasons since the forecasts are to be constructed for a long historic period and used for simulation purposes it has been chosen to generate forecasts based on auto-regressive models, i.e. models not applying external information. With this model type the future value is forecast using historically observed values only. The models are estimated adaptively and recursively, i.e. the underlying estimates are updated as new observations arrive and the influence of old observations are gradually removed. In this way the coefficients of the models adapts to the actual situation, a property which presumably is important for potentially non-stationary processes as electricity prices.

The models are applied in order to forecast the imbalance unit costs with focus on providing expected values conditional on information available at forecast time. As an example consider an AR(1) model applied for 1-step predictions. The model can be written

$$x_{t+1} = \phi_0 + \phi_1 x_t + e_{t+1} \tag{16}$$

where x_t is either the down- or up-regulation imbalance unit cost at time t, $\phi_.$ are coefficients to be estimated from data and e_t represents a zero-mean, but otherwise unpredictable, noise term. The estimates adaptively updated and available at time t are denoted $\hat{\phi}_{0t}$ and $\hat{\phi}_{1t}$ and hence the 1-step forecast available at time t can be written

$$\hat{x}_{t+1|t} = \hat{\phi}_{0t} + \hat{\phi}_{1t} x_t \tag{17}$$

Which given (16) equals the expected value of x_{t+1} given the information available at time t, which in this case is represented by the updated coefficients and x_t.

5 Evaluation of results

The forecast method is applied to both the actual hourly imbalance unit costs and the simulated 5 minute imbalance unit costs based on the simulated 5 minute prices as described in ENFOR/08EKS0004A003-A[3]. In both cases a forgetting factor corresponding to 6 weeks is applied. For the hourly data other forgetting factors has been investigated as well.

As seen from e.g. (14) and (15) in Section 2 focus should be on evaluating the quality of the forecasts as estimates of the conditional expectations occurring in these equations. This is the focus in Section 5.1 below.

In the FlexPower simulation work package forecasts of the 5 minute prices for horizons up to 12 hours assuming the spot prices to be known must be produced. These are constructed based on forecasts of the 5 minute imbalance unit costs and known spot prices. In Section 5.2 these forecasts are evaluated using traditional forecast evaluation techniques.

5.1 Evaluation based on original imbalance unit costs

This section focus on evaluating the quality of the forecasts as estimates of the conditional expectations occurring in e.g. (14) and (15) in Section 2.

The forecast values are divided in ten groups of equal size using quantiles corresponding to 0%, 10%, ..., 90%, 100%. For each of these groups the mean of the forecast imbalance unit costs and the mean of the actual imbalance unit costs are calculated. Subsequently, these two sets of mean values are plotted against each other. Under the assumption that the forecast reproduce the actual mean values the ten points mentioned should, except for random variation, lie on the line of identity. The random variation may be addressed by calculating two times the standard error of the mean. However, due to serial correlation in the data this will under estimate the uncertainty of the estimated means of the actual prices. Nevertheless it is used as a simple indication of the uncertainty.

As mentioned in Section 3 the data covers the period from 2000-10-01 to 2011-12-31. However, in order to allow the recursive and adaptive estimation methods to be initialized, the three months in 2000 are not considered when evaluation the results.

Figure 1 show the evaluation results as described above for the period 2001–2011 when using a forgetting factor corresponding to 6 weeks. Figures 15–18 (pages 31–34) in Appendix B depicts the results when split by year.

It is seen that generally the observed means agrees well with the foretasted means. This is also true when considering the yearly results. However, for the yearly results the standard error of the observed means are quite large for some years, notably in the end of the period. Based on the overall evaluation it is seen that from horizons of 6 hours or longer the observed means lie below the line of identity indicating that the forgetting factor is too small whereby extra noise is introduced into the forecast unit penalties. Figure 2 show results obtained when using a forgetting factor corresponding to 24 weeks. Comparing to Figure 1 it is seen that, for horizons longer than 6 hours, the line segments practically touch the line of identity. Furthermore, the spread on the 1st axis is reduced. Both observations confirm the hypothesis posed above.

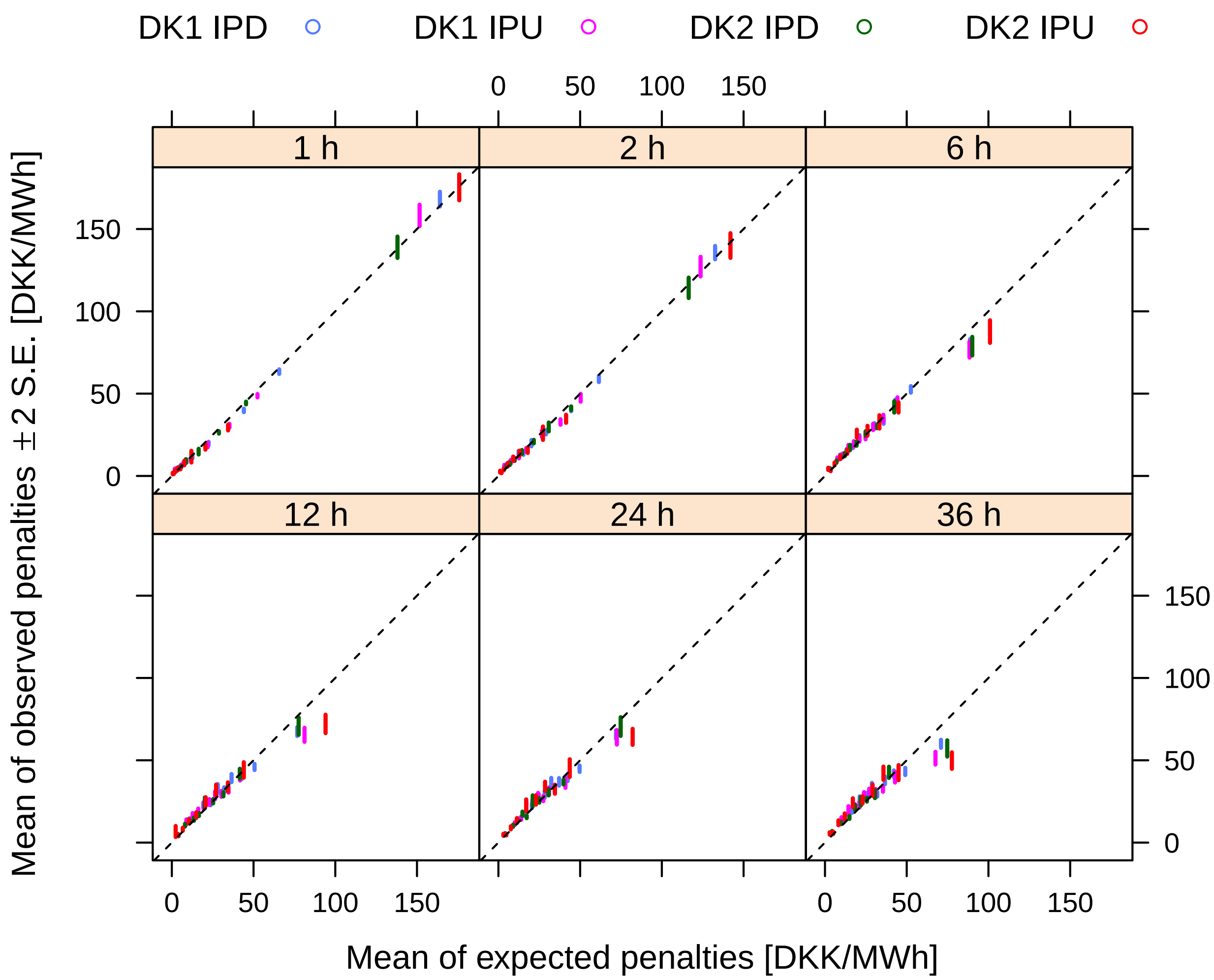

Figure 1: Imbalance unit cost; mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ±2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 6 weeks.

For a price taker the optimal bid is given by (15) for which the right hand side defines the probability which in turn defines the quantile in the conditional distribution of the future wind power production which should be bid on the spot marked. The probability is defined by the forecasts (expected means) of the imbalance unit costs. In order to evaluate these probabilities directly the data is again grouped in ten equally sized groups, here defined by the bid quantiles (actually probabilities). For each of these groups the observed means are calculated and the bid quantiles based on the observed means are found. The two sets of quantiles are plotted against each other and should lie near the line of identity.

The result of this evaluation is depicted in Figure 3 for a forgetting factor corresponding to 6 weeks and in Figure 4 for a forgetting factor corresponding to 24 weeks. It is seen that the evaluation results lie close to the line of identity, for longer horizons the forgetting factor

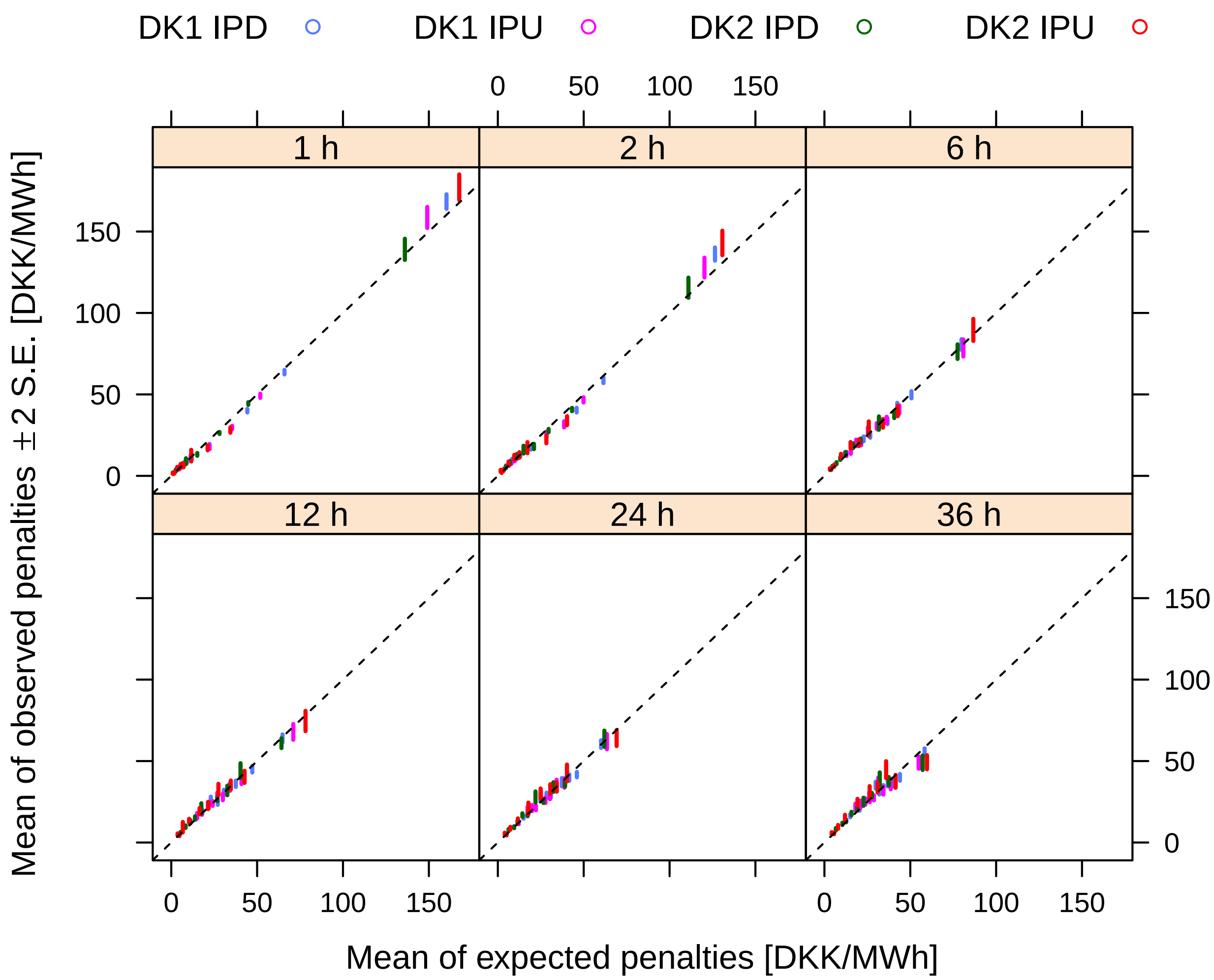

Figure 2: Imbalance unit cost; mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ±2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 24 weeks.

corresponding to 24 weeks is beneficial. Figures 19 and 20 (pages 35 and 36) show corresponding yearly evaluation results for a forgetting factor corresponding to 6 weeks.

Finally, for the evaluation of day ahead horizons the reader is referred to Appendix C.

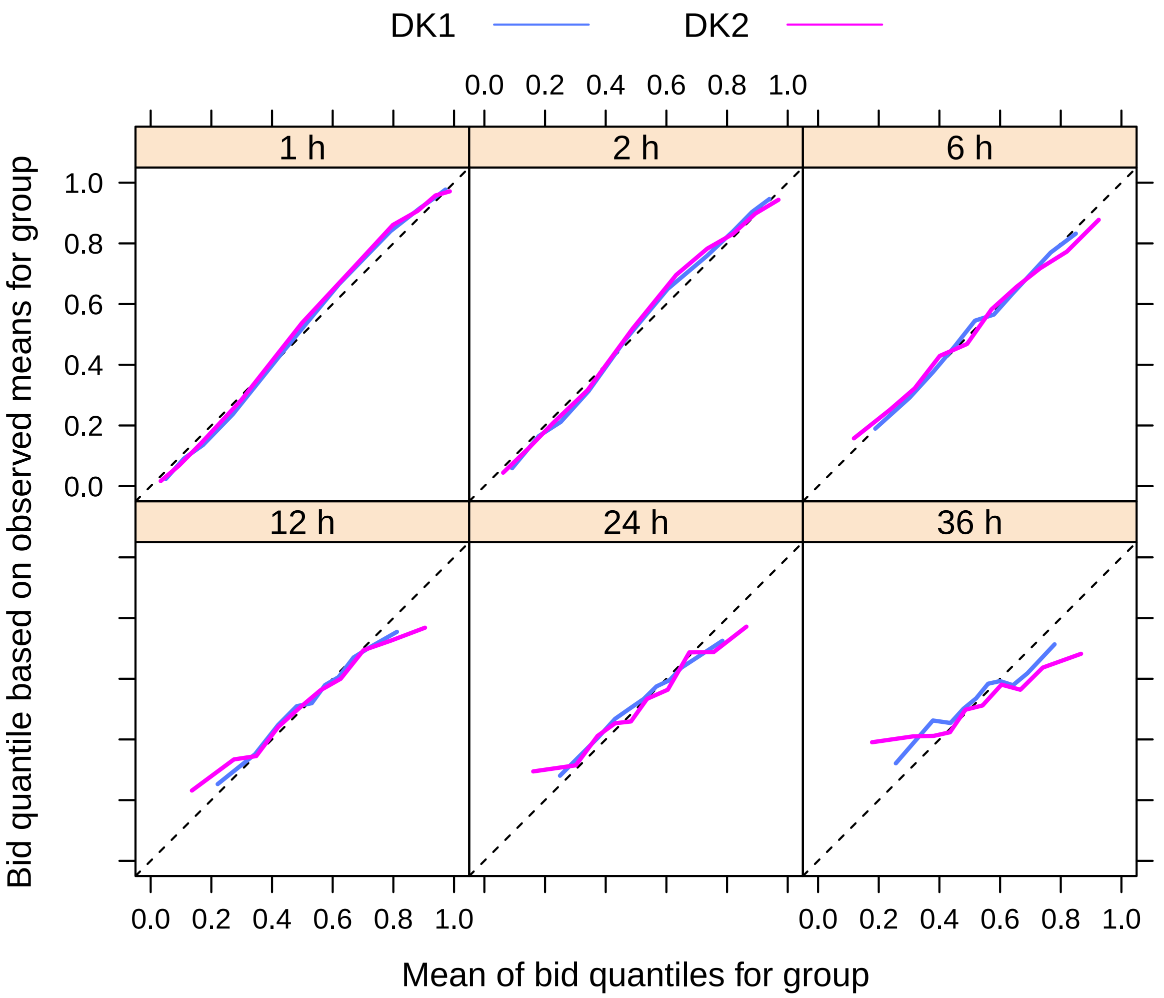

Figure 3: Probability of bid quantile based on observed means versus mean of probability of bid quantile based on expected (forecast) penalties. The results are obtained using a forgetting factor corresponding to 6 weeks.

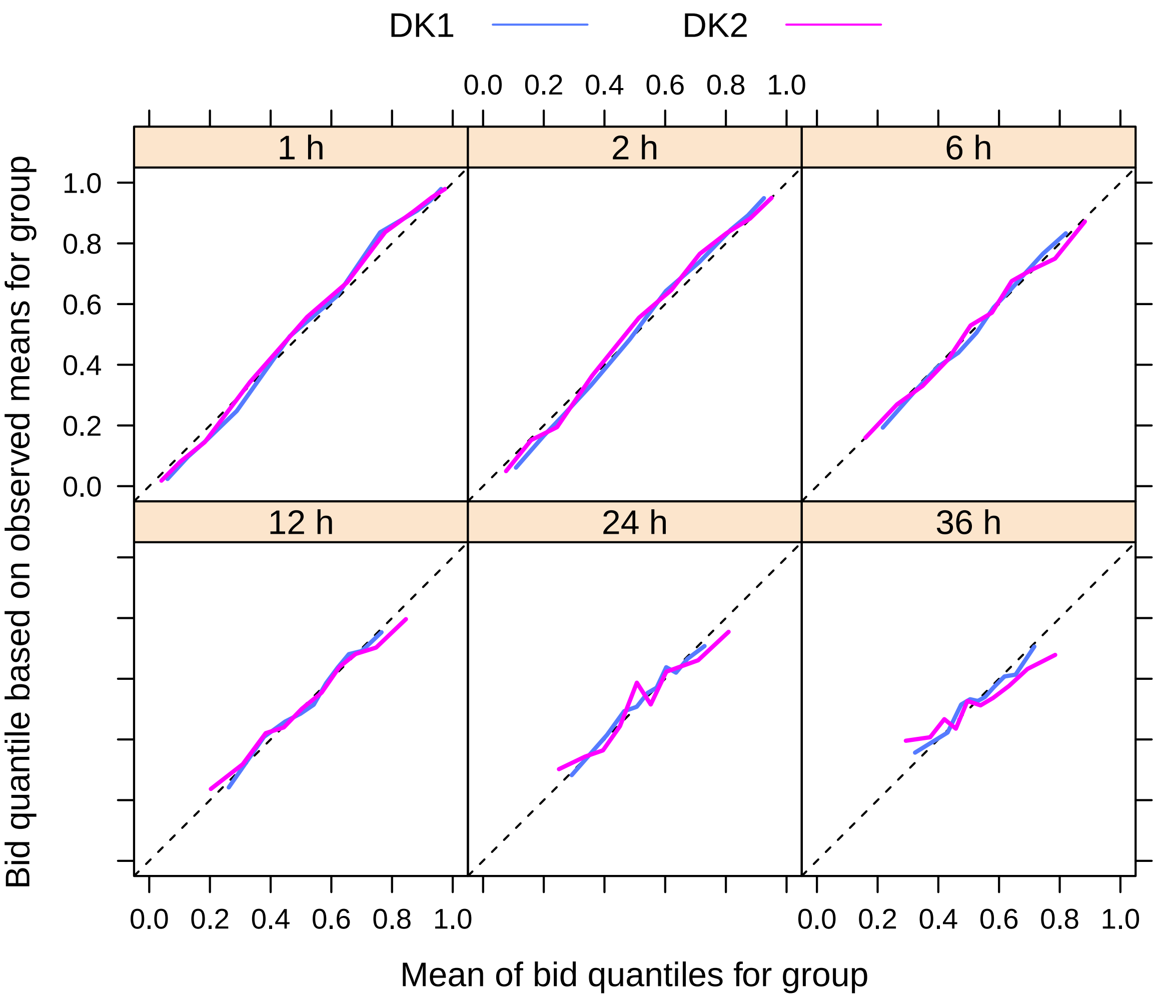

Figure 4: Probability of bid quantile based on observed means versus mean of probability of bid quantile based on expected (forecast) penalties. The results are obtained using a forgetting factor corresponding to 24 weeks.

5.2 Evaluation of price forecasts based on simulated 5 minute price signals

Section 5.1 considered evaluation of forecasts of hourly imbalance unit costs. For FlexPower forecasts of the actual 5 minute electricity price must be supplied for horizons up to 12 hours and hence the spot price is assumed to be known. Here the 5 minute forecasts of the simulated data described in ENFOR/08EKS0004A003-A[3] are considered for the case where a forgetting factor corresponding to 6 weeks is used.

The forecasts are produced by first forecasting the imbalance unit costs C_{Ut} and C_{Dt} as considered above and then constructing the forecast of the 5 minute price P_{t+k} as

$$\widehat{P}_{t+k|t} = P_{S,t+k} + \widehat{C}_{U,t+k|t} - \widehat{C}_{D,t+k|t}\,, \tag{18}$$

where $P_{S,t+k}$ is the spot price at time $t+k$ which is assumed known at time t.

Evaluation is performed using the standard forecast performance measures; bias, MAE (Mean Absolute Error), and RMSE (Root Mean Square Error) for each horizon separately. Evaluation is performed for each calendar year separately and the forecast errors are normalized with the average spot price for the particular calendar year. Figure 5 depicts the average spot prices for each price area and year.

Considering a specific forecast horizon, time indices $t = 1, \ldots, N$ corresponding to one year, and the corresponding average spot price $\bar{P}_S$. The normalized bias can be written

$$\frac{1}{N}\sum_{t=1}^{N} e_t/\bar{P}_S\,, \tag{19}$$

the normalized MAE can be written

$$\frac{1}{N}\sum_{t=1}^{N} |e_t/\bar{P}_S|\,, \tag{20}$$

and the normalized RMSE can be written

$$\sqrt{\frac{1}{N}\sum_{t=1}^{N} (e_t/\bar{P}_S)^2}\,, \tag{21}$$

where e_t is the forecast error.

The evaluation results are shown in Figures 6–8 for the DK1 price area and in Figures 9–11 for the DK2 price area. Please note that the axes are identical within performance measures across price areas, but not between performance measures. The measures are reported in percent, i.e. the numbers are those from (19)–(21) multiplied with 100%. Furthermore, Figures 12 and 13 show the cumulative error, the cumulative absolute error, and the cumulative square error for the forecasts based on one-hourly data as described in Section 5.1, using a forgetting factor corresponding to 6 weeks. One-hourly data are used here in order to reduce the number of points underlying these plots.

In general large variations between years occur. Overall the bias is quite small, generally within 1%. Both MAE and RMSE are measures of variability and it is seen that this variability of forecast errors increase rapidly within the first 2–4 hours, where after the variability stabilize. For RMSE / DK2 this is not exactly true for years 2006 and 2009. This is because the squared error in (21) for short periods of time contributes significantly to the error measure. This can be seen from the lower plot in Figure 13 which show significant contributions for 4 and 6 hour forecasts for a short period in the beginning of 2006 and in the end of 2009. In general the cumulative absolute error plots shown that no short periods of time contributes significantly to the MAE performance measure. Therefore, in general, the error in DKK/MWh is of the same magnitude for longer periods.

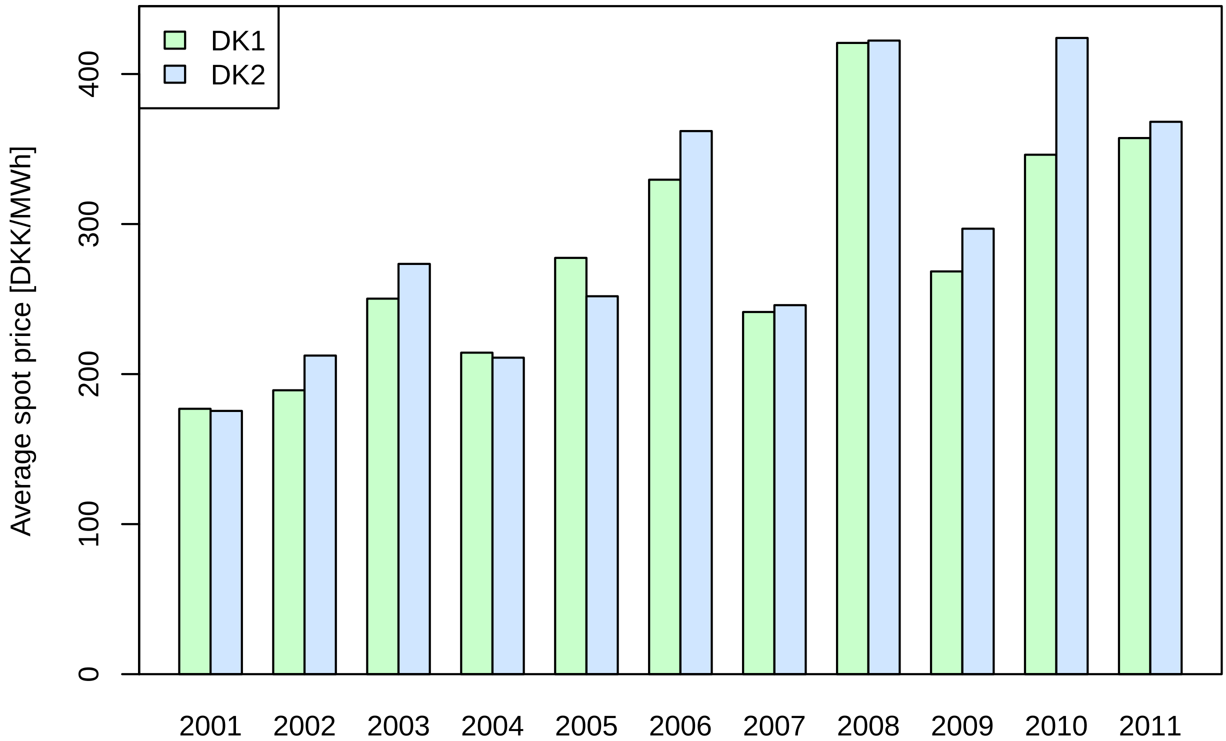

Figure 5: Yearly average spot price for Nordpool Spot price areas DK1 and DK2.

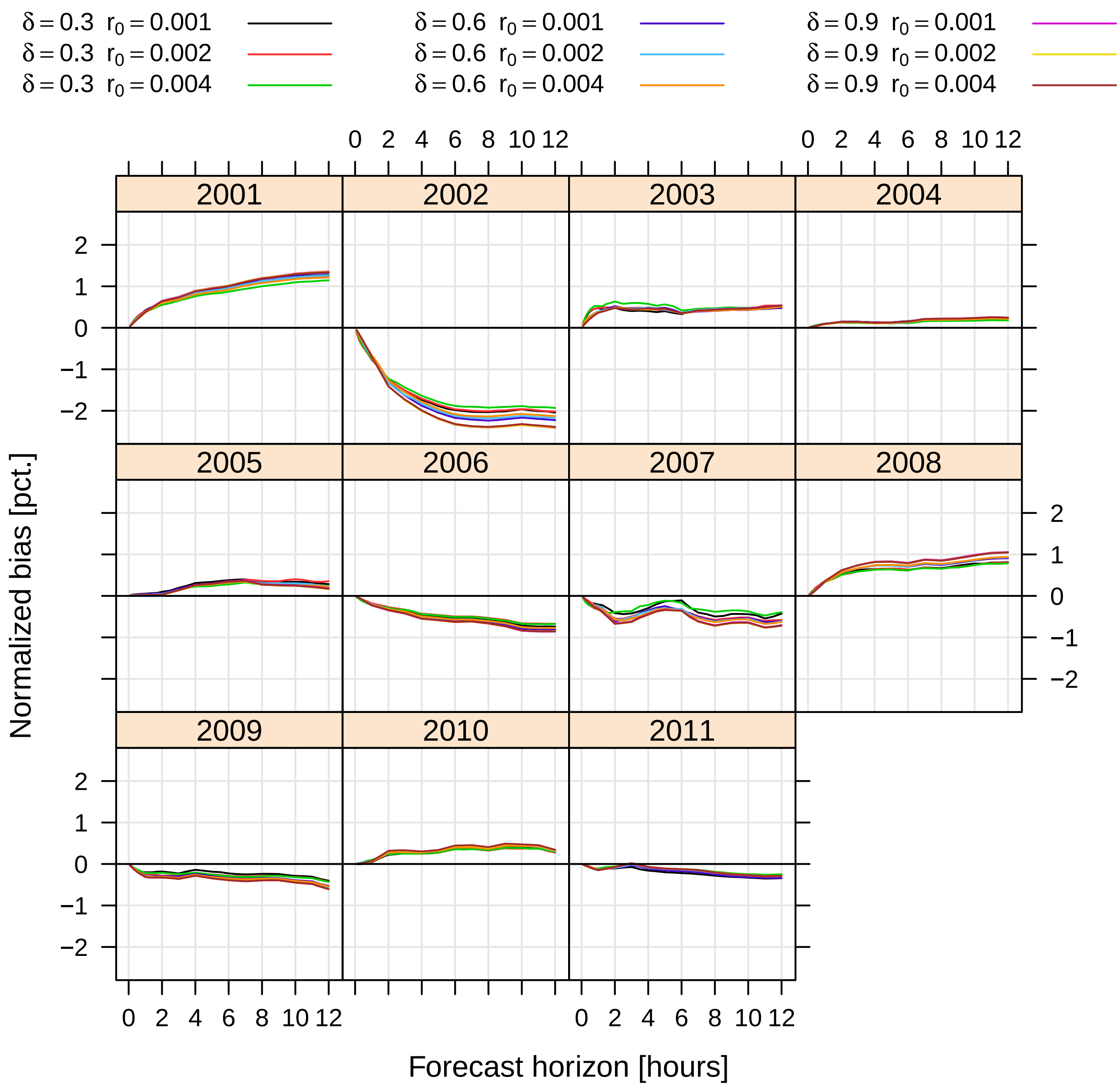

Figure 6: Price area DK1: Normalized bias against forecast horizon for each calendar year for the nine simulated data sets as indicated in the legend, cf. ENFOR/08EKS0004A003-A[3]. The normalization constant is the average spot price for the particular year.

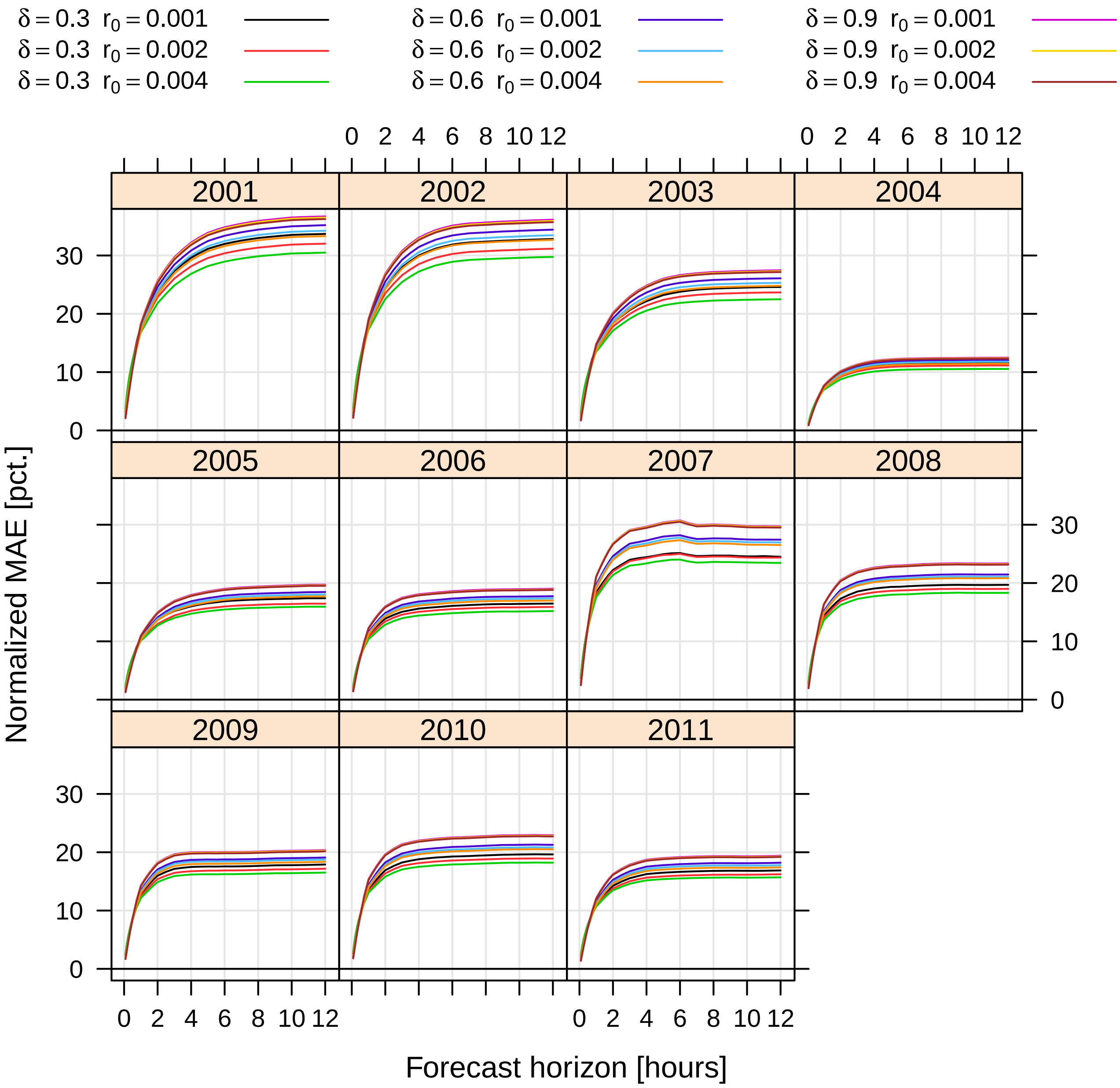

Figure 7: Price area DK1: Normalized MAE (Mean Absolute Error) against forecast horizon for each calendar year for the nine simulated data sets as indicated in the legend, cf. ENFOR/08EKS0004A003-A[3]. The normalization constant is the average spot price for the particular year.

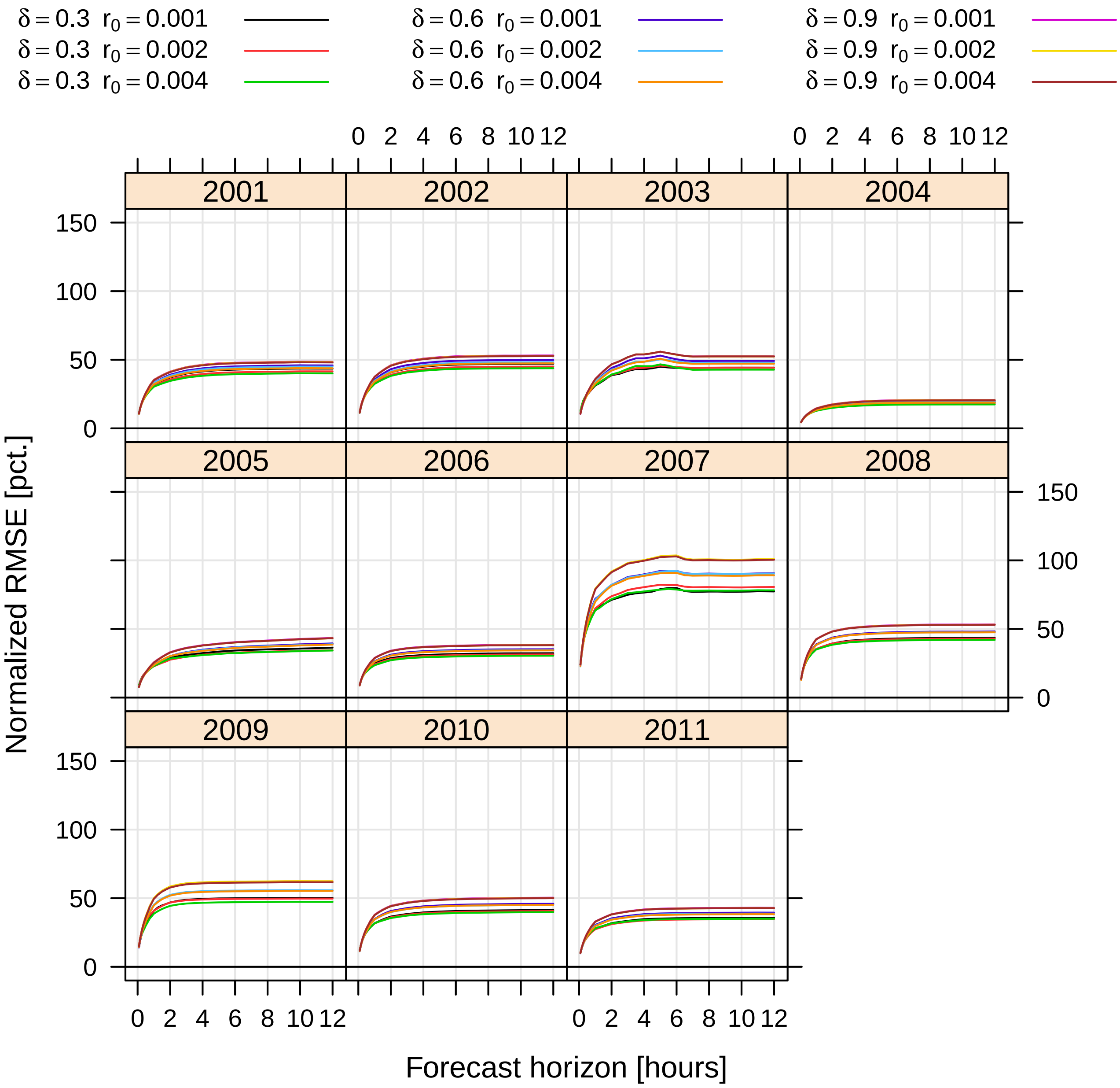

Figure 8: Price area DK1: Normalized RMSE (Root Mean Square Error) against forecast horizon for each calendar year for the nine simulated data sets as indicated in the legend, cf. ENFOR/08EKS0004A003-A[3]. The normalization constant is the average spot price for the particular year.

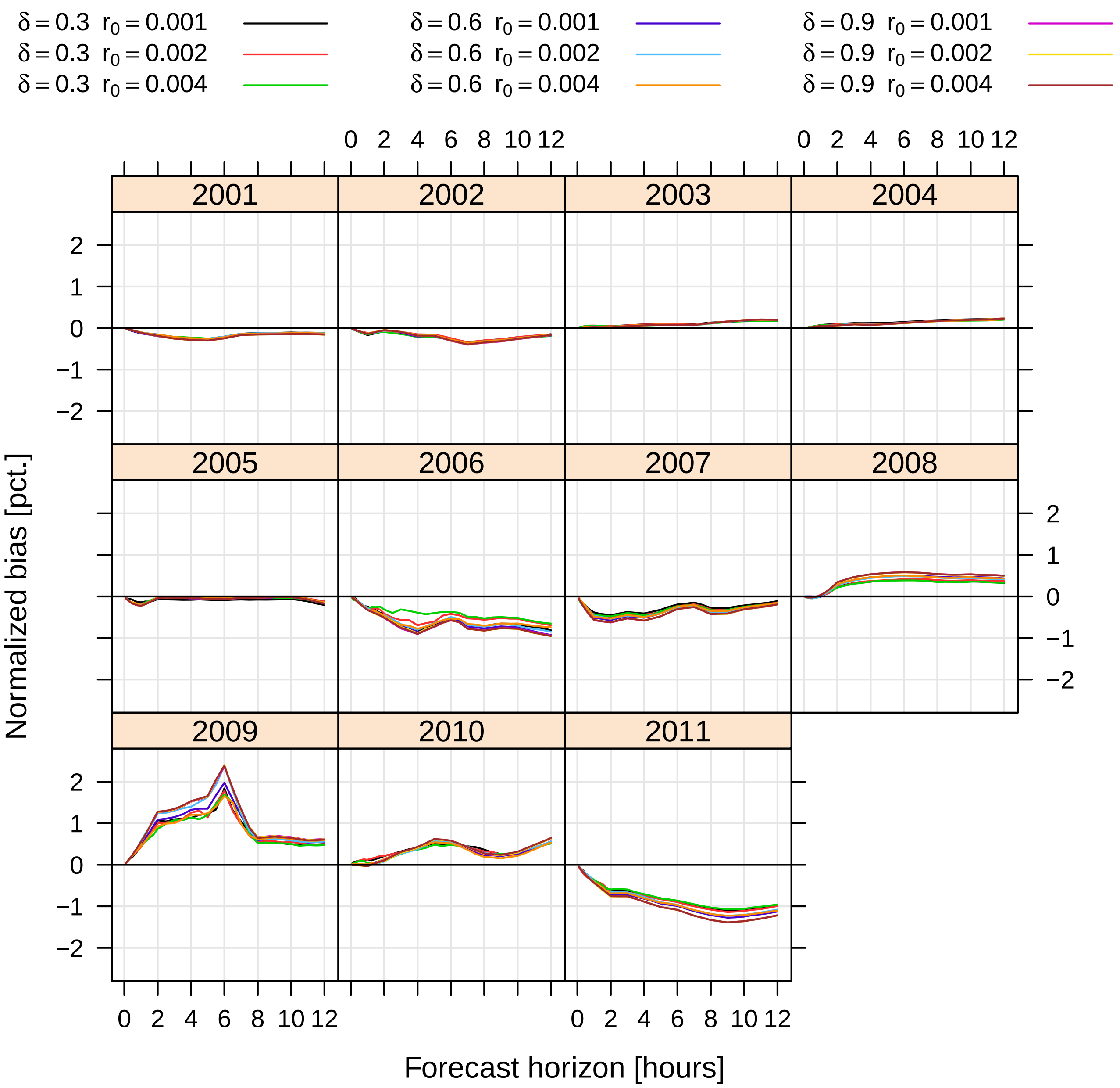

Figure 9: Price area DK2: Normalized bias against forecast horizon for each calendar year for the nine simulated data sets as indicated in the legend, cf. ENFOR/08EKS0004A003-A[3]. The normalization constant is the average spot price for the particular year.

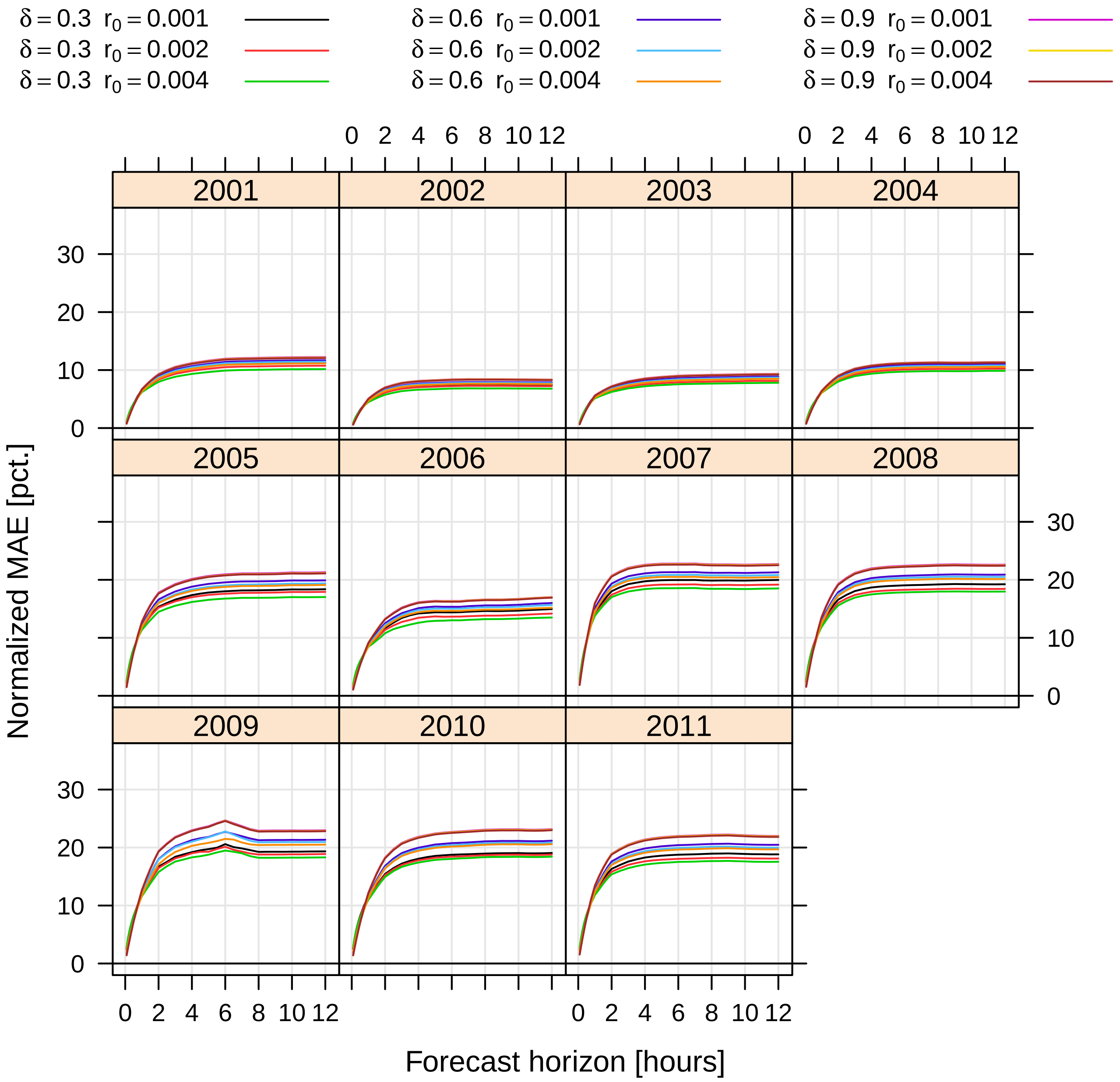

Figure 10: Price area DK2: Normalized MAE (Mean Absolute Error) against forecast horizon for each calendar year for the nine simulated data sets as indicated in the legend, cf. ENFOR/08EKS0004A003-A[3]. The normalization constant is the average spot price for the particular year.

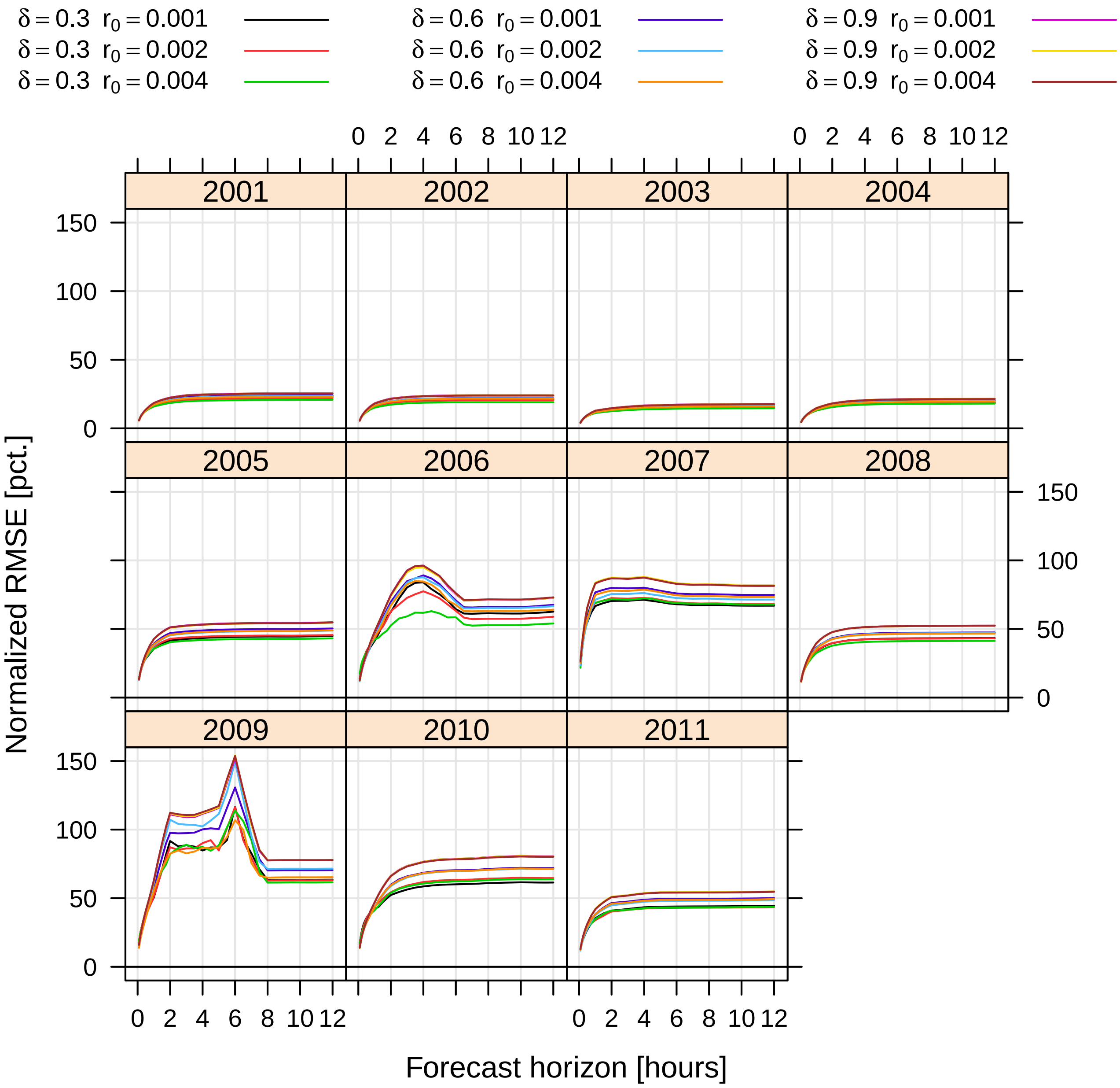

Figure 11: Price area DK2: Normalized RMSE (Root Mean Square Error) against forecast horizon for each calendar year for the nine simulated data sets as indicated in the legend, cf. ENFOR/08EKS0004A003-A[3]. The normalization constant is the average spot price for the particular year.

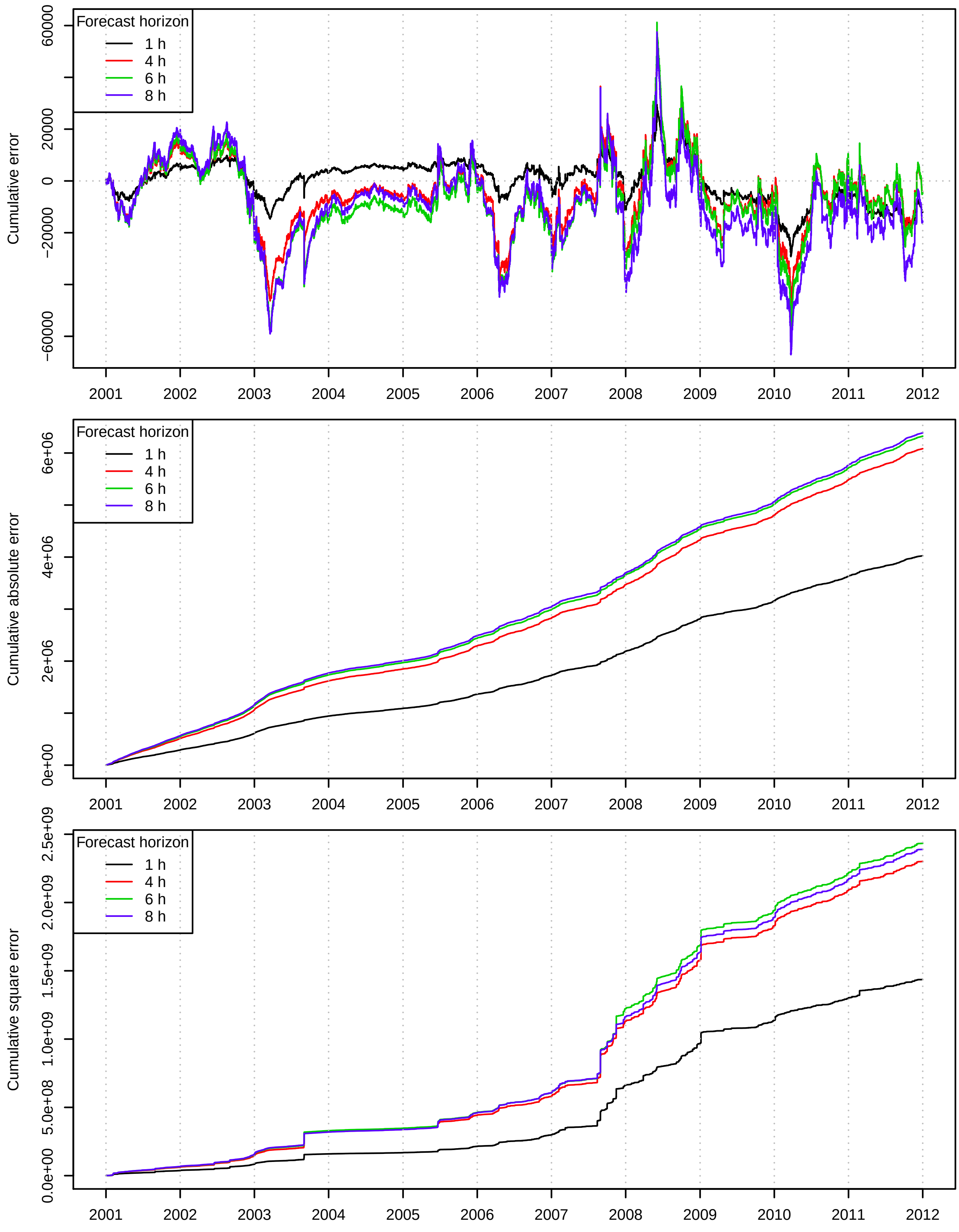

Figure 12: Price area DK1: Cumulative error sums for the one-hourly forecasts described in Section 5.1, when using a forgetting factor corresponding to 6 weeks.

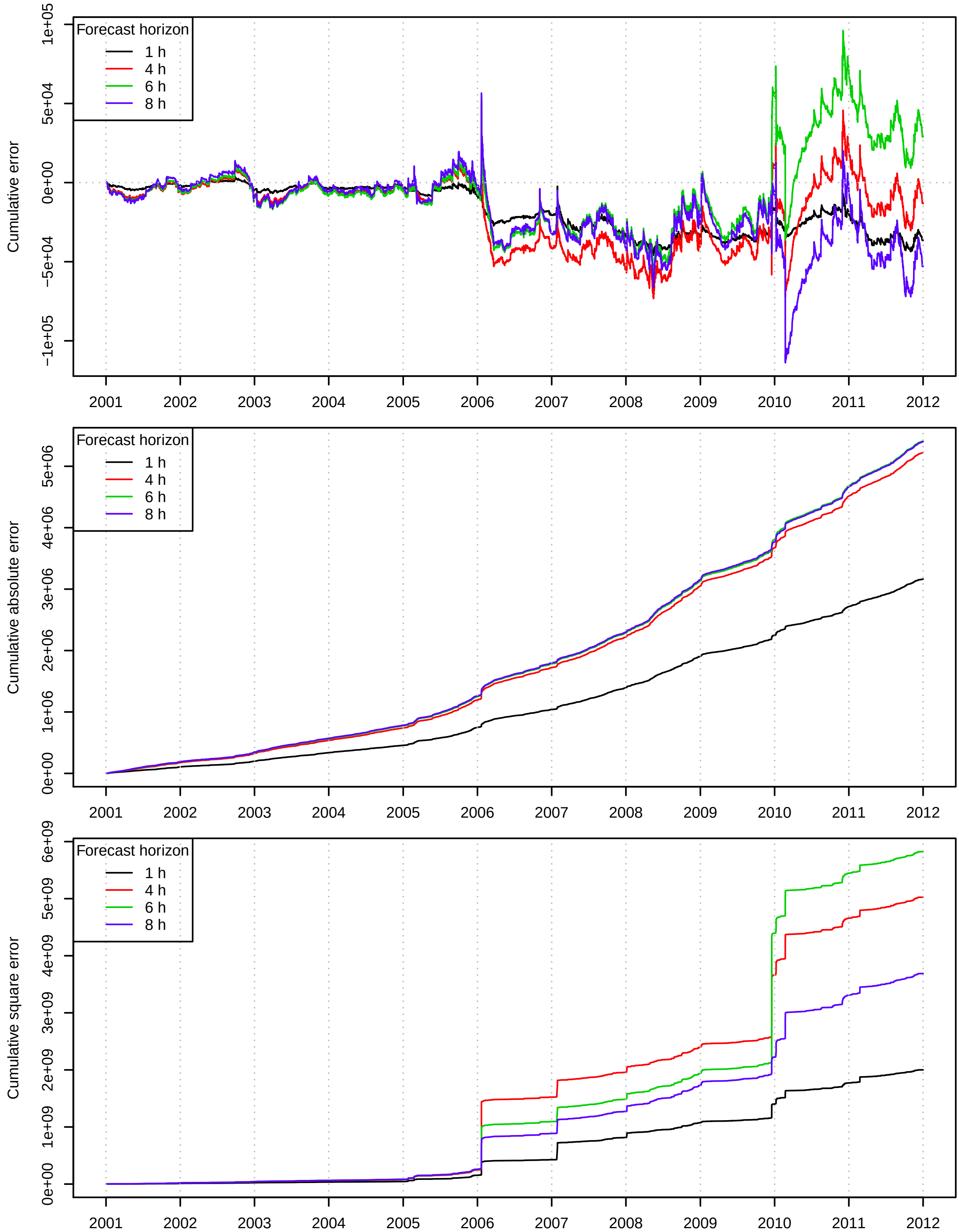

Figure 13: Price area DK2: Cumulative error sums for the one-hourly forecasts described in Section 5.1, when using a forgetting factor corresponding to 6 weeks.

6 Artificial forecast data for system simulation purposes

For system simulation purposes it might be relevant to address the influence of forecast quality on simulation results. A simple way to address this is to define a scalar $\alpha \in [0, 1]$, multiply the actual forecast error by this quantity and add the result to the forecast, resulting in the artificial forecasts:

$$\widehat{P}^{(art)}_{t|t-k} = \widehat{P}_{t|t-k} + \alpha(P_t - \widehat{P}_{t|t-k}) = \alpha P_t + (1-\alpha)\widehat{P}_{t|t-k} \, . \tag{22}$$

With this formulation $\alpha = 0$ corresponds to the raw forecasts ($\widehat{P}^{(art)}_{t|t-k} = \widehat{P}_{t|t-k}$) contained in the data sets and $\alpha = 1$ corresponds to perfect forecasts ($\widehat{P}^{(art)}_{t|t-k} = P_t$).

7 Data sets containing forecasts of 5 minute prices

The forecasts for the simulated data sets described in ENFOR/08EKS0004A003-A[3] as evaluated in Section 5.2 are provided as binary files which can be loaded into R (`www.r-project.org`). Table 1 show the available data sets. When loaded into R the data will be available as an object named `P5mFcs`, which is a list with elements `dk1` and `dk2` containing data and forecasts for the two Danish Nordpool Spot price areas.

δ	r_0	Name of data set
0.3	0.001	`P5mFcs,sdf=1,phi=0.8,delta=0.3,r0=0.001.rda`
0.3	0.002	`P5mFcs,sdf=1,phi=0.8,delta=0.3,r0=0.002.rda`
0.3	0.004	`P5mFcs,sdf=1,phi=0.8,delta=0.3,r0=0.004.rda`
0.6	0.001	`P5mFcs,sdf=1,phi=0.8,delta=0.6,r0=0.001.rda`
0.6	0.002	`P5mFcs,sdf=1,phi=0.8,delta=0.6,r0=0.002.rda`
0.6	0.004	`P5mFcs,sdf=1,phi=0.8,delta=0.6,r0=0.004.rda`
0.9	0.001	`P5mFcs,sdf=1,phi=0.8,delta=0.9,r0=0.001.rda`
0.9	0.002	`P5mFcs,sdf=1,phi=0.8,delta=0.9,r0=0.002.rda`
0.9	0.004	`P5mFcs,sdf=1,phi=0.8,delta=0.9,r0=0.004.rda`

Table 1: Data sets containing forecasts of 5 minute data using a forgetting factor corresponding to 6 weeks. See ENFOR/08EKS0004A003-A[3] for the definition of ϕ, δ, and r_0. Each file is approximately of size 500 MB.

Table 2 shows part of one element of a data set. The columns contains the following variables:

`Time` : End time point of 5 minute time interval.

`PS` : Spot price for the hour, which the 5 minute interval belongs to.

Time	PS	P1h	P5m	k=1	k=2	k=3	k=4
2011-12-01 01:05:00	219.29	282.41	270.8323	238.0249	238.2373	238.0712	237.9049
2011-12-01 01:10:00	219.29	282.41	270.8323	269.0925	238.2368	238.0706	237.9041
2011-12-01 01:15:00	219.29	282.41	270.8323	269.0933	267.2909	238.0699	237.9033
2011-12-01 01:20:00	219.29	282.41	270.8323	268.8553	266.8323	264.8234	237.0049
2011-12-01 01:25:00	219.29	282.41	270.8323	268.8564	266.8341	264.8344	262.8131
2011-12-01 01:30:00	219.29	282.41	270.8323	268.8576	266.8363	264.8454	262.8245
2011-12-01 01:35:00	219.29	282.41	270.8323	269.0999	267.3126	265.5446	263.7434
2011-12-01 01:40:00	219.29	282.41	270.8323	269.1007	267.3146	265.5475	263.7551
2011-12-01 01:45:00	219.29	282.41	270.8323	269.1015	267.3162	265.5505	263.7590
2011-12-01 01:50:00	219.29	282.41	270.8323	269.1023	267.3178	265.5529	263.7629
2011-12-01 01:55:00	219.29	282.41	282.4100	269.1031	267.3194	265.5552	263.7661
2011-12-01 02:00:00	219.29	282.41	282.4100	279.9314	267.3210	265.5576	263.7692

Table 2: Structure of each element of the R list `P5mFcs`*, here a subset for DK1 ($\delta = 0.3$ and $r_0 = 0.001$). Not all columns are shown, the full set of columns are:* `Time`*,* `PS`*,* `P1h`*,* `P5m`*,* `k=1`*,* `k=2`*,* `k=3`*,* `k=4`*,* `k=5`*,* `k=6`*,* `k=9`*,* `k=12`*,* `k=15`*,* `k=18`*,* `k=21`*,* `k=24`*,* `k=30`*,* `k=36`*,* `k=42`*,* `k=48`*,* `k=54`*,* `k=60`*,* `k=66`*,* `k=72`*,* `k=78`*,* `k=84`*,* `k=90`*,* `k=96`*,* `k=102`*,* `k=108`*,* `k=114`*,* `k=120`*,* `k=126`*,* `k=132`*,* `k=138`*,* `k=144`*.*

`P1h` : Final price for the hour, which the 5 minute interval belongs to (spot price plus up-regulation price minus down-regulation price).

`P5m` : Simulated price for the particular 5 minute time interval.

`k=1` : Forecast of `P5m` for time `Time` based on information up to time `Time` $- 1 \times 5$ minutes

`k=2` : Forecast of `P5m` for time `Time` based on information up to time `Time` $- 2 \times 5$ minutes

$\vdots$

`k=144` : Forecast of `P5m` for time `Time` based on information up to time `Time` $- 144 \times 5$ minutes

Note that in order to limit the file size not all 5 minute time steps between 1 and 144 are included in the data sets. In order to obtain forecasts for horizons not included interpolation must be used. When doing this interpolation all forecast available at a given time point should be extracted and the spot price (`PS`) subtracted. Hereafter, the interpolation in time is performed and the spot price is added again. Appendix D contains an example R-script performing this interpolation and exporting data to csv-format.

8 Operational setup

The operational 5 minute forecast setup is based on the operational 5 minute simulated data for the Eastern Danish price area (DK2) as described in ENFOR/08EKS0004A003-A[3]. As described in the report mentioned time stamps are shifted 6 hours in order to be able operationally to issue prices in the *beginning* of the 5 minute intervals. The 5 minute price forecasts apply the same time shift. Figure 14 depicts an overview of the operational setup.

Figure 14: Overview of the operational setup used in order to generate forecasts of the 5 minute prices covering a horizon of 12 hours. The procedure outlined is repeated every 5. minutes. The final forecast is constructed as shown in (18) on page 16.

9 Conclusion and discussion

With the current market structure optimal bids of wind power production on the Nordpool Spot Market requires knowledge of the expected values of up- and down-regulation imbalance unit costs conditional on information available at decision time. For a price taker this leads to the conclusion that the bid should be a quantile in the conditional distribution of the future wind production for which the probability level is defined by the expected values of the imbalance unit costs. In the general case the same hold, but in that case the producer must be able to specify how the bid affects the spot price.

In this report it is shown how this conclusion can be reached by considering maximisation of the expected revenue. The derivation requires some approximations, which are described. In the derivation 1st order approximations are used, the possibility of using 2nd order approximations are briefly mentioned. The derivation is similar to that of Bremnes [2] with the exception that here we explicitly consider the stochastic aspect and suggest how a price maker could use the results. Furthermore, the deviation used here is similar to that used for a price taker by Jónsson [5, Paper D], which base the derivation on the (negative) imbalance cost. Here we also consider cases where the bid is influencing the spot price. This influence is considered in two variants, where the bid in one case is assumed not to affect the regulating prices, and in an other case is assumed not to affect the imbalance unit cost, i.e. the positive difference between spot and regulating prices. It is argued that the later assumption might be most consistent with the assumed influence of the bid on the spot price. Jónsson [5, Paper E] also considers the case where the bid affects the spot and regulating prices.

Based on data from the two Danish Nordpool Spot areas covering the period 2001-2011 a simple model not using external inputs are formulated and evaluated for imbalance unit cost forecasting. The evaluation focus on the ability of the forecast method to produce the conditionally expected values of the imbalance unit costs as required by the decision rule. Generally, the observed means agrees well with the forecast values. For yearly evaluations this is also true, but the uncertainty of the observed means as estimates of a true underlying mean is quite large in this case, notably in the end of the period. This indicates that the financial risk of the bidding approach for yearly periods might be significant. A similar analysis is performed for the quantile probability level for a price taker. This analysis leads to similar conclusions.

Forecasts are also generated for the nine data sets containing simulated prices ENFOR/08EKS0004A003-A[3]. In this case the final electricity price is forecast for horizons up to 12 hours assuming the spot price to be known, which will be true in practice except for a short period just after noon. In this case normalized performance measures are reported separately for each year. Furthermore, cumulative error plots highlights extreme events. This information should be used when selecting periods for system simulation.

Operationally, 5 minute prices are simulated. These prices are available on the ftp-server (`host4.enfor.dk`) in the *beginning* of the time interval for which they are valid ENFOR/08EKS0004A003-A[3]. Forecasts of these 5 minute prices are generated operationally and the setup is briefly described here.

A Alternative derivation of the optimal bid

Using (1) and (12) the revenue for hour t can be expressed as

$$R_t = P_{S,t}A_t - C_{D,t}(A_t - B_t)I(A_t > B_t) + C_{U,t}(A_t - B_t)I(A_t < B_t)\,, \tag{23}$$

Using 1st order approximations the conditionally expected revenue can be expressed as

$$\begin{aligned} E[R_t|\mathcal{X}_{t_0}] =& E[P_{S,t}|\mathcal{X}_{t_0}]E[A_t|\mathcal{X}_{t_0}] \\ & - E[C_{D,t}|\mathcal{X}_{t_0}]E[(A_t - B_t)I(A_t > B_t)|\mathcal{X}_{t_0}] \\ & + E[C_{U,t}|\mathcal{X}_{t_0}]E[(A_t - B_t)I(A_t < B_t)|\mathcal{X}_{t_0}] \end{aligned} \tag{24}$$

where the first term involves an approximation which is not needed in Section 2.

Since the properties of the random variable representing the actual future production $A_t|\mathcal{X}_{t_0}$ is unaffected by the bid the partial derivative w.r.t. the bid can be expressed as

$$\frac{\partial}{\partial B_t}E[R_t|\mathcal{X}_{t_0}] = -E[A_t|\mathcal{X}_{t_0}]a_p + E[C_{D,t}|\mathcal{X}_{t_0}] - (E[C_{D,t}|\mathcal{X}_{t_0}] + E[C_{U,t}|\mathcal{X}_{t_0}])F_{A_t|\mathcal{X}_{t_0}}(B_t) \tag{25}$$

where $a_p = -\partial E[P_{S,t}|\mathcal{X}_{t_0}]/\partial B_t$. Furthermore, it is assumed that $\partial E[C_{D,t}|\mathcal{X}_{t_0}]/\partial B_t = 0$ and $\partial E[C_{U,t}|\mathcal{X}_{t_0}]/\partial B_t = 0$, this is in contrast to a similar assumption in Section 2 involving the regulating prices instead of the imbalance unit costs. Since the spot price affects the level of the regulating prices the assumption in the deviation of (25) may be more appropriate than the one used in Section 2.

Based on (25) the bid maximizing the revenue can be expressed as

$$(E[C_{D,t}|\mathcal{X}_{t_0}] + E[C_{U,t}|\mathcal{X}_{t_0}])F_{A_t|\mathcal{X}_{t_0}}(B_t) = E[C_{D,t}|\mathcal{X}_{t_0}] - E[A_t|\mathcal{X}_{t_0}]a_p \tag{26}$$

or

$$F_{A_t|\mathcal{X}_{t_0}}(B_t) = \frac{E[C_{D,t}|\mathcal{X}_{t_0}] - E[A_t|\mathcal{X}_{t_0}]a_p}{E[C_{D,t}|\mathcal{X}_{t_0}] + E[C_{U,t}|\mathcal{X}_{t_0}]} \tag{27}$$

Comparing (25) and (13) it is seen that the alternative approximations and assumptions used here result in a slightly different solution for the general case where $a_p > 0$. However, when B_t is close to $E[A_t|\mathcal{X}_{t_0}]$ the two expressions are similar since $E[A_t|\mathcal{X}_{t_0}]a_p - 2a_pB_t$ from (11) approximately equals $-E[A_t|\mathcal{X}_{t_0}]a_p$ from (25).

B Evaluation of results individually for each year

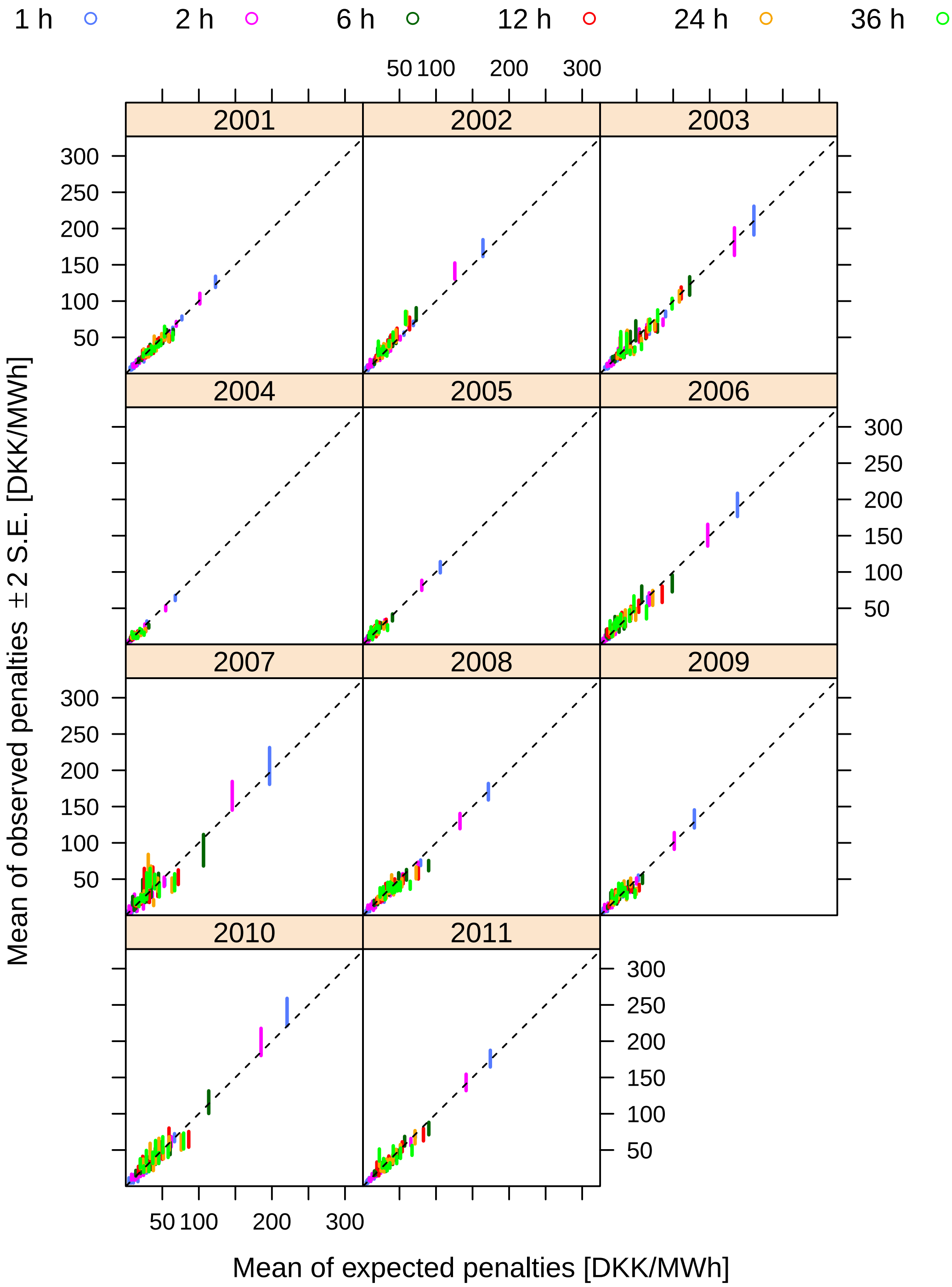

Figure 15: Down-regulation imbalance unit cost for DK1; mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ±2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 6 weeks.

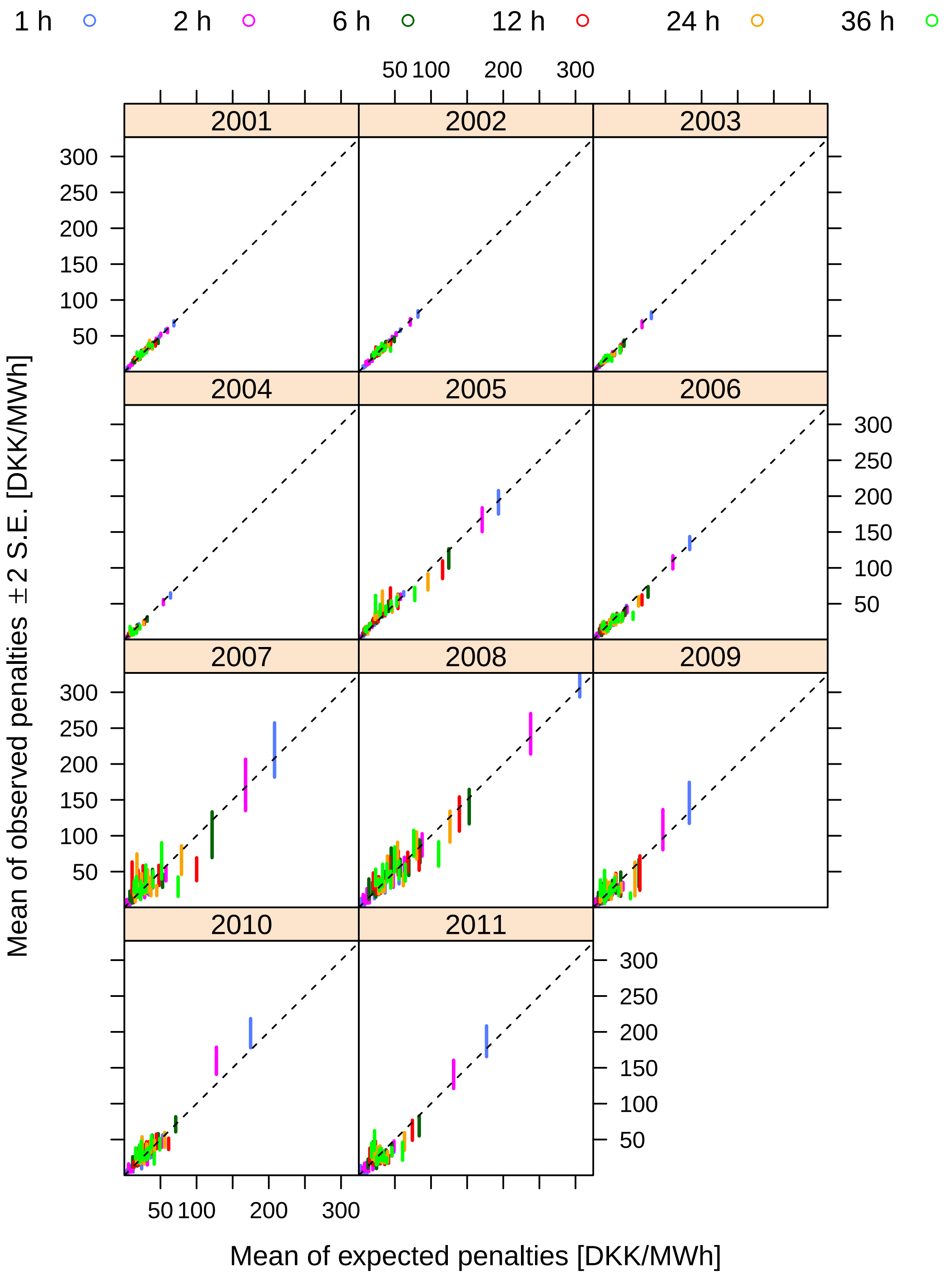

Figure 16: Up-regulation imbalance unit cost for DK1; mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ± 2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 6 weeks.

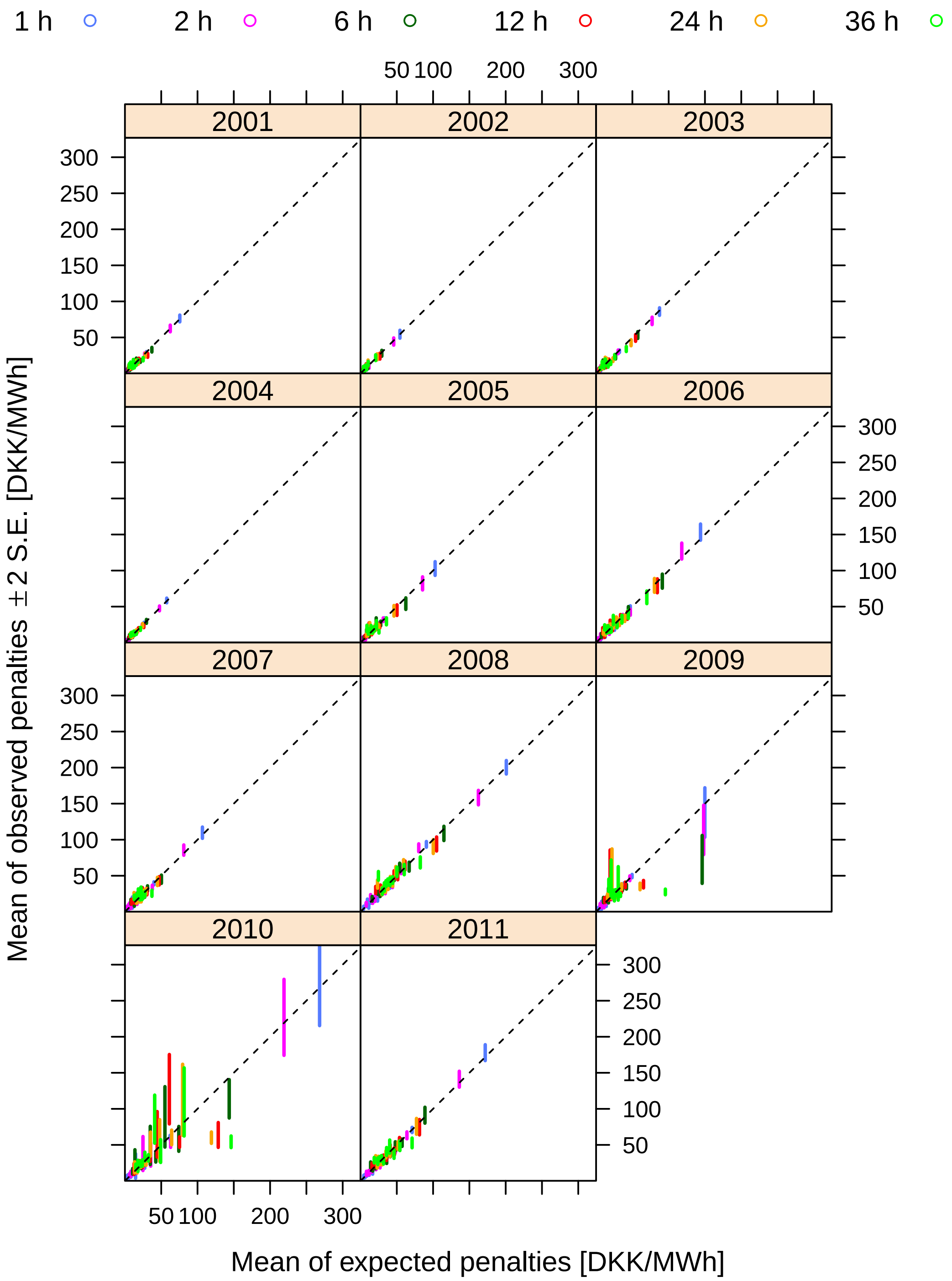

Figure 17: Down-regulation imbalance unit cost for DK2; mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ± 2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 6 weeks.

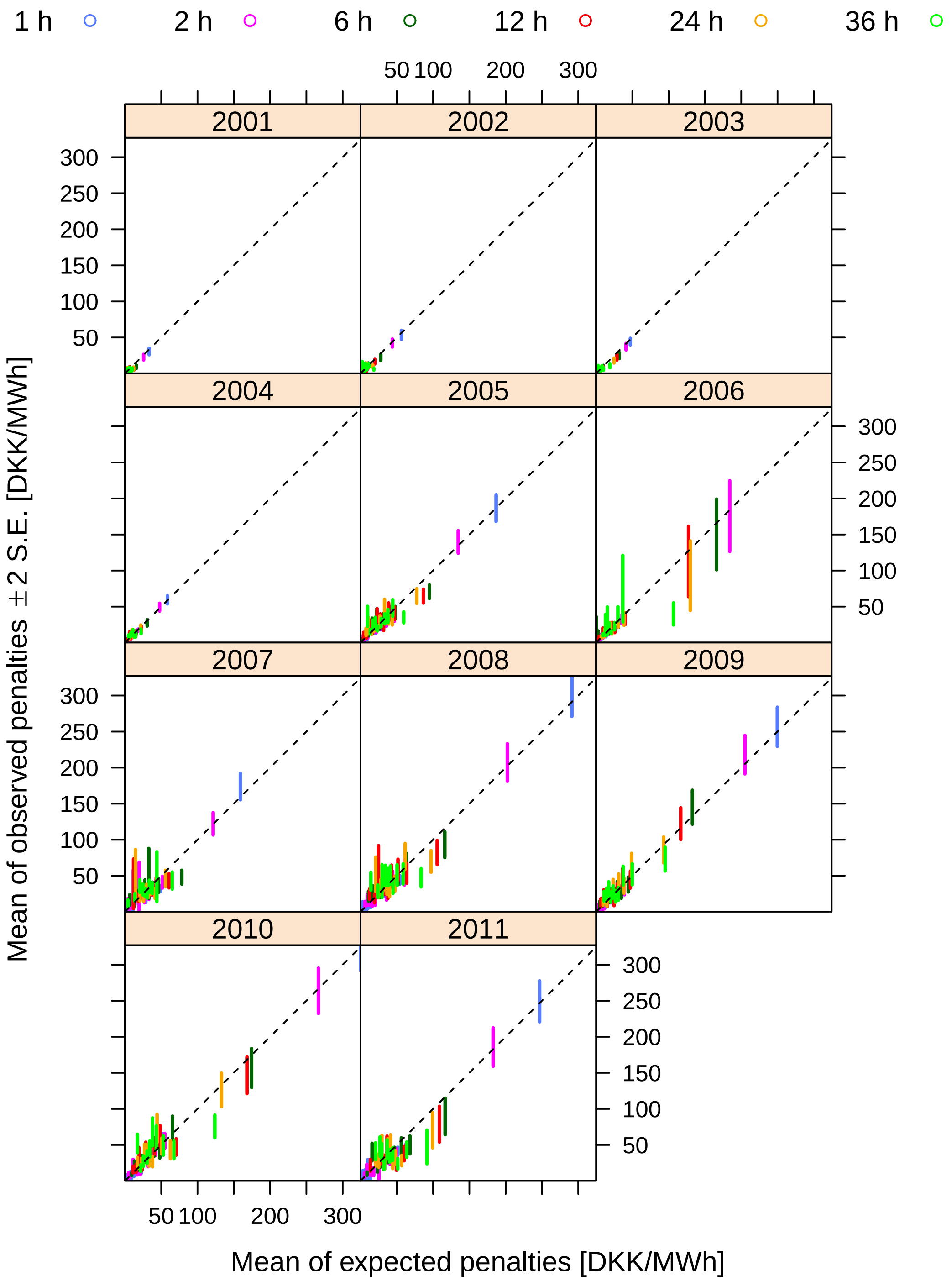

Figure 18: Up-regulation imbalance unit cost for DK2; mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ±2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 6 weeks.

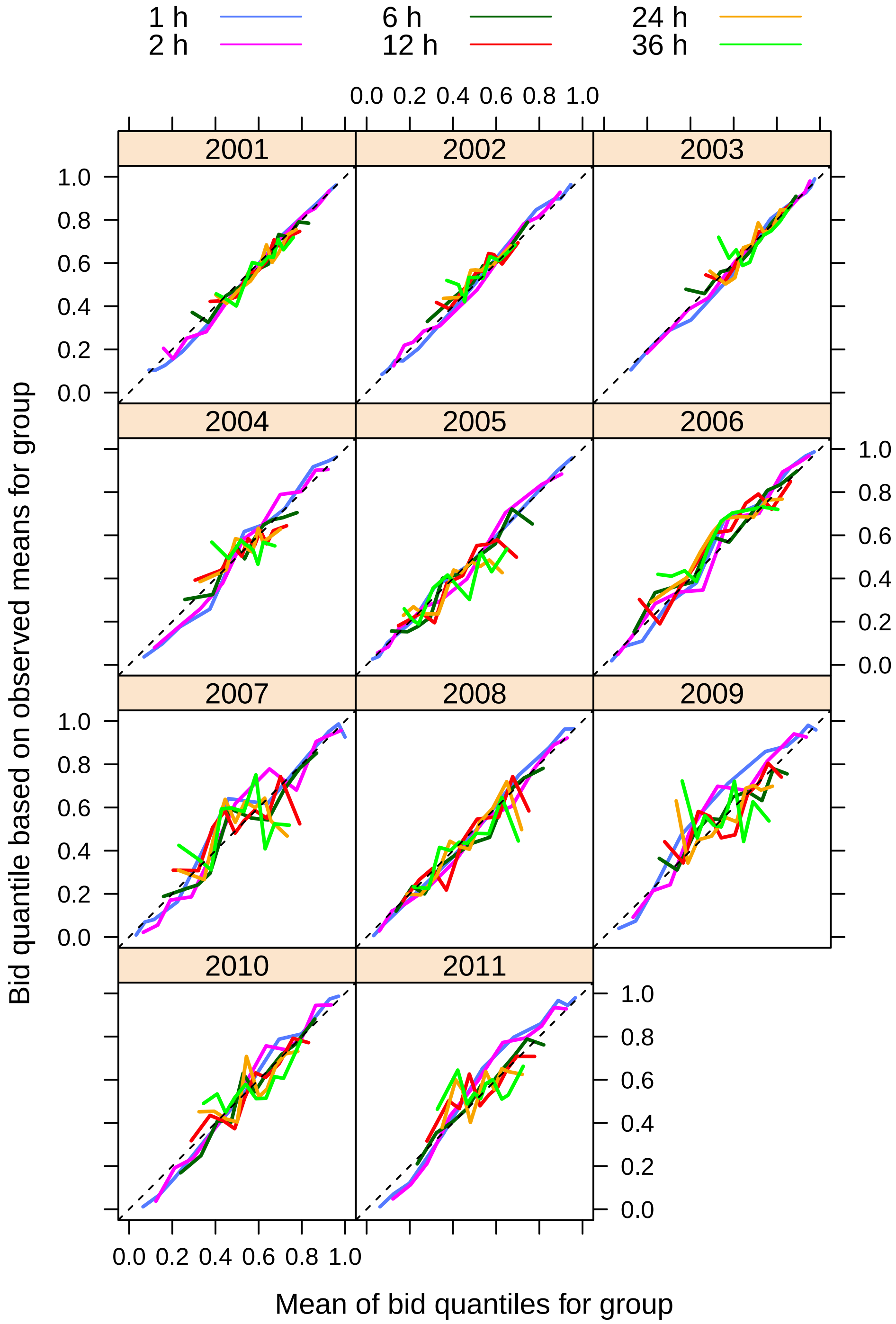

Figure 19: Probability for bid quantile for DK1; probability of bid quantile based on observed means versus mean of probability of bid quantile based on expected (forecast) penalties. The results are obtained using a forgetting factor corresponding to 6 weeks.

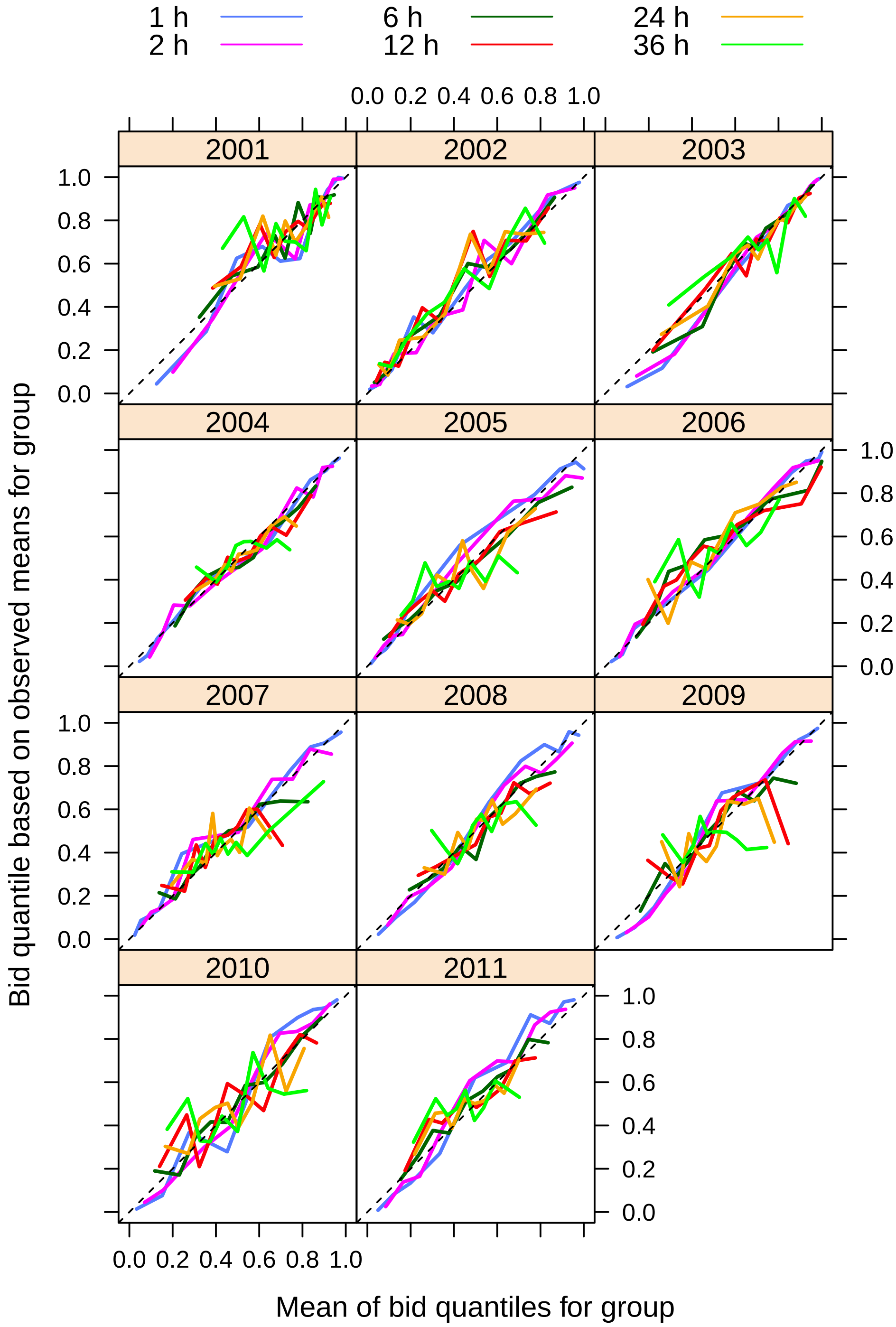

Figure 20: Probability for bid quantile for DK2; probability of bid quantile based on observed means versus mean of probability of bid quantile based on expected (forecast) penalties. The results are obtained using a forgetting factor corresponding to 6 weeks.

C Evaluation of day ahead horizons

For bidding on the current spot market forecasts of the day ahead imbalance unit costs are of main interest, cf. Section 2. In this Appendix corresponding results are presented. By day ahead forecasts we mean the forecasts of imbalance unit costs available in the morning before the day of delivery starting at midnight. Since the spot bid must be ready by 12 noon local time we consider forecasts which are ready at 10 am local time. Furthermore, based on experience, we assume a 2 hour delay when receiving the regulating prices, see ENFOR/08EKS0004A003-A[3] for further details. Hence, the day ahead forecasts are based on data up to 8 am local time. The results are presented for a forgetting factor corresponding to 24 weeks, c.f. Section 5.1-

Figures 21 and 22 show evaluation results for the full period 2001-2011 for conditional means and bid quantiles respectively. Figures 23 and 24 show the evaluation results for the conditional means for the individual years for DK1 and DK2, respectively. Figures 25 and 26 show the evaluation results for the bid quantiles for the individual years for DK1 and DK2, respectively.

For the full period both the evaluation of the conditional means (forecast penalties) and the bid quantiles are quite satisfactory. The conditional means when evaluated on a yearly basis show large variation between years and also the uncertainty of the observed means as an estimate of the true underlying mean is quite large. However, for many of the years the bid quantile probabilities agrees quite well with quantile probabilities based on observed means.

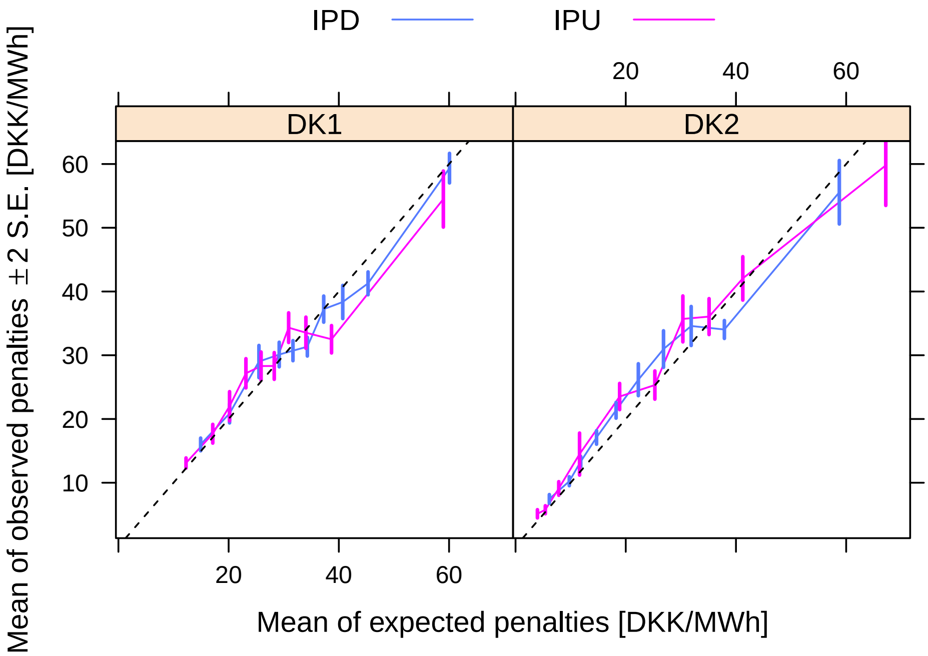

Figure 21: Imbalance unit cost for day ahead horizons; mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ±2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 24 weeks.

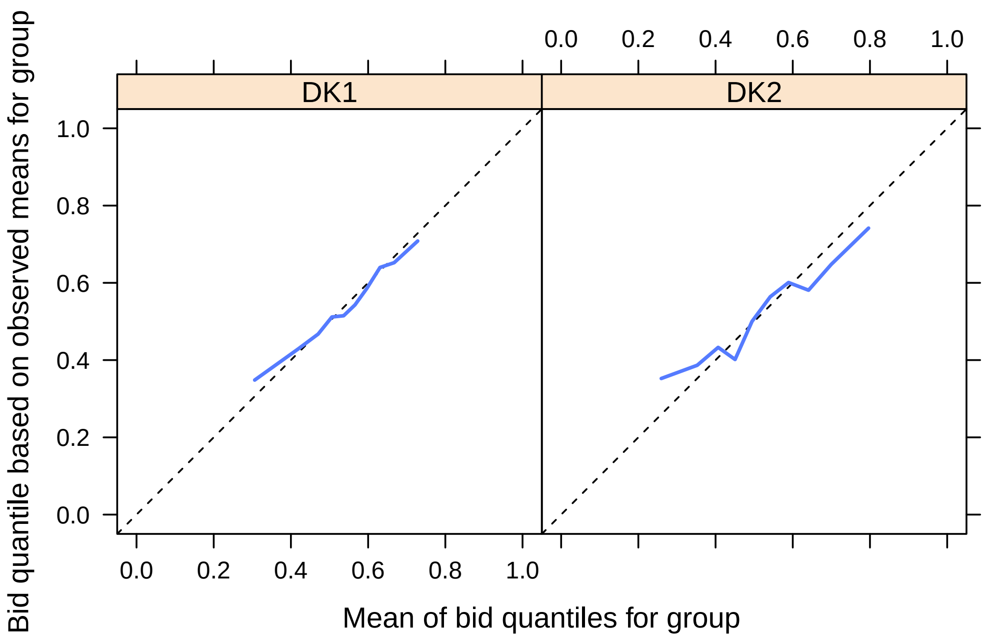

Figure 22: Day ahead horizons; probability of bid quantile based on observed means versus mean of probability of bid quantile based on expected (forecast) penalties. The results are obtained using a forgetting factor corresponding to 24 weeks.

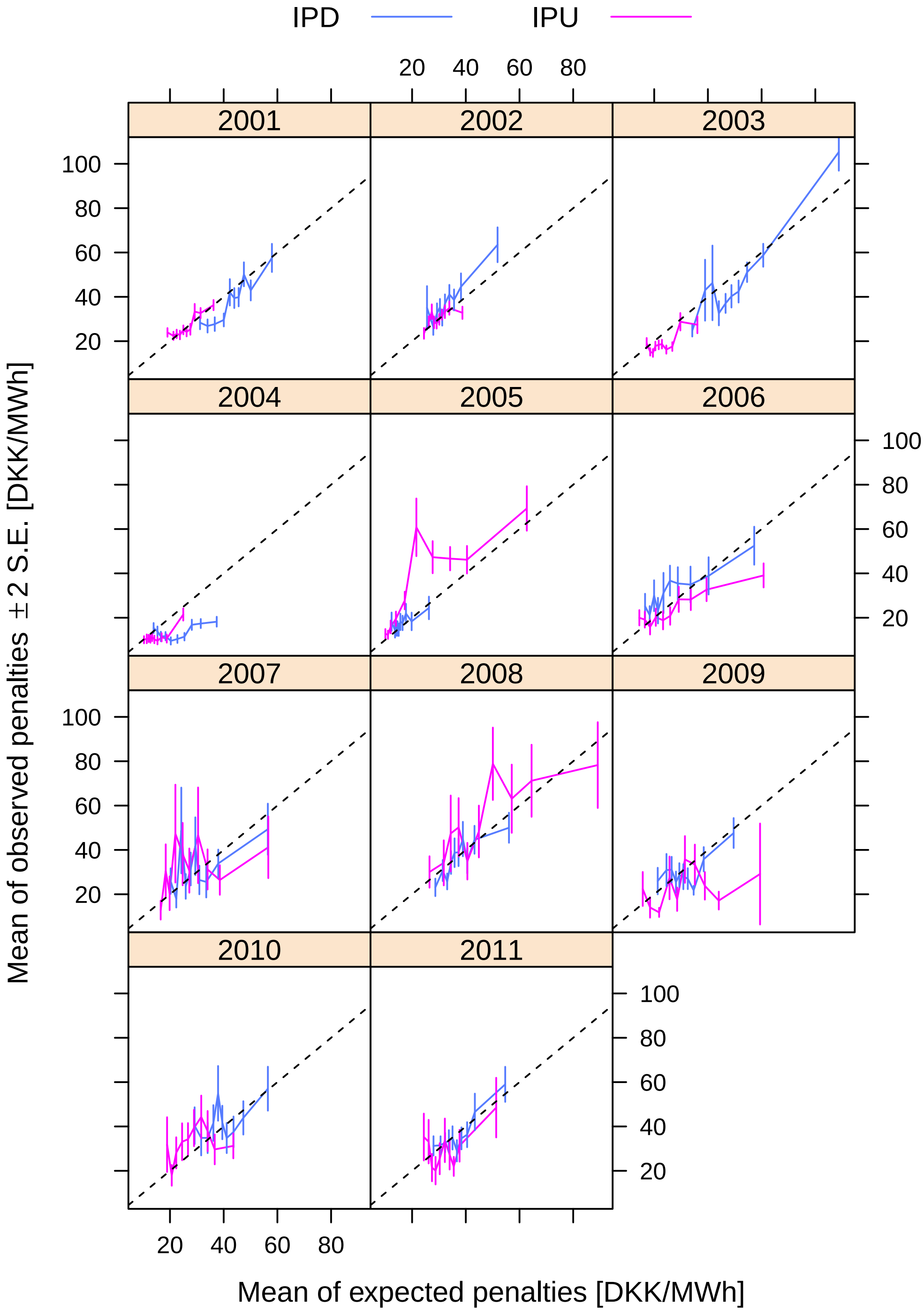

Figure 23: Imbalance unit cost for DK1 (day ahead horizons); mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ±2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 24 weeks.

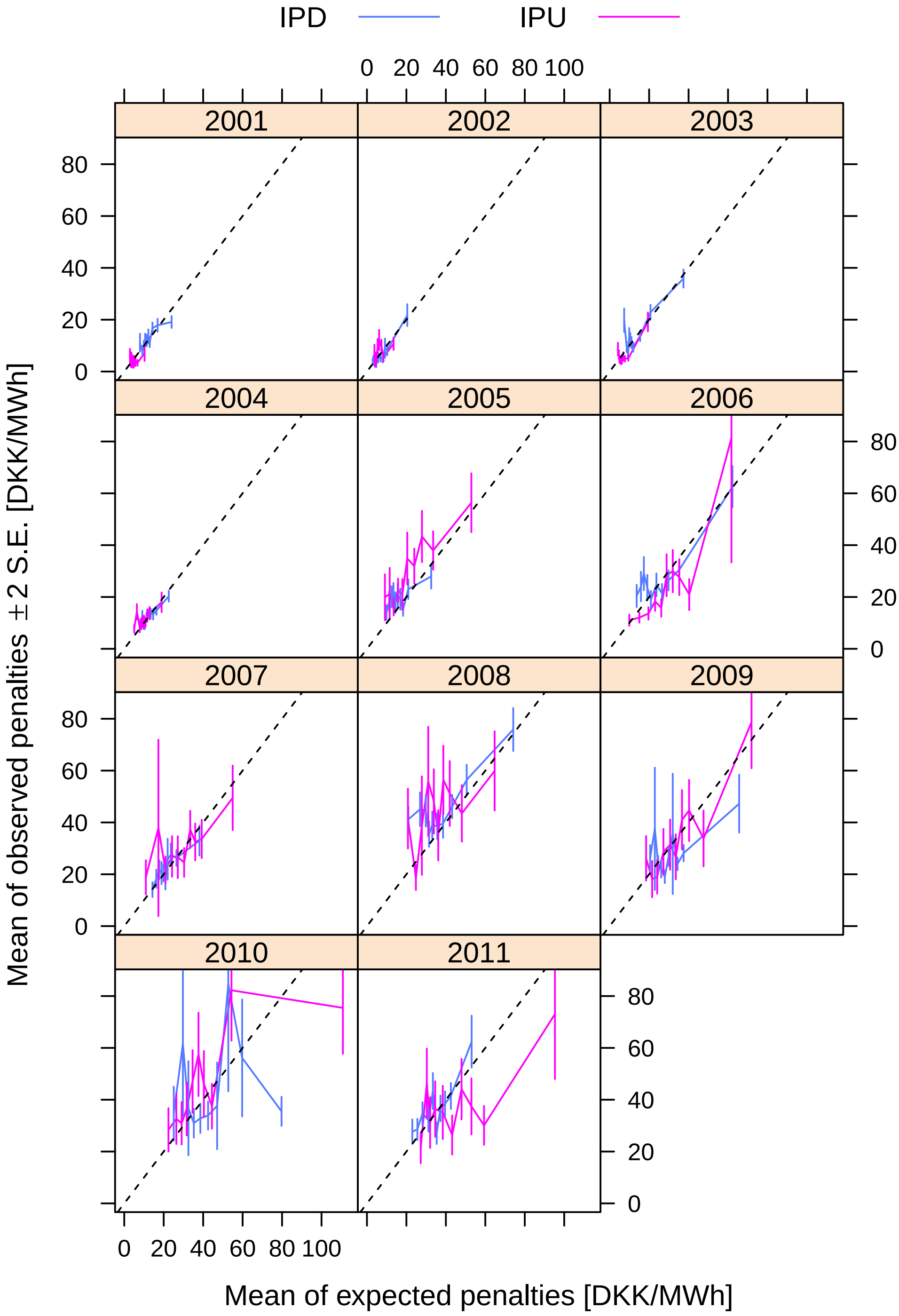

Figure 24: Imbalance unit cost for DK2 (day ahead horizons); mean of observed penalties versus mean of expected (forecast) penalties. The verticals line indicates ±2 standard errors of the observed means. The results are obtained using a forgetting factor corresponding to 24 weeks.

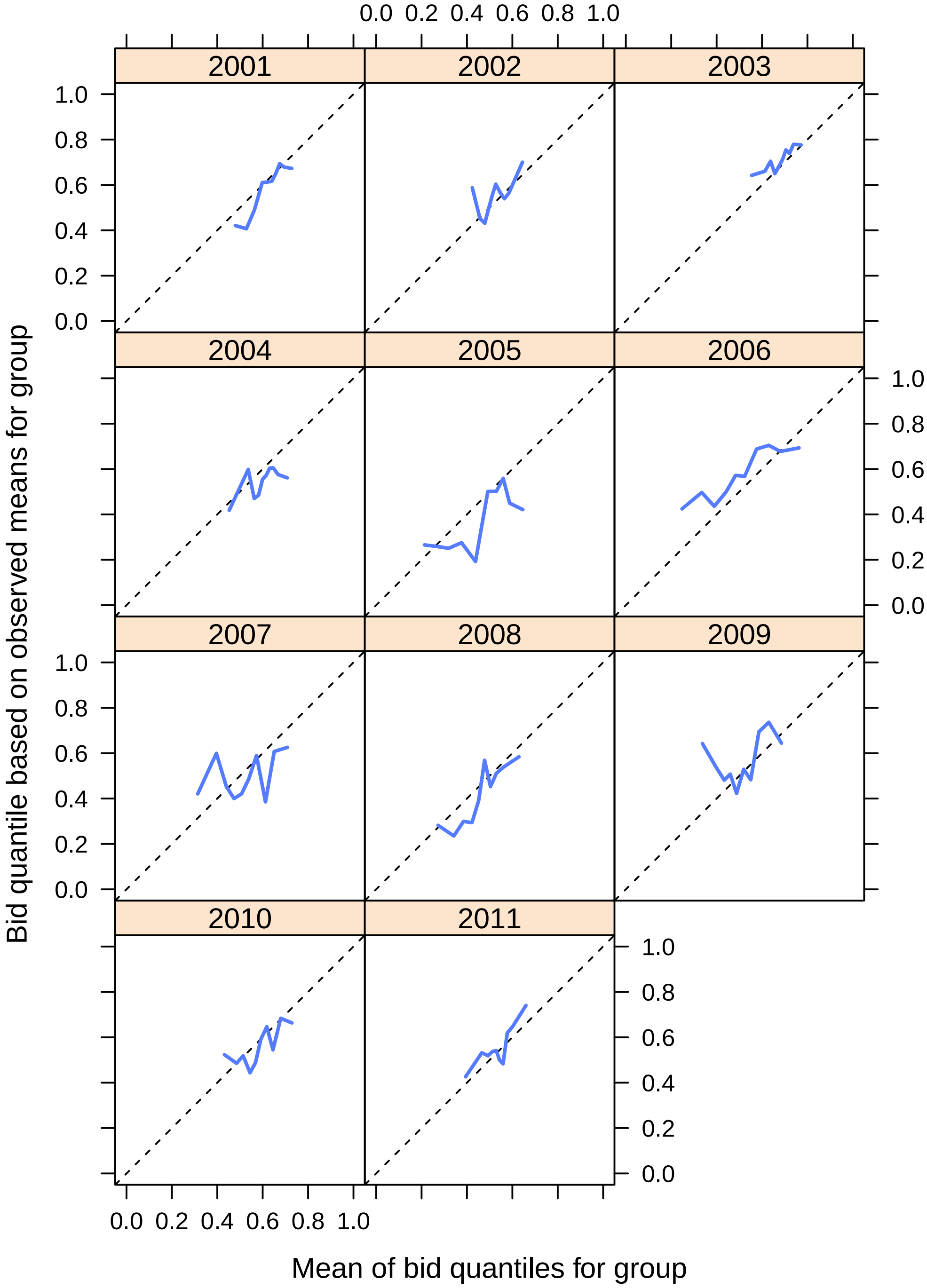

Figure 25: Probability for bid quantile for DK1 (day ahead horizons); probability of bid quantile based on observed means versus mean of probability of bid quantile based on expected (forecast) penalties. The results are obtained using a forgetting factor corresponding to 24 weeks.

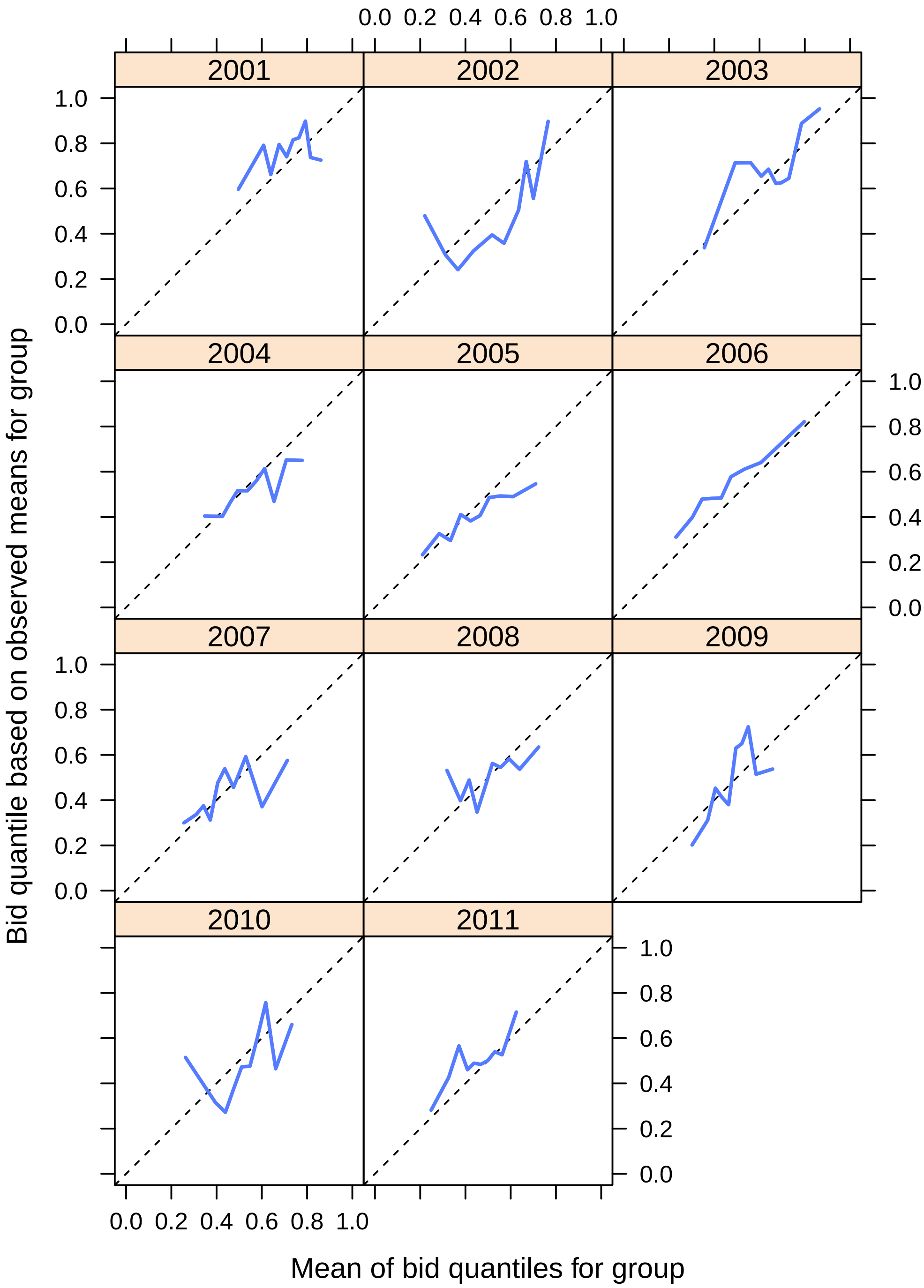

Figure 26: Probability for bid quantile for DK2 (day ahead horizons); probability of bid quantile based on observed means versus mean of probability of bid quantile based on expected (forecast) penalties. The results are obtained using a forgetting factor corresponding to 24 weeks.

D Example R-script for exporting data

```
dir       = "fcsSimPrices5m" # where the rda-files containing forecasts are placed
file      = "P5mFcs,sdf=1,phi=0.8,delta=0.3,r0=0.001.rda" # The particilar rda-file
area      = "dk2" # dk1 or dk2
TimeBegin = "2011-01-01 00:00" # beginning of period; 1st time stamp will be 5 min. later
TimeEnd   = "2011-01-02 00:00" # end of period; last time stamp will be this
expDir    = "." # Directory for export
expFile   = "prices.csv" # File name for export

## The data files can be obtained via FTP from login.enfor.dk, using account
## 'flexpower'. Please contact Henrik Aalborg Nielsen to obtain the password
## and arrange that data are placed on the FTP server.

## The interpolation method interpolates the imbalance unit costs and
## subsequently ads the spot price back on. The implementation of the
## interpolation method used here is to be considered as an example
## illustrating the principle.

## The implementation does not scale well in terms of computer
## time. Therefore, it is better to extract a number of short periods and
## combine these subsequently. In this case the code should be modified to
## load the data only once.

######################################################################

load(file.path(dir, file)) # => P5mFcs

tmfmt <- "%Y-%m-%d %H:%M"

P5mFcs <- P5mFcs[[area]]
P5mFcs <- subset(P5mFcs,
                 Time >= as.POSIXct(TimeBegin, tz="Europe/Copenhagen") - 12*3600 &
                 Time <= as.POSIXct(TimeEnd,   tz="Europe/Copenhagen") )
rownames(P5mFcs) <- format(P5mFcs$Time, format=tmfmt, tz="UTC")

P5mFcsOut <-
  matrix(nrow=nrow(P5mFcs),ncol=144,dimnames=list(rownames(P5mFcs), paste("k=", 1:144, sep="")))

i5mHat <- P5mFcs[,grep("^k=", names(P5mFcs))]
i5mHat <- i5mHat - P5mFcs[,rep("PS", ncol(i5mHat))]
PTime <- P5mFcs[,rep("Time", ncol(i5mHat))] -
  300*matrix(as.numeric(substring(names(i5mHat),3)), byrow=T, ncol=ncol(i5mHat), nrow=nrow(i5mHat))

upt <- sort(unique(do.call(c, PTime)))

for(jj in 1:length(upt)) {
  if(0 == jj %% 72) cat(jj,"\n")
  idx <- format(upt)[jj] == as.matrix(PTime)
  if(sum(idx) == ncol(i5mHat)) {
    Time.jj <- P5mFcs$Time[idx %*% rep(1,ncol(idx)) > 0]
    TimeOut.jj <- seq(min(Time.jj), max(Time.jj), 300)
    i5mHatOut.jj <- approx(x=Time.jj, y=i5mHat[idx], xout=TimeOut.jj)$y
    PS.jj <- P5mFcs[format(TimeOut.jj, format=tmfmt, tz="UTC"), "PS"]
    P5mFcsOut.jj <- PS.jj + i5mHatOut.jj
    diag(P5mFcsOut[format(TimeOut.jj, format=tmfmt, tz="UTC"),]) <- P5mFcsOut.jj
  } else {
    i5mHatOut.jj <- P5mFcsOut.jj <- rep(NA, 144)
  }
}

P5mFcsOut <- cbind(P5mFcs[,-grep("^k=", names(P5mFcs))], P5mFcsOut)

P5mFcsOut <-
  subset(P5mFcsOut,
```

```
        Time >  as.POSIXct(TimeBegin, tz="Europe/Copenhagen") &
        Time <= as.POSIXct(TimeEnd,   tz="Europe/Copenhagen") )

P5mFcsOut$Time <- format(P5mFcsOut$Time, "%Y-%m-%d %H:%M", tz="Europe/Copenhagen")
write.csv(P5mFcsOut, file=file.path(expDir, expFile), quote=F, row.names=F, na="")

## Columns:
##
## Time  : End of 5 minute interval
## PS    : Actual spot price [DKK/MWh]
## P1h   : Actual price; up-regulation price if above PS, down-regulation if below PS [DKK/MWh]
## P5m   : Actual (simulated) 5 minute price [DKK/MWh]
## k=1   : Simple forecast of P5m based on information avalilable time at 'Time -   5 minutes'
## k=2   : Simple forecast of P5m based on information avalilable time at 'Time -  10 minutes'
## .
## .
## k=k   : Simple forecast of P5m based on information avalilable time at 'Time - k*5 minutes'
## .
## .
## k=144 : Simple forecast of P5m based on information avalilable time at 'Time -  12 hours'
```

References

[1] C. Bang, F. Fock, and M. Togeby. Design of a real time market for regulating power. Technical report, Ea Energy Analyses, 2011. FlexPower WP1 – Report 3.

[2] John Bjørnar Bremnes. Probabilistic wind power forecasts using local quantile regression. *Wind Energy*, 7(1):47–54, 2004.

[3] ENFOR/08EKS0004A003-A. Simulation of 5 minute prices based on actual 1 hour data, 2013.

[4] Harley Flanders. Differentiation under the integral sign. *The American Mathematical Monthly*, 80(6):615–627, 1973.

[5] Tryggvi Jónsson. *Forecasting and decision-making in electricity markets with focus on wind energy*. PhD thesis, Department of Informatics and Mathematical Modelling, Technical University of Denmark, 2012.