Technical University of Denmark

DTU

The Effect of Wind Power on Electricity Prices in Denmark

Jonsson, Tryggvi; Madsen, Henrik

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

The Effect of Wind Power on Electricity Prices in Denmark

Tryggvi Jónsson
Henrik Madsen

November 29, 2007

Contents

CHAPTER 1

Introduction

This report is the result of a special course taken by the author at IMM DTU under the guidance of professor Henrik Madsen. The aim of the project is to analyze the influence wind energy has on the electricity spot price in Western Denmark and investigate how information about wind power production can be used to model the electricity spot price. Various model types were tried, giving very different performance. Here, only the models that performed best are discussed in order to keep focus on the projects goal.

1.1   Background

The Western Danish (from now on denoted DK-West) electricity system along with the Norwegian, Swedish, Finnish and Eastern Danish ones, make up the Nordic power system. In the system there is a common power exchange, Nordpool, where hourly system price is set for 24 hours at a time based on supply and demand bids. The bids are sent in at noon for the time interval from midnight to midnight, that is the price is set $12-36$ hours ahead in time. The system price is valid in all areas of the Nordic system unless the transmission capacities are not sufficient, then individual areas can make up their own price area. For DK-West, the typical circumstances for this to happen is when wind power production is high and when transmission capacities to Norway and Sweden are fully utilized. *Figure 1.1* shows the difference between the common system price and the DK-West price in 2006. Although it should be noted that 2006 is not typical for this difference, due to extremely little rain in the summer of 2006, the level of water in the Norwegian and Swedish hydro power reservoirs, which has by far the greatest influence on the Nordic spot price, was way below average and therefore the common spot price was unusually high in the late summer and early autumn. Normally the price on DK-West is a bit higher than the system price. However what is common with 2006 and other years is that the DK-West price has a lot more variance than the system price. This can be explained by the large penetration of wind power in Denmark. Wind power production is much more unstable than hydro-, thermal- and nuclear power plants and that results in a more varying power price. The advantage of wind power is however that despite rather high initial capital of a wind farm, the production cost is very low and therefore once the farms are up, it is always beneficial to keep them producing.

For agents biding into a day-ahead markets such as Nordpool and most other energy spot markets, an accurate prediction of the price to come is a key instrument in maximizing the profit. Therefore in Denmark's case, making it possible to further decrease the government support to the wind energy sector. Now a days, agents biding into the spot market are
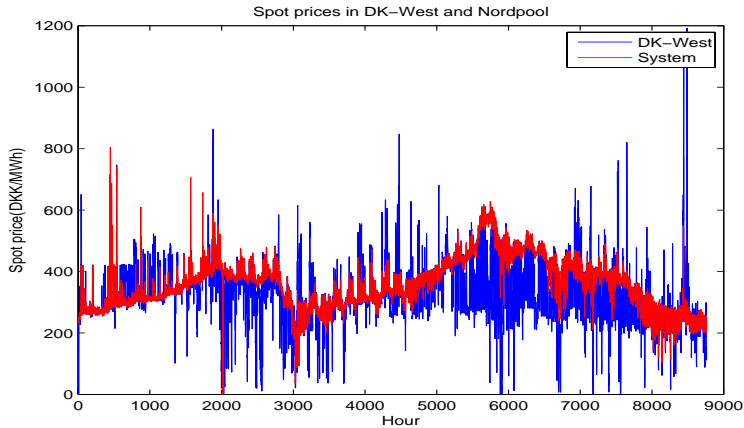
Figure 1.1: Spot prices in DK-West and Nordpool in 2006

often forced to bid wind energy into the market at the price of 0 in order to prevent the wind farms to be forced to shut down. With good knowledge of the prices to come this could be changed. Statistical models have been used in other countries such as Spain with good results and therefore it seems obvious that they could be of use in Denmark also. What makes it more difficult to predict the spot prices in Denmark however is that, wind power makes up for much more part of the energy production in Denmark than for example Spain[4][1] and therefore are spot prices more unstable.

1.2   The Data

Energinet has made available through their home-page hourly data for the electricity prices, production(central production and wind power production separately), and trading from several years back and until today. The data is kept separated for the two electricity grids in Denmark, Jutland and Fyn on the one hand and the islands on the other hand. When it is appropriate the data for the whole Nordpool area is also included. The data used for evaluating the models in report was obtained from this homepage and span the whole year of 2006. Furthermore data on predictions of wind power production was also obtained from Energinet and used to make predictions for the prices. These data are however confidential and therefore not available on Eneginet's homepage.

It is a known fact that energy consumption varies over 24 hours and this poses a small problem when including the predicted wind power production in the models. Namely that an hour with big predicted production has much more effect during the night than it has during the day as it will make up for much more of the total energy production. It is therefore concluded that it would be desirable to include the predicted proportion of wind power in the system at each time instead of just the wind power predictions directly. By utilizing the fact that energy consumption is practically constant for a given time of the year and the day, one can get a good estimate of the predicted penetration of wind power in the system. This is done by calculating how big proportion the predicted wind power would be of the measured total energy consumption each hour. Thereby the predicted

---

[1]Source in [4]: www.iaea.org and www.ens.dk

wind power production is transformed to the predicted wind power penetration in the system. In the following both versions are tried for the sake of comparison.

1.3   Previous work

IBT-Wind, which is a research center for wind power founded by Århus University and The Business and Engineering university in Herning in collaboration with numerous companies, universities and institutes in Denmark, published in December 2006 a report on the influence of the wind power on the DK-West spot prices. The IBT-Wind report contains mostly a static analysis of how the spot prices change with wind power production in 2005 along with discussion of the characteristics of the spot market. Here, similar static analysis is made for the year 2006 and the results used to build up models for the spot prices. However here the influence of the predicted wind power production is estimated instead of the actual production. Furthermore are there used non-parametric kernel methods which are not used in IBT-Wind´s report along with the discrete methods.

CHAPTER 2

Static analysis of the spot prices

As is stated in [3], when the effects of wind power on danish electricity prices are to be measured, the right question to ask is: " What would the prices have been, if the wind hadn't blown?", but not: " What would the prices have been if there were no wind power in the danish electricity system?". Since the demand for electricity would be the same whether or not wind energy is produced, the energy would have to be produced in some other manner. In the following the aim is to come up with some answers to this question.

2.1 Daily variability of the spot prices

It is a well known fact that demand for energy varies over the day. The majority of people is awake during the day and sleeps in the night and therefore more energy is needed during the day when shops, offices and institutes are open. Furthermore some of the most energy consuming household equipment are in the kitchen or the laundry room, which are mainly used in the morning and after work hours. Therefore typically for the Western world, the use of electricity is noticeably higher during the day than in the night with demand peaks in the morning and around dinner time. Here Denmark is no exception and according to the principle of supply and demand it seems natural to assume that this demand peaks should result in a higher price at peak hours. *Figure 2.1* shows the average spot price in 2006 for each hour of the day for different amount of wind power predicted to be in the system. There the peaks spoken about before are well detectable and furthermore it is noticed that the prices seem to go down as the wind power input increases. This seems reasonable, since wind energy, despite it's high initial capital, has very low production cost as mentioned earlier and therefore when the wind farms produce much energy, there is more cheap energy in the system.

Now *figure 2.1*, only shows the average price for roughly categorized wind power input and therefore assumes linear changes between categories. In order to obtain a more smoother estimate of how the spot prices change over 24 hours, Locally Weighted Polynomial Regression(LWPR) is used. By doing that the spot prices are assumed to be a non-linear function of time of day and the predicted wind power input. A further discussion of LWPR can be found in *Section 3.1.2*. The smooth estimate is very consistent with the real one and displays the same characteristics of the spot prices as mentioned earlier. Since it is obvious that predictions of wind power inputs do influence the price, it can be concluded that the predictions could be of use when modeling the spot price. However some other aspects need to be examined before that is done.
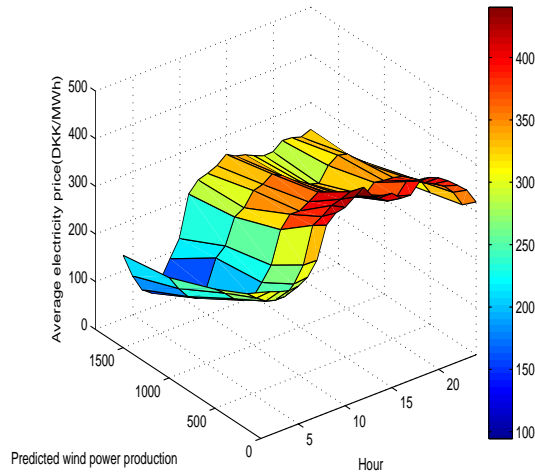
Figure 2.1: The relationship between the time of the day, wind production and price
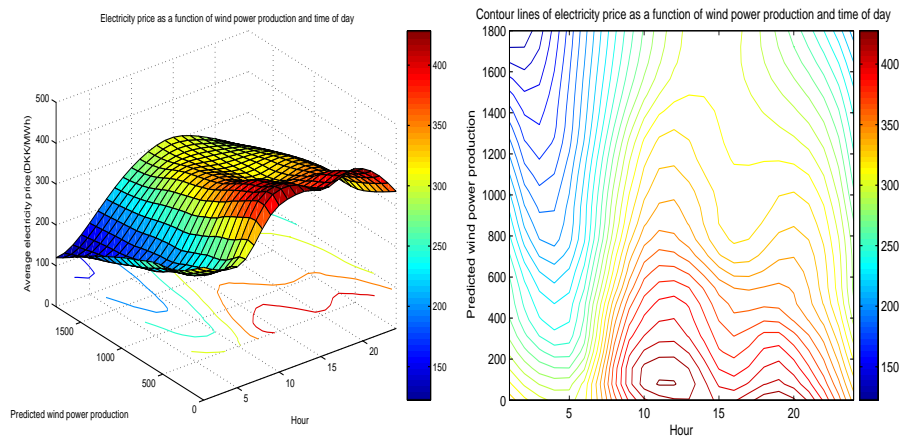

Figure 2.2: The relationship between the time of the day, wind production and price estimated with LWPR

2.2 Variability of the spot prices over the year

As said earlier the Nordpool spot prices are, in the long run, mainly influenced by the level of water in the Norwegian and Swedish hydro power reservoirs. However the changes in the water level are relatively slow so for a model with short prediction interval, these changes can most likely be accounted for by adaptive parameters. The water level of these reservoirs is dependent on how much and what kind of precipitation falls. Obviously it rains and snows more in the winter time than in the summer time, and therefore the water level is generally lower during the summer than it is in the winter and normally reaches a year minimum in the late summer. Therefore one would assume that prices will be higher during the summer. By inspecting the graph displayed in *figure 2.3* a noticeable

difference in the average spot price can be seen. The prices seem to be higher during the
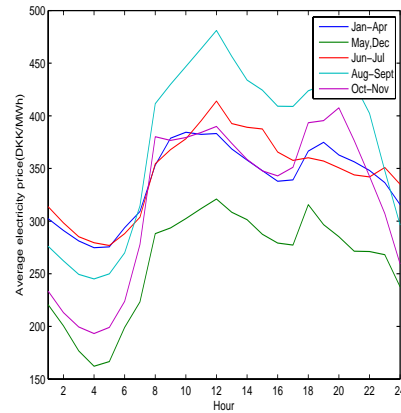

Figure 2.3: Average electricity price for different months in 2006

summer as was argued for before. However the effect of the low water level is smaller than anticipated since energy consumption is much lower during the summer. People are traveling abroad, and also there are approximately twice as many daylight hours during the summer compared to the winter.

Another thing of interest about the spot prices is that when looking at the spot prices from *figure 1.1* for DK-West, one notices that the prices are more stable during the summer than in the winter time. The reason is among other things that in the winther, the wind blows more and it rains more which both have increasing effect on wind power production. So in the winter there is more penetration of wind power in the system, leading to a more unstable prices.

2.3   The effects of wind power on the spot prices

Having established that wind power penetration does in fact influence the electricity prices, the remaining question is: "How?".

As said earlier, whether the wind blows or not, the demand for energy is practically a constant at any given time. So an estimate of the savings from wind energy can be obtained just by looking at the price as a function of production. In *figures 2.4 & 2.5* the average price for a given hour and a production for two periods in the summer and during the winter are shown respectively. The graphs clearly show that as the production of wind power increases, the price gets lower. It is also worth noticing that with increased wind power the two peaks become less extreme.

Now although the energy demand is very stable for a given time of the day it varies throughout the day as mentioned earlier. Therefore the approach taken above is maybe not entirely correct. Since there is much less need for electricity during the night, an hour of big production during the night will make up for much more of the demand than it would do in the day and therefore have much more influence on the spot price. An approach that will account for this is to view the prices as a function of the ratio between the total energy demand and the wind power production, i.e. as a function of
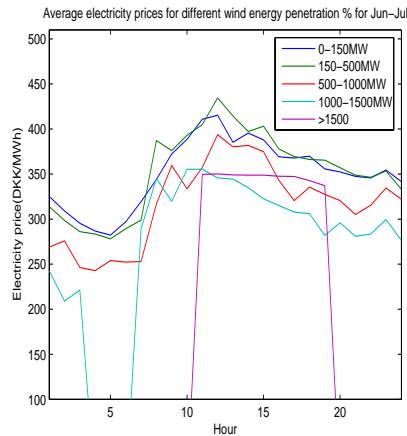
Figure 2.4: Average spot prices for different predicted wind power production in Jun and Jul 2006
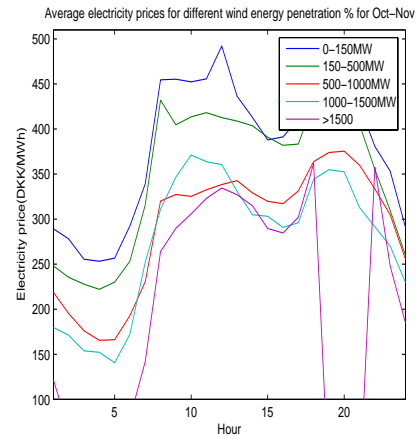
Figure 2.5: Average spot prices for different predicted wind power production in Oct and Nov 2006

the proportion of the wind energy in the system. *Figures 2.6 - 2.7* show the average spot prices in the same periods as before as a function of the wind power penetration. Furthermore the average price through out the year can be viewed in *figure 2.8*. These
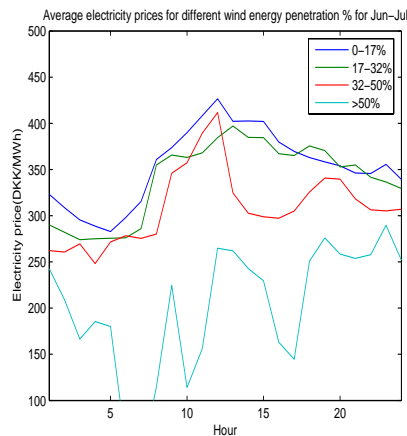

Figure 2.6: Average spot prices for different predicted wind power penetration in Jun and Jul 2006
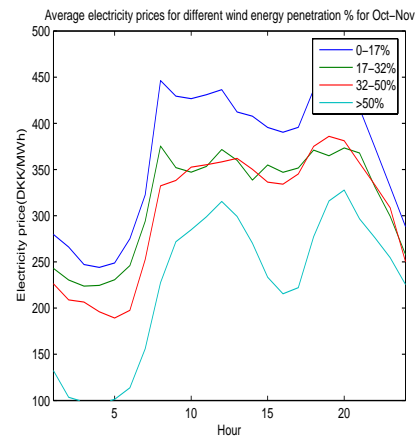
Figure 2.7: Average spot prices for different predicted wind power penetration in Oct and Nov 2006

figures illustrate even further how the wind power lowers the spot prices.

To get a better estimate of how much the prices differ compared to if the energy was to be gotten from other available energy sources, one can view the bar plots in *figures 2.9 - 2.14*, where the average prices and average difference is summed up for the same periods as before. In these plots a production of $150MW$ or less or a wind penetration of about
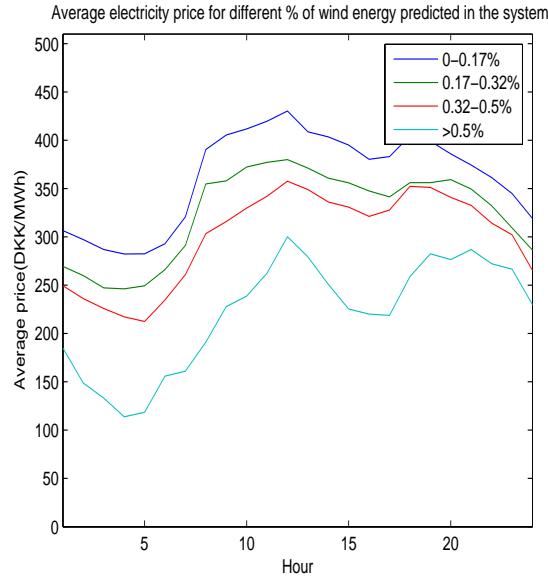
Figure 2.8: Average spot prices for different wind power production in 2006

4% or less is taken to be "no wind power". This is assumed to be reasonable since prices do not change much for the production interval $0 - 150MW$. Similar plots for other parts of the years and plots from the more naive approach can be viewed in Appendix A. From
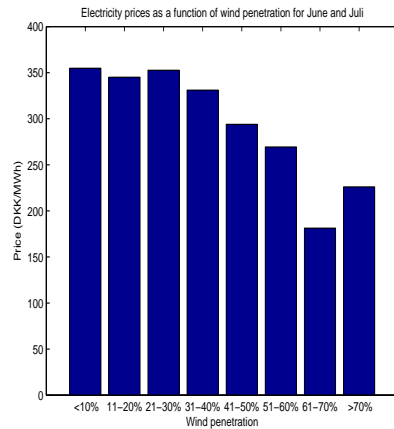

Figure 2.9: Average spot prices for a given wind penetration(Jun/Jul)
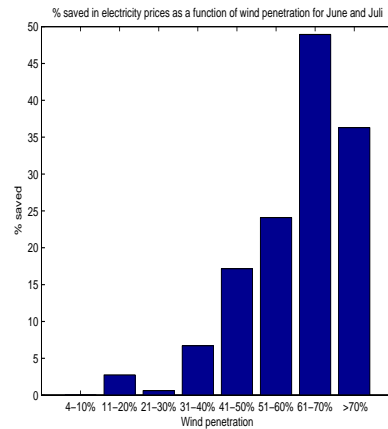
Figure 2.10: Average price difference in % for a given wind penetration(Jun/Jul)

those figures it is clear that the difference, between the prices when wind penetration is high and when wind power makes up for a small proportion of the energy, is considerable. Especially is the difference large during the winter, when the wind blows more. Though it has to be said that wind penetration above 50% is rather rare and therefore difference
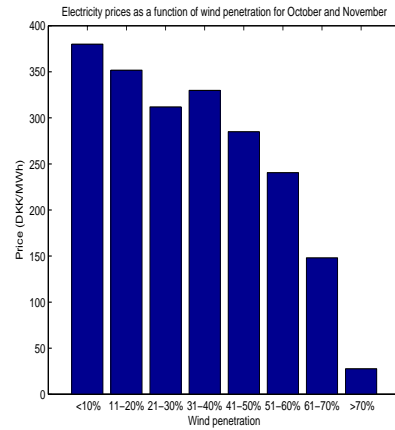
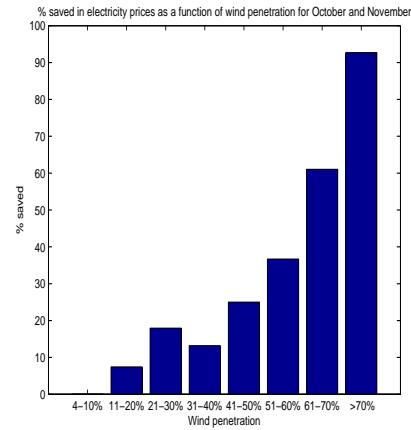Figure 2.11: Average spot prices for a given wind penetration(Oct/Nov)

Figure 2.12: Average price difference in % for a given wind penetration(Oct/Nov)
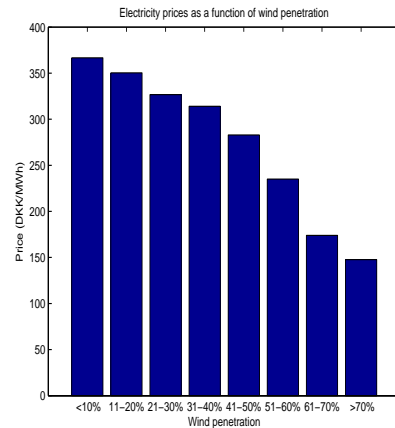
Figure 2.13: Average spot prices for a given wind penetration(2006)
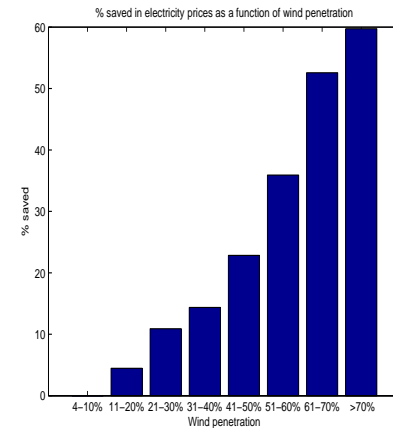
Figure 2.14: Average price difference in % for a given wind penetration(2006)

of 40% and above therefore are far from the average.

A summary of the price differences in 2006 is given in *figures 2.15 & 2.16*. Still wind penetration under 4% is taken as "no wind", and *figure 2.15* shows how much lower the spot prices were on average when the wind penetration exceeded those 4%. The difference is considerable especially when it is taken into consideration that the money, energy companies buy energy for runs on a few billion DKK per year. *Figure 2.16* shows the difference on the average spot price for "no wind" and when the wind is blowing, where the previous statements about the main effects are during the winter are confirmed further.

Figure 2.15: % saved in energy prices in 2006

Figure 2.16: Average spot prices for 2006

Time invariant and adaptive models for the spot price

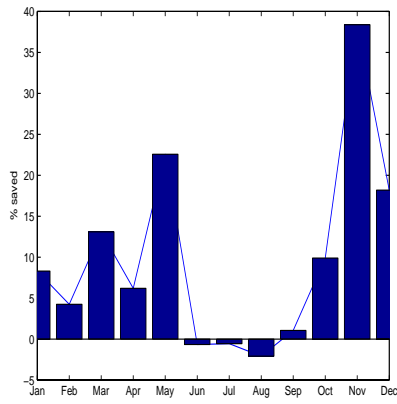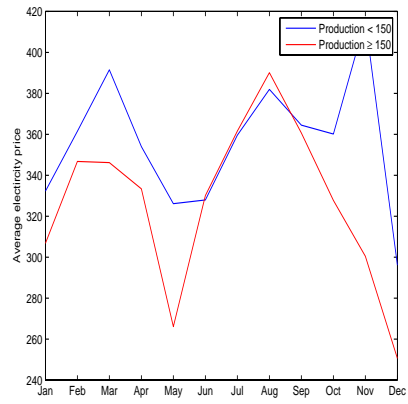With the information from previous chapter, it seems obvious that information about wind power production and therefore the wind penetration in the system will be of great help when predicting the spot prices. In this chapter different possibilities of building a model for the spot price and their performance is estimated. Both time invariant and adaptive ways are tried. Derivations of modeling algortihms are obtained from [1][1] and [2][2].

3.1 Model construction

The spot prices can be described as a linear system in discrete time as

$$Y_t + \phi_1 Y_{t-1} + \ldots \phi_k Y_{t-k} = \omega_1 U_{t-1} + \ldots + \omega_m U_{t-m} + \epsilon_t$$

Where Y_t is the spot price (output signal), ϵ_t is white noise and U_t is an uncorrelated input signal. By introducing the vectors

$$\mathbf{X_t} = [-Y_{t-1}, \ldots, -Y_{t-k}, U_{t-1}, \ldots, U_{t-m}]^T$$
$$\theta_{\mathbf{t}} = [\phi_1, \ldots, \phi_k, \omega_1, \ldots, \omega_m]^T$$

The system can be described as the linear regression model

$$Y_t = \mathbf{X_t^T} \theta + \epsilon_t$$

The least squares estimate of the model parameters, θ is the solution to

$$\hat{\theta} = \arg \min_{\theta} S(\theta)$$

where

$$S(\theta) = \sum_{s=1}^{n} \left(Y_s - \mathbf{X_s^T} \theta\right)^2$$

---

[1]Chapter 9
[2]Chapter 10

This estimate is often referred to as the *ordinary least squares estimate* and can be found as.

$$\hat{\theta} = \left(\mathbf{X}\mathbf{X^T} \right)^{-1} \mathbf{X}\mathbf{Y}$$

where $\mathbf{X}$ is a matrix with $\mathbf{X_t}$ as the tth column and $\mathbf{Y}$ is a column vector with Y_t as the tth variable.

Since this model is static, it does not account for any changes or seasonal variation that happen in the long run, compared to the prediction horizon. One way of dealing with such effects is to estimate the parameters recursively so the parameters, θ become time dependent (θ_t) and adapt to the relatively slow changes in the system. The recursive estimate of the parameters in $\theta_\mathbf{t}$ is found as the solution to

$$\hat{\theta}_\mathbf{t} = \arg\min_\theta S_t(\theta)$$

where

$$S_t(\theta) = \sum_{s=1}^{t} \left(Y_s - \mathbf{X_t}^T \theta_\mathbf{t} \right)^2$$

The *off-line* solution to this equation is found as

$$\hat{\theta}_\mathbf{t} = \mathbf{R_t}^{-1} \mathbf{h_t}$$

where

$$\mathbf{R_t} = \sum_{s=1}^{t} \mathbf{X_s}\mathbf{X_s^T}$$

$$\mathbf{h_t} = \sum_{s=1}^{t} \mathbf{X_s} Y_s$$

and from this the recursive updating formulas for $\mathbf{R_t}$ and $\mathbf{h_t}$ are easily derived

$$\mathbf{R_t} = \sum_{s=1}^{t-1} \mathbf{X_s}\mathbf{X_s^T} + \mathbf{X_t}\mathbf{X_t^T} = \mathbf{R_{t-1}} + \mathbf{X_t}\mathbf{X_t^T}$$

$$\mathbf{h_t} = \sum_{s=1}^{t-1} \mathbf{X_s} Y_s + \mathbf{X_t} Y_t = \mathbf{h_{t-1}} + \mathbf{X_t} Y_t$$

And thereby, the updating formula for $\hat{\theta}$ is found to be

$$\hat{\theta}_\mathbf{t} = \hat{\theta}_\mathbf{t-1} + \mathbf{R_t^{-1}}\mathbf{X_t} \left[Y_t - \mathbf{X_t^T}\hat{\theta}_\mathbf{t-1} \right]$$

To avoid the expensive calculations of matrix inversion the matrix $\mathbf{P_t} = \mathbf{R_t}^{-1}$ is introduced and by the matrix inversion lemma

$$\left[\mathbf{A} + \mathbf{BCD} \right]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B} \left[\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C} \right]^{-1} \mathbf{DA}^{-1}$$

an updating formula for $\mathbf{P_t}$ is found to be

$$\mathbf{P_t} = \mathbf{P_{t-1}} - \frac{\mathbf{P_{t-1}}\mathbf{X_t}\mathbf{X_t}^T\mathbf{P_{t-1}}}{1 + \mathbf{X_t}^T\mathbf{P_{t-1}}\mathbf{X_t}}$$

At last the gain matrix, $\mathbf{K_t}$ is introduced as

$$\mathbf{K_t} = \mathbf{R_t^{-1}X_t} = \frac{\mathbf{P_{t-1}X_t}}{1 + \mathbf{X_t^T P_{t-1} X_t}}$$

Thus, the RLS algorithm can now be written as

$$\hat{\theta}_\mathbf{t} = \hat{\theta}_\mathbf{t-1} + \mathbf{K_t}\left[Y_t - \mathbf{X_t}^T\hat{\theta}_\mathbf{t-1}\right]$$
$$\mathbf{K_t} = \frac{\mathbf{P_{t-1}X_t}}{1 + \mathbf{X_t}^T\mathbf{P_{t-1}X_t}}$$
$$\mathbf{P_t} = \mathbf{P_{t-1}} - \frac{\mathbf{P_{t-1}X_tX_t}^T\mathbf{P_{t-1}}}{1 + \mathbf{X_t}^T\mathbf{P_{t-1}X_t}}$$

Although the RLS algorithm provides parameters that adapt to the changes in the system, it still assumes that the system behaves linearly on the time interval from $t = 0$ to $t = t$ and provides parameter estimates that are consistent in the time space and converge to a point in parameter space.

3.1.1 Recursive Least Squares with forgetting factor

In many cases, it is desirable to have time varying parameters instead of the time consistent parameters provided by the RLS algorithm. The forgetting factor or exponential forgetting technique is a simple extension to the RLS algorithm and handles time varying parameters by discounting old prediction errors in the loss function, $S(\theta)$. In other words, the prediction error in the most recent observations weights more than the older one, when estimating the parameters. Although the change to the algorithm seems small, this leads to drastic changes to the properties of the algorithm. The parameter estimates do not converge anymore and instead, the parameter sequence can be described as a stochastic process. This process is non-Gaussian but the parameter estimation error will still have a mean of 0.

The algorithm for RLS with exponential forgetting still involves finding a solution to the *weighted least squares estimator*

$$\hat{\theta}_\mathbf{t} = \arg\min_\theta S_t(\theta)$$

but now a weight function has been added to the loss function so

$$S_t(\theta) = \sum_{s=1}^{t} \beta(t,s)\left(Y_s - \mathbf{X_s^T}\theta\right)^2$$

where it is assumed that the sequence of weights satisfies

$$\beta(t,s) = \lambda(t)\beta(t-1,s) \quad 1 \leq s \leq t-1$$
$$\beta(t,t) = 1$$

which means that

$$\beta(t,s) = \prod_{j=s+1}^{t} \lambda(j)$$

That is, the weight of the squared residual at time s in the computation of the parameter estimates at time t is the product of all intermediate weighting factors.
The solution for $\hat{\theta}_\mathbf{t}$ is still

$$\hat{\theta}_\mathbf{t} = \mathbf{R_t^{-1} h_t}$$

but now

$$\mathbf{R_t} = \sum_{s=1}^{t} \beta(t,s) \mathbf{X_s X_s^T}$$

$$\mathbf{h_t} = \sum_{s=1}^{t} \beta(t,s) \mathbf{X_s} Y_s$$

so the updating formulas become

$$\mathbf{R_t} = \lambda(t) \mathbf{R_{t-1}} + \mathbf{X_t X_t^T}$$
$$\mathbf{h_t} = \lambda(t) \mathbf{h_{t-1}} + \mathbf{X_t} Y_t$$

From this an algorithm for RLS with a forgetting structure is derived in the same notation as before as

$$\hat{\theta}_\mathbf{t} = \hat{\theta}_\mathbf{t-1} + \mathbf{K_t} \left[Y_t - \mathbf{X_t}^T \hat{\theta}_\mathbf{t-1} \right]$$

$$\mathbf{K_t} = \frac{\mathbf{P_{t-1} X_t}}{1 + \mathbf{X_t}^T \mathbf{P_{t-1} X_t}}$$

$$\mathbf{P_t} = \frac{1}{\lambda(t)} \left[\mathbf{P_{t-1}} - \frac{\mathbf{P_{t-1} X_t X_t}^T \mathbf{P_{t-1}}}{\lambda(t) + \mathbf{X_t}^T \mathbf{P_{t-1} X_t}} \right]$$

If $\lambda(t) = \lambda$ then $\beta(t,s) = \lambda^{t-s}$ and if $0 < \lambda < 1$ then λ is called the forgetting factor. Then the squared errors are weighted exponentially and then the number of effective observations can be found as.

$$T_0 = \frac{1}{1-\lambda}$$

From this can be seen that forgetting factor $\lambda = 1$ is the same as not including a forgetting factor.

3.1.2   Locally Weighted Polynomial regression

Estimating a complicated non-linear relationship between variables can be a tricky business. However if the relationship is linear or a low order polynomial the problem is much more convenient to deal with. One would simply solve the weighted least squares problem

$$\arg\min_{\theta} \frac{1}{N} \sum_{s=1}^{N} w_s(\mathbf{x})(Y_s - \theta)^2$$

One way of estimating non-linear relationship between variables is to use LWPR. By assuming the relationship to be of low order on a small interval one can easily find the relationship on that interval by regression. The size of the interval or the bandwidth, h ($0 \leq h \leq 1$), is given as the proportion of total observations used to estimate the relationship in each point. In other words when a bandwidth of 0.2 is chosen it means that the

20% of the observations that are closest to each point are taken into consideration when the relationship between the variables is estimated around that point. The bandwidth is to be chosen as one that is assumed appropriate for the problem at hand. One must find an appropriate trade-off between the bias and the variance of the resulting model. For the models constructed here a bandwidth of $h = 0.3$ was used as estimates were desired for as small interval as possible in each point and $h = 0.3$ was the smallest one which did not result in singular matrices at any time, i.e. the smallest h which resulted in enough dissimilarity between observations so an estimation could be made. It ought to be mentioned that in order to get a good estimation from LWPR, fairly large data set has to be available. Fortunately this is the case here and therefore LWPR is used to estimate various relationship between variables in the model. In those analysis, a tri-cube kernel was used as a weight function i.e.

$$w(u) = \left\{ \begin{array}{cc} (1 - u^3)^3 & u \in [0; 1) \\ 0 & u \in [1; \infty) \end{array} \right.$$

where u is the relative distance between the point, which the relationship is to be estimated around and other points within the bandwidth.

3.1.3 k-step ahead prediction

In order to use the RLS-algorithm to predict more than one step ahead in time, the pseudo prediction error is used and it is defined as

$$\tilde{Y}_{t|t-k}^{pseudo} = Y_t - \mathbf{X_{t-k}}^T \hat{\theta}_{\mathbf{t-1}}$$

And the k-step prediction is calculated as

$$\hat{Y}_{t+k|t} = \mathbf{X_t}^T \hat{\theta}_{\mathbf{t}}$$

So predictions are not made for 1 hour at a time and repeated k times, but one prediction is made for k time steps ahead. Therefore different parameters are found for every k.

3.1.4 Performance comparison

Numerous ways exist to compare the performance of different models. Two of them are used here to compare the different models, the coefficient of determination (R^2) and the root mean squared error ($RMSE$). R^2 is a measurement of how much of the variability in the data is accounted for in the model and is calculated as

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T}$$

where SS_T is the total sum of squares, SS_R is the explained sum of squares and SS_E is the sum of squared residuals

$$SS_T = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$SS_R = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

$$SS_E = \sum_{i=1}^{n} (y_i - \hat{y})^2$$

$RMSE$ is a measure of how much the model deviates from the real values on average and therefore it decreases if a term is added to the model that provides the estimator with information that otherwise would not be accounted for in the prediction. $RMSE$ is calculated as

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(Y_t - \hat{Y}_t \right)^2}$$

3.2 Time invariant models of the spot price

To start with a time invariant linear autoregressive model of the spot price is built for the whole year. So the parameters, ϕ, of the model are assumed to be constant and are estimated off-line using a part of the data set and then tested on the remaining part. 50% or the first 6 months of the observations were taken for model estimation and then performance tested on the latter half of the year. Now in order to estimate which past observations are to be used in the model, a look is taken at the autocorrelation function (ACF) and partial autocorrelation functions of the series. The ACF shows how
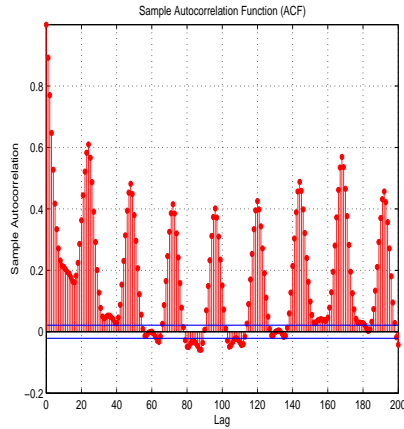
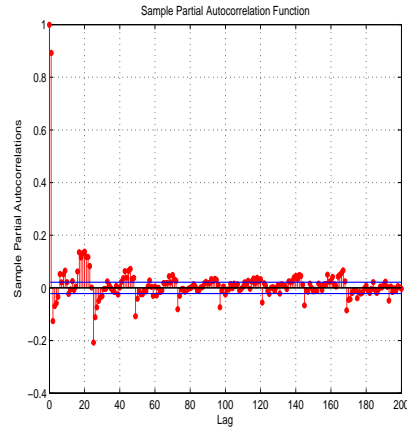

Figure 3.1: Autocorrelation function of the spot price

Figure 3.2: Partial autocorrelation function of the spot price

the spot price is correlated to earlier prices while the partial ACF shows how much of that correlation is not explained by observations closer to the time t. From those plots, it is concluded that disregarding the seasonal variation, the autoregressive part of the model is dominated by a 2nd order model. This is concluded since the autocorrelation of the 1st two lags are considerably higher than the confidence interval. For the same reasons it is decided to include a daily seasonal parameter and a weekly seasonal parameter. The model is therefore on the form

$$Y_t = -\phi_0 - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \Phi_1 Y_{t-24} - \Phi_2 Y_{t-168} + \sum_{i=1}^{q} \omega_i X_i + \epsilon_t$$

Where X_i is external input i. Different external signals were also tried and the performance estimates for 1 hour predictions of the models tried can be viewed in *table 3.1*. As can

External input	R^2	R^2_{ad}	$RMSE$
-	0.7694	0.7689	50.9554
Time	0.7713	0.7708	50.7395
Wind production	0.7751	0.7746	50.3160
Wind penetration	0.7783	0.7778	49.9601
Time & Wind Prod.	0.7762	0.7755	50.2013
Time & Wind pen.	0.7793	0.7786	49.8541

Table 3.1: Performance of the AR model with different external signals

be seen, the model where the predicted wind penetration and the time of day is included gives the best result. This model gives a $R^2_{ad} = 0.7786$ and $RMSE = 49.8541$.
Several possibilities for making the model better exist. It was tried to make a non-linear autoregressive model which did not outperform the AR model and to split the data into smaller intervals and make a model for each of them. By splitting the data set into the 12 months and then combine months that had similar behavior, 5 data sets were obtained each containing $2 - 4$ months. These data sets were

1. January, February, March, April

2. May, December

3. June - Juli

4. August -September

5. October - November

It seems strange that May and December are similar especially since they are not alike the months before and after. This is most likely a coincident since no logical explanation can be found and when a look is taken at other years a split after alike months does not give this result. However due to other flaws in this split, it was decided to leave it at that and focus on the performance in other months and then pursue other and better ways of modeling the spot price. The data for May and December are non the less included in the performance estimate. Since the split leads to discontinuity in the data set, May was taken as a estimation set and December for validation. For other data sets 75% of the data were taken for estimation and the remaining 25% were taken for validation. Models for each period is now constructed with the time of the day and predicted wind power prediction as external signals. Their performance estimates can be viewed in *table 3.2*. Plots of the average spot prices in each month in each set and the average can be viewed in *figures B.1 - B.6* in Appendix B. The models perform better now apart from that the

Data set	R^2	R^2_{ad}	$RMSE$
1	0.8268	0.8249	41.4609
2	0.4412	0.4242	31.4678
3	0.7547	0.7470	26.9791
4	0.8973	0.8940	31.9708
5	0.7276	0.7190	37.4439

Table 3.2: Performance of the AR for the splittet data sets

R^2 coefficients for the $2nd$ data set is extremely low. This is due to the discontinuity of the data set mentioned earlier. The performance is especially good during the late summer which can be explained by the much lower variance in the prices during the summer.
Splitting the data set up like this raises the question of how to make the split. Making a split by months has some serious disadvantages since the month is not a direct factor on the spot prices. For example in [3] a data set for 2005 is splittet up by the same principles and results in a very different month combinations. Therefore much more data is need for making these models credible in the real world and also a split by some other more relevant factor should be considered. Other measures that could be taken in order to enhance the performance of the model could be for example to differentiate the time series in order to make it more stationary.

3.3   Adaptive models of the spot price

One way of dealing with non-stationarity is to make the models adaptive, so parameters are updated on-line before every prediction. Models of this kind are now implemented for the spot prices in two ways.

3.3.1   Recursive Least Squares model

First the off-line model constructed in previous section is made adaptive and instead of including the external inputs on a polynomial form, a non-parametric estimate of the spot price is included in the model. This non-parametric estimate is found by LWPR with the hour and wind power prediction as inputs. Furthermore a MA-part is added to the model since there is a little exponential decay in the partial ACF for the price. The residuals included in the model are $\epsilon_{t-1}, \epsilon_{t-24}$ so the MA-part is of order one and a daily seasonal parts of order one. So the model is on the form

$$Y_t = -\phi_{t,0} - \phi_{t,1}Y_{t-1} - \phi_{t,2}Y_{t-2} - \Phi_{t,1}Y_{t-24} - \Phi_{t,2}Y_{t-168} + \omega_t U_t + \theta_{t,1}\epsilon_{t-1} + \Theta_{t,1}\epsilon_{t-24} + \epsilon_t$$

The next step is to choose a suitable forgetting factor. This is done by fitting a model to the data with various forgetting factors, λ. In *tables 3.3 & 3.4* the performance estimates of the model for the λ-values tried can be seen. Performance was estimated for $1, 12, 24$ and 36 hour predictions. As can be seen from the tables, a forgetting factor of 0.999 is

λ	R^2	R^2_{ad}	RSE
1.0000	0.8230	0.8228	43.8067
0.9990	0.8235	0.8233	43.7515
0.9950	0.8204	0.8202	44.1348
0.9900	0.8167	0.8165	44.5823
0.9750	0.7993	0.7991	46.6472
0.9500	0.7374	0.7372	53.3576

Table 3.3: Performance of the RLS model with $k = 1$

the one that gives the best results. This means that previous 1000 observations are used to estimate the parameter at each time. So spot prices from past 6 weeks are the ones that are mostly related to the spot price at any given time.

λ	k=12			k=24			k=36		
	R^2	R^2_{ad}	$RMSE$	R^2	R^2_{ad}	$RMSE$	R^2	R^2_{ad}	$RMSE$
1.0000	0.7439	0.7436	52.6993	0.7388	0.7385	53.2158	0.7345	0.7342	53.6532
0.9990	0.7472	0.7470	52.3524	0.7421	0.7418	52.8860	0.7372	0.7369	53.3791
0.9950	0.7380	0.7377	53.3040	0.7361	0.7358	53.4968	0.7342	0.7339	53.6896
0.9900	0.7379	0.7376	53.3089	0.7383	0.7381	53.2661	0.7388	0.7385	53.2233
0.9750	0.7288	0.7285	54.2272	0.7358	0.7355	53.5234	0.7427	0.7424	52.8196
0.9500	0.7101	0.7098	56.0697	0.7088	0.7084	56.1959	0.7075	0.7071	56.3222

Table 3.4: Performance of the RLS model for $k = 12, 24, 36$

Numerous other versions of the model were tried out in order to find the best one, all which were worse than this one.

3.3.2   Non-parametric model with a error regression

The other way of constructing an adaptive model for the spot price that was tried, is to make a non-parametric estimate of the spot price with more inputs than used before, and then regress on the error of that estimate in order to account for effects from other factors. The non-parametric estimate was as before obtained by LWPR and the model that gave the best result was

$$\hat{Y}_t^{np} = f(Hour, Month, Wind\,prediction)$$

Next the residuals of this non-parametric model were investigated. By looking at plots of the residuals from the model, shown in *figure 3.3*, one notices that they behave very similar to the spot prices it self, just on a smaller scale. Another thing that is interesting is that the autocorrelation plots show a much less dominant seasonal part than before, indicating that the non-parametric model accounts for a large proportion of the seasonal variation. All this is a good indication of that the non-parametric model provides a good prediction basis for the spot price but obviously other factors have to be accounted for. Two possibilities of introducing a correcting term in the model are now considered. The first one is to regress on the error itself and the other one is to regress on the difference between errors. By taking the difference of the error time series once, the resulting series will be the one shown in *figure 3.4*. This series is much closer to be stationary and therefore should be easier to predict the differentiated values.
 As it turns out the differentiated series gives a better result and therefore it is decided to go on with that model. The choice of which residuals to include in the model is made after inspecting the ACF and the partial ACF of the differentiated residuals. The ACFs can be seen in *figures 3.5 & 3.6* and it is concluded that including $\epsilon_{t-1}, \epsilon_{t-24}$ and ϵ_{t-48} is appropriate. The model for the spot price then has the form

$$Y_t = \hat{Y}_t^{np} + \epsilon_{t-1} + \theta_{t,1}(\epsilon_{t-1} - \epsilon_{t-2}) + \theta_{t,2}(\epsilon_{t-24} - \epsilon_{t-25}) + \theta_{t,3}(\epsilon_{t-48} - \epsilon_{t-49}) + \epsilon_t$$

 The forgetting factor is chosen in the same manner as before and the results from that procedure can be seen in *tables 3.5 & 3.6* As for the RLS-model, the best performance is achieved by setting $\lambda = 0.999$ leading to the same conclusions as before about effective observations in the parameter estimation. Even though the performance of the model for one hour predictions are not quite as good as for the RLS-model the performance for predictions further ahead is better. This is because the model is not nearly as dependent
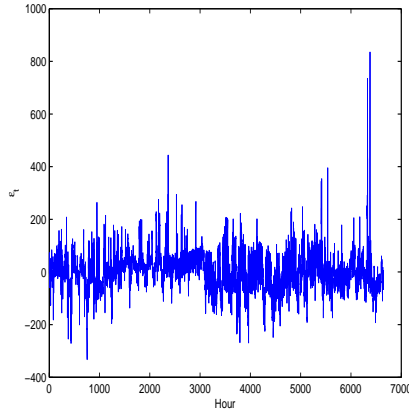
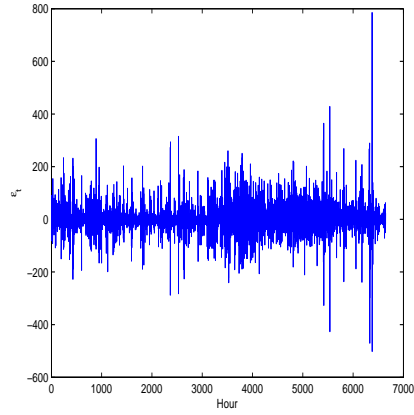Figure 3.3: Time series plot of the residuals from the non-parametric model

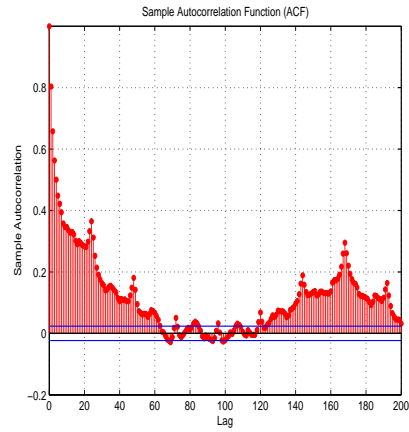Figure 3.4: Time series plot of the residuals from the non-parametric model differentiated

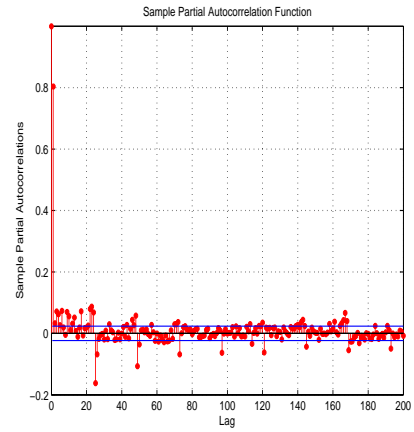Figure 3.5: ACF of the differentiated residuals

Figure 3.6: PACF of the differentiated residuals

λ	R^2	R^2_{ad}	RSE
1.0000	0.8151	0.8151	45.0567
0.9990	0.8153	0.8153	45.0367
0.9950	0.8021	0.8020	46.3296
0.9900	0.7736	0.7735	49.0026
0.9750	0.6556	0.6555	59.0141
0.9500	0.3332	0.3330	81.9700

Table 3.5: Performance of the non-parametric model for $k = 1$

λ	k=12			k=24			k=36		
---	R^2	R^2_{ad}	RSE	R^2	R^2_{ad}	RSE	R^2	R^2_{ad}	RSE
1.0000	0.7861	0.7860	48.5758	0.7903	0.7902	48.1000	0.7710	0.7709	50.2667
0.9990	0.7877	0.7876	48.3975	0.7918	0.7917	47.9275	0.7726	0.7725	50.0851
0.9950	0.7827	0.7826	48.9660	0.7673	0.7672	50.6682	0.7679	0.7679	50.5967
0.9900	0.7640	0.7639	51.0287	0.7780	0.7779	49.4903	0.7496	0.7495	52.5588
0.9750	0.6576	0.6575	61.4574	0.7287	0.7286	54.7117	0.6437	0.6436	62.6910

Table 3.6: Performance of the non-parametric model for $k = 12, 24, 36$

on data from recent past as the RLS-model is. So one must conclude that this model is better suited for predicting longer ahead in time.

3.4   Prediction with adaptive models

Predictions are now made 12, 24 and 36 hours ahead in time with the non-parametric model as it gave better results for long time predictions.

3.4.1   12 hour predictions

January through March are taken as initialization months and then predictions are made for the remaining months of the year. In *figures 3.7 & 3.8* the predictions are shown for July and November respectively. Along with the predictions, the actual values along with 95% confidence interval of the predictions are shown. The predictions seem to be rather good since they imitate the observed data very well. The daily fluctuations of the spot price are well detectable in the plots since approximately every 24 hours the price goes down. Furthermore the weekly changes are also detectable although they are a bit more difficult to see.

In *figure 3.9* the evolvement of the model parameters is shown. The seasonal variation in
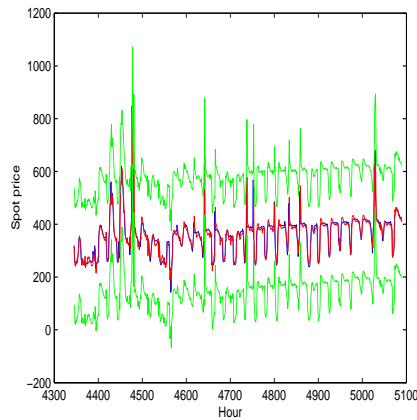
Figure 3.7: Predicted values 12 hours ahead in Juli and 95% confidence interval
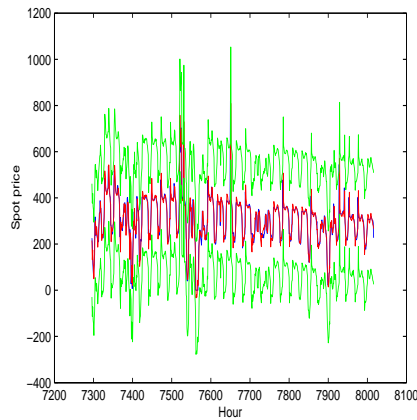
Figure 3.8: Predicted values 12 hours ahead in November and 95% confidence interval
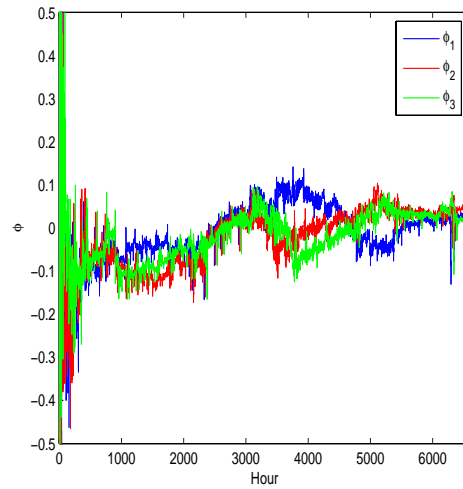
Figure 3.9: Evolvement of parameters for 12 hour predictions

the variability of the spot price is well detectable form this plot as during the summer the parameters are much higher than in the winter time. This is fully in accordance with what was previously said about how much more stable the spot price is during the summer. Since the prices are more stable in the summer, the systematic difference between the real value and the non-parametric estimate is also more stable and therefore more correlated with previous errors. This results in higher parameters.

A histogram and a scatter plot of the residuals against fitted values are shown in *figures*
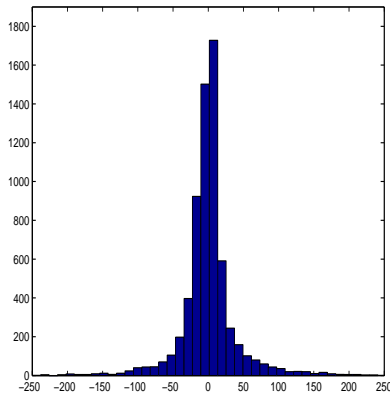
Figure 3.10: Residual histogram for 12 hour predictions
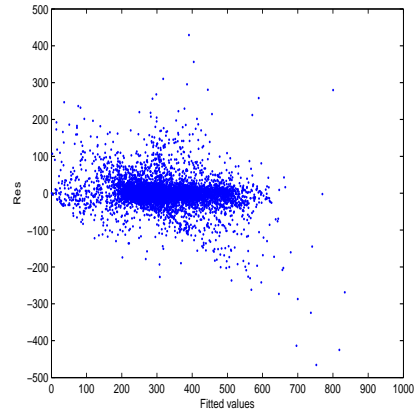
Figure 3.11: Residuald vs. Fitted values for 12 hour predictions

3.10 & 3.11. The residuals seem to be quite close to be normally distributed but not entirely as the histogram is more heavy tailed than a normal distribution. The scatter

plot shows that there is no systematic error in the model as there is no correlation between the fitted values and the residuals. Furthermore the mean error is 0.03 so negative and positive residuals are equally likely.

3.4.2   24 hour predictions

For the 24 hour predictions the same 3 months are taken for initialization as before and predictions made for the rest of the year. Predictions for the same periods as were shown for the 12 hour predictions are shown in *figures 3.12 & 3.13*. The predictions are still quite good and are close to the real values at all times. The confidence interval is very much alike the one for the 12 hour predictions. This is in accordance with the very similar performance estimates of the model on different time horizons. Although the parameters change and therefore the system in the error, the model can almost equally well predict for a longer time horizons and therefore does the confidence interval not become wider. The evolvement of the parameters is shown in *figure 3.14*. The parameters show somewhat
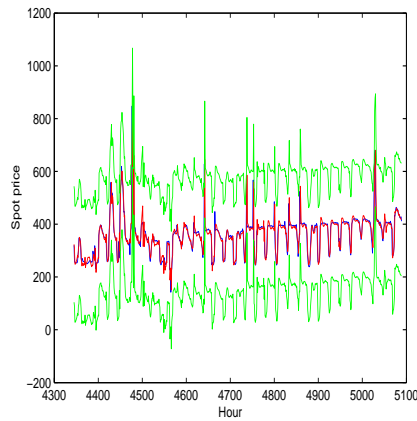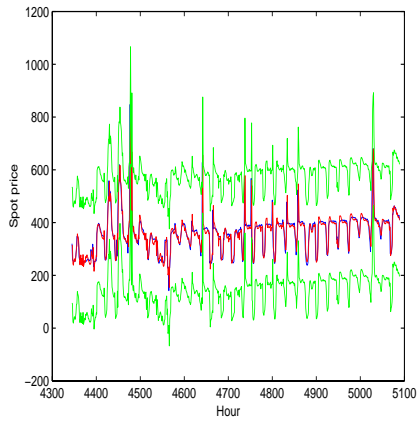

Figure 3.12: Predicted values 24 hours ahead in Juli and 95% confidence interval
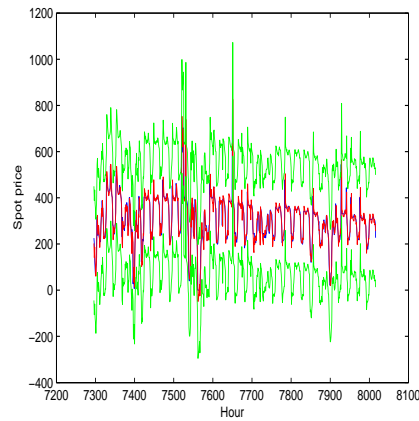
Figure 3.13: Predicted values 24 hours ahead in November and 95% confidence interval

different behavior now as ϕ_1 is negative here while it was positive during the summer for the 12 hour predictions. The parameters are still larger during the summer, for the same reasons as were mentioned before. A histogram of the residuals is shown in *figure 3.15* and shows very similar distribution properties as the 12 hour one. Viewing the scatter plot of the residuals vs. fitted values also leads to the same conclusions as were drawn for the 12 hour residuals. So the residuals are close to normally distributed around 0 but with a larger tail and no correlation between residuals and fitted values is detected.

3.4.3   36 hour predictions

At last predictions are made for 36 hours ahead. The initialization period is taken to be the same and predictions are shown for the same periods in *figures 3.17 & 3.18*. The predictions are still quite good and are close to the real values at all times. The confidence interval is still very similar to the previous ones, for the same reasons as were mentioned earlier.
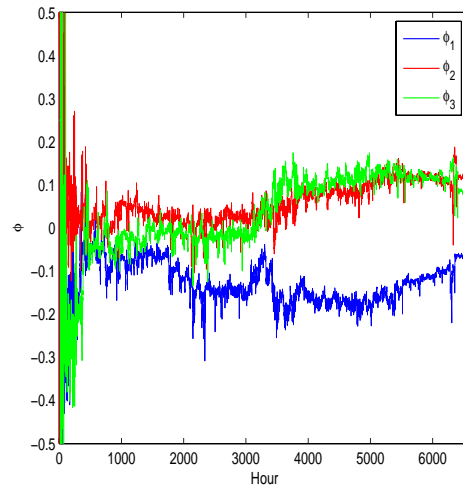
Figure 3.14: Evolvement of parameters for 24 hour predictions
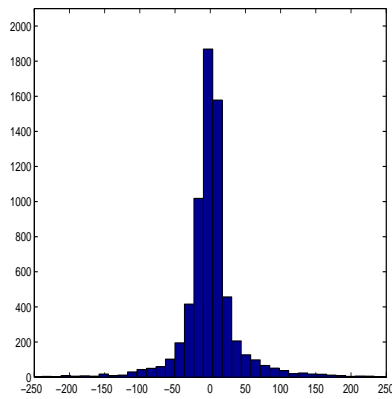
Figure 3.15: Residual histogram for 24 hour predictions
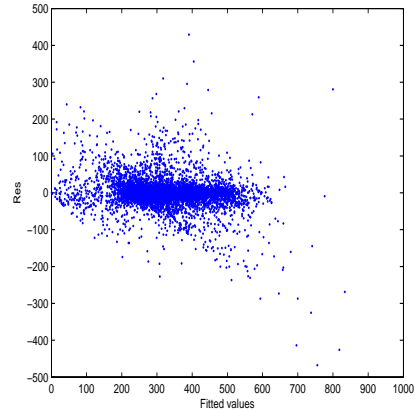
Figure 3.16: Residuals vs. Fitted values for 24 hour predictions

The evolvement of the model parameters is quite similar for the 12 hour predictions and the 36 hour predictions. The only difference is that the parameters for the 36 hour predictions are a bit smaller than the 12 hour ones. One would think that when predicting further ahead, predictions will rely more on the non-parametric prediction and less on the errors and therefore this shrinkage in the parameters seems reasonable. The same plots of the residuals as before are shown in *figures 3.20 & 3.21* and they show the same characteristics as the other predictions.

Figure 3.17: Predicted values
36 hours ahead in Juli and 95%
confidence interval
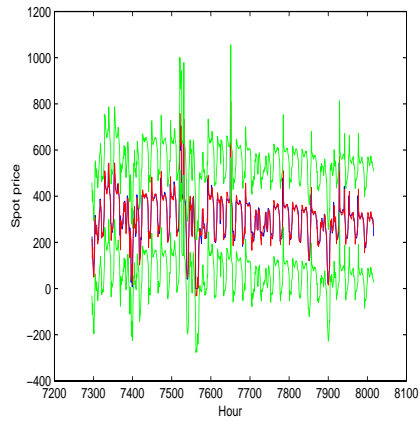
Figure 3.18: Predicted values
36 hours ahead in November
and 95% confidence interval
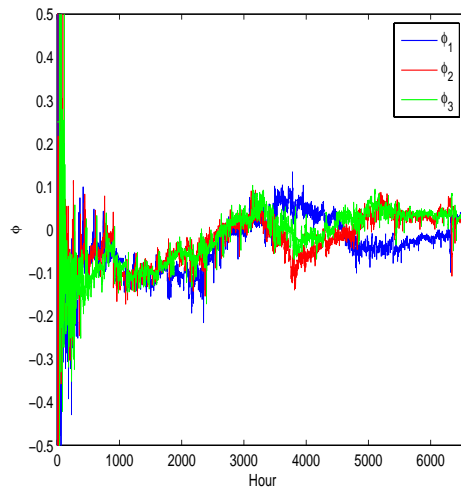
Figure 3.19: Evolvement of parameters for 36 hour predictions
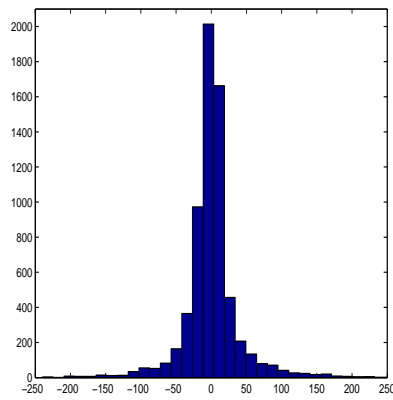
Figure 3.20: Residual histogram for 36 hour predictions
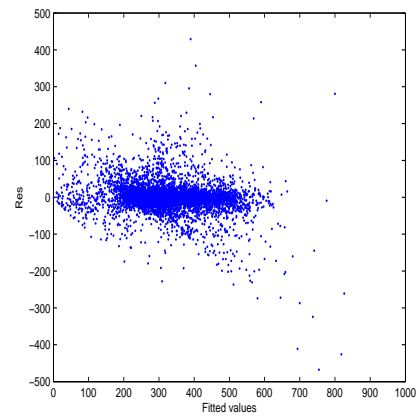
Figure 3.21: Residuals vs. Fitted values for 36 hour predictions

Conclusion

4.1 General conclusions

There is no doubt that the use of wind energy of in the western danish electricity system results in a lower energy prices, compared to if other available energy production methods are used. The difference is considerable, even when it is taken into account that the wind farms receive a government support. By having a better prediction model of the spot prices, chances are that the effects would be even more positive for the danish power companies and therefore the danish public, as those model would help agents when bidding on the Nordpool spot market.

Numerous methods, both static and dynamic, have been tried to predict the electricity spot price on Nordpool. Rather good predictions could be made with a static model if separate models were constructed for $2-4$ months at a time. However as mentioned before, split by months is probably not the ideal way of dividing the data, since climate, which in many ways affects the spot price, can vary a lot between years in the same month. So if a time invariant model is to be used, some other more relevant splitting probably would be of better use. What comes to mind is for example a SETAR-model where a split would be made by previous spot prices or even water level in the hydro power plants' reservoirs in Norway and Sweden.

The adaptive models made gave a better overall performance and predictions made by them were rather good. It is well possible that those models could be of some help when making bids into Nordpool. From the estimated performance, one can conclude that for very few hours, a Recursive least squares model could be of good use, while for more practical predictions with a time horizon of $12-36$ hours a non-parametric model, corrected with a MA-part is more useful.

4.2 Future work

It is very likely that the performance of the adaptive models can be improved considerably by also including other external inputs. Some of the inputs that could be of use are the price of CO_2 release kvota and other things necessary to produce electricity by heat. Furthermore since the methods used here seem to be fitting for predictions of this kind, similar method are likely to provide at least good basis for predictions of the price in other markets in Nordpool, such as the "Regulerings market".

This will be the subject of the autor's master thesis which will be worked on from August 2007 to June 2008.

Estimated price difference for other periods

Estimated price difference for other individual periods than displayed in the report along with the naive analysis, where the demand is not taken into account can be viewed in *figures A.1 - A.12*
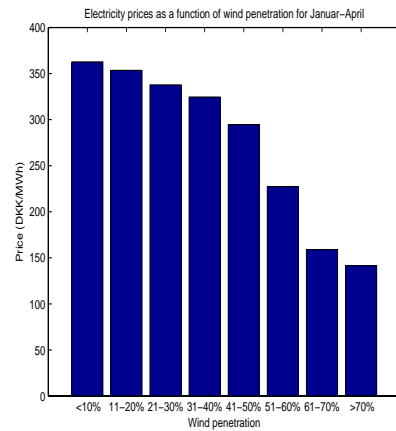

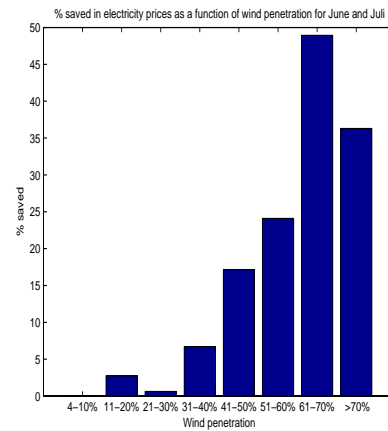Figure A.1: Average spot prices for a given wind penetration(Jan-Apr)

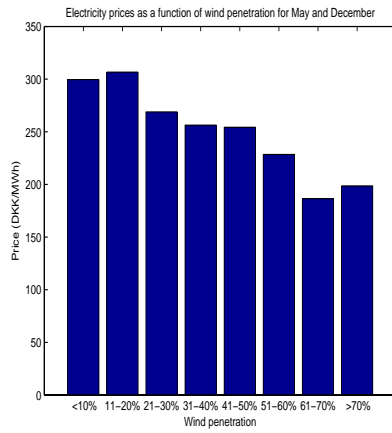Figure A.2: Average price difference in % for a given wind penetration(Jan-Apr)

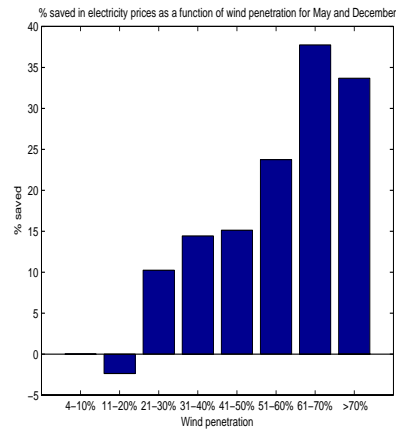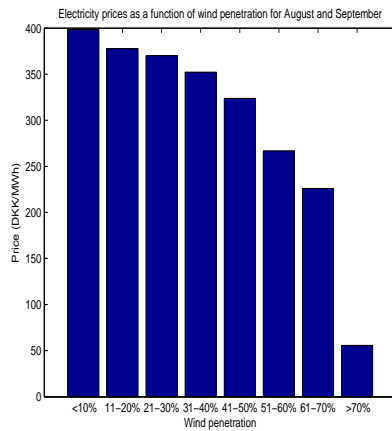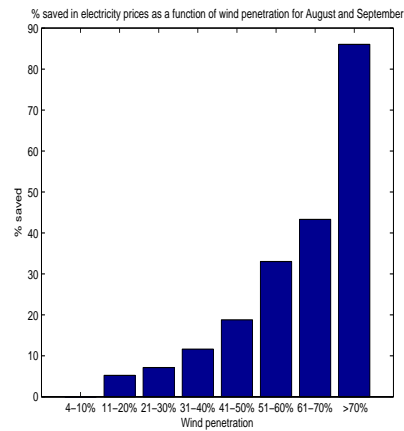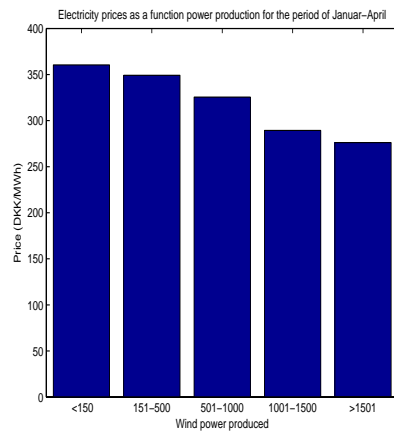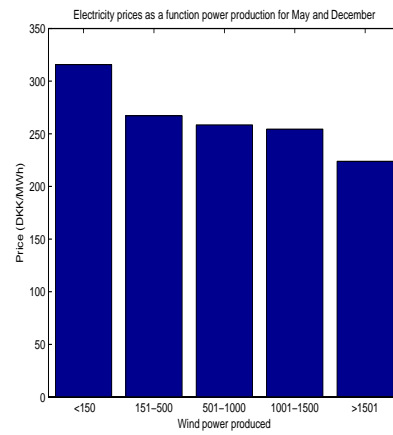Figure A.3: Average spot prices for a given wind penetration(May/Dec)

Figure A.4: Average price difference in % for a given wind penetration(May/Dec)

Figure A.5: Average spot prices for a given wind penetration(Aug/Sep)

Figure A.6: Average price difference in % for a given wind penetration(Aug/Sep)

Figure A.7: Average spot prices for a given wind power production(Jan-Apr)

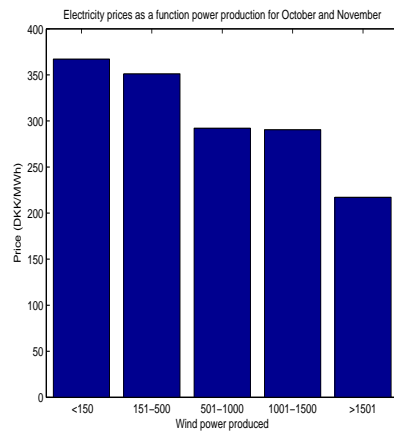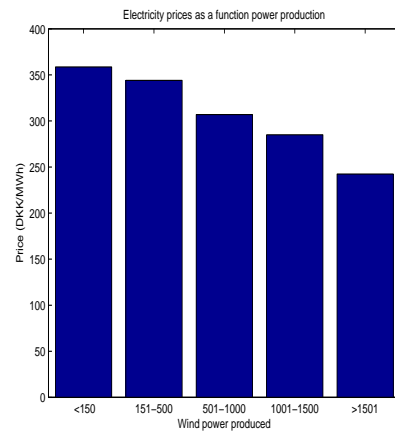Figure A.8: Average spot prices for a given wind power production(May/Dec)
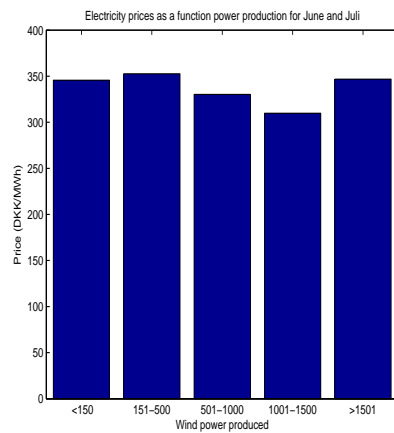
Figure A.9: Average spot prices for a given wind power production(Jun/Jul)
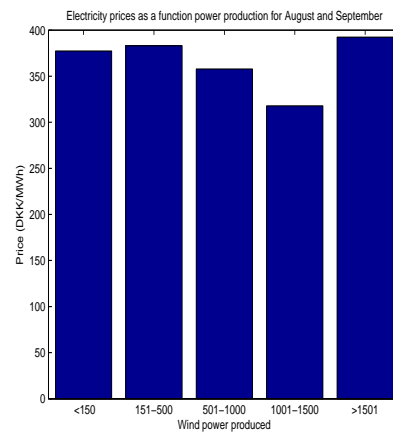
Figure A.10: Average spot prices for a given wind power production(Aug/Sep)

Figure A.11: Average spot prices for a given wind power production(Oct/Nov)

Figure A.12: Average spot prices for a given wind power production(2006)
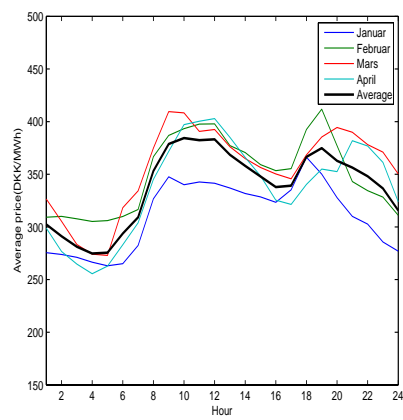
Data spilt and combinations for AR-models

Figure B.1: Average spot prices for a given period(Jan-Apr)
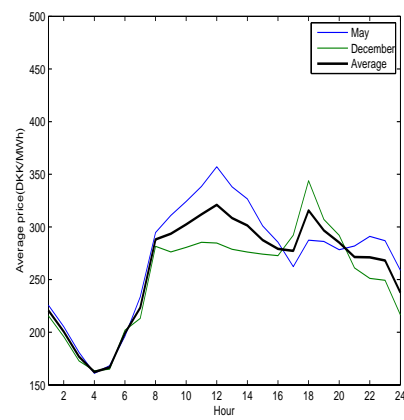
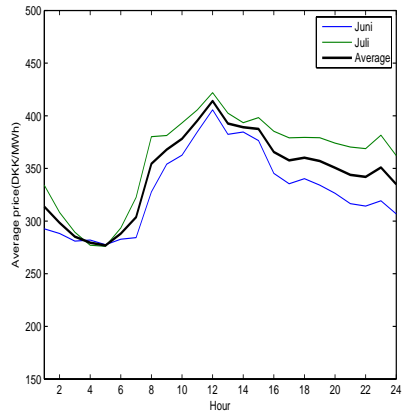Figure B.2: Average spot prices for a given period(May,Dec)

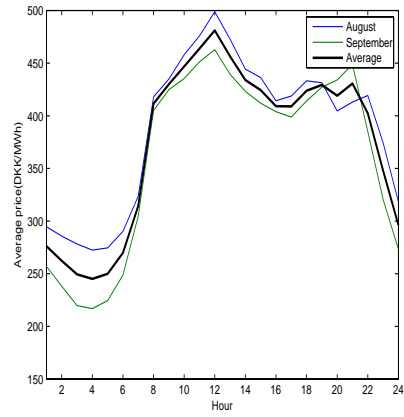Figure B.3: Average spot prices
for a given period(Jun,Jul)

Figure B.4: Average spot prices
for a given period(Aug,Sep)
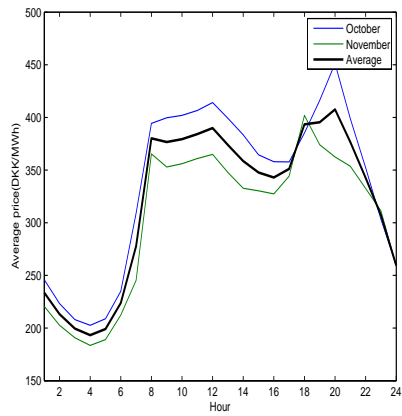

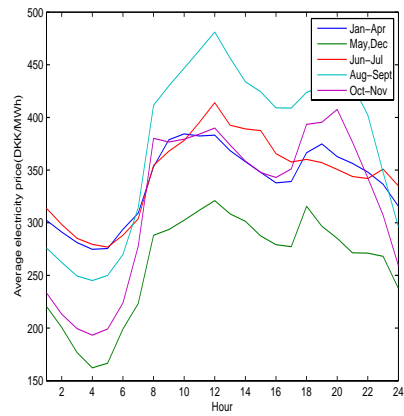Figure B.5: Average spot prices
for a given period(Oct,Nov)

Figure B.6: Average spot prices
for combined data sets(2006)

References

[1] H.Madsen *Time Series Analysis*, IMM, DTU 2001

[2] H.Madsen & J.Holst *Modelling Non-Linear and Non-Stationary Time Series* IMM, DTU, 2000

[3] S.W.Enevoldson, P.A.Østergaard, P.E.Morthorst, R.Moesgaard *Vindkraftens betydning for elprisen i Danmark* IBT-Wind, 2006

[4] C.Schaumburg-Muller, *Modelling the Western Danish Electricity Spot Price Using Time Series Analysis*, report IMM, 2007