Wind Power Forecasting

Henrik Madsen
hm@imm.dtu.dk

Section for Math. Statistics
Technical University of Denmark (DTU)
DK-2800 Lyngby

www.imm.dtu.dk/~hm
Wind Power Forecasting in Denmark
Methods used for predicting the wind power
Configuration example for a large system
Spatio-temporal modelling
Use of several providers of MET forecasts
Uncertainty and confidence intervals
Scenario forecasting
Value of wind power forecasts
Wind Power Forecasting in Denmark

WPPT (Wind Power Prediction Tool) is one of the wind power forecasting solutions available with the longest historic of operational use.

- WPPT has been continuously developed since 1993 – initially at DTU (Technical University of Denmark and since 2006 by ENFOR – in close co-operation with:
  - Energinet.dk,
  - Dong Energy,
  - The ANEMOS projects and consortium (since 2002)
  - DTU (since 2006).
- WPPT has been used operationally for predicting wind power in Denmark since 1996.
- WPPT is now used all over Europe, Australia and North America.

Now in Denmark (DK1): Wind power covers on average about 25 pct of the system load (this is a world record).
Goal: 50 pct of the needed power should come from wind in 2025.
Prediction Performance

New performance measures. 13 large parks. Installed power: 1064 MW. Period: August 2008 - March 2010. MET input is in all cases ECMWF. Criterion used: RMSE.

RMSE

Nominal power: 1064 MW

Forecast horizon (hours)

Normalized power

RPPT–PC
Simple model
Commercial model 1
Commercial model 2
Prediction of wind power

In areas with high penetration of wind power such as the Western part of Denmark and the Northern part of Germany and Spain, reliable wind power predictions are needed in order to ensure safe and economic operation of the power system.

Accurate wind power predictions are needed with different prediction horizons in order to ensure

- (a few hours) efficient and safe use of regulation power (spinning reserve) and the transmission system,
- (12 to 36 hours) efficient trading on the Nordic power exchange, NordPool,
- (days) optimal operation of eg. large CHP plants.

Predictions of wind power are needed both for the total supply area as well as on a regional scale and for single wind farms.

Today also reliable methods for ramp forecasting is provided by most of the tools.
Modelling approach – the inputs

Depending on the configuration WPPT can take advantage of input from the following sources:

- Online measurements of wind power prod. (updated every 5min. – 1hr).
- Online measurements of the available production capacity.
- Online “measurements” of downregulated production.
- Aggregated high resolution energy readings from all wind turbines in the groups defined above (updated with a delay of 3-5 weeks).
- MET forecasts of wind speed and wind direction covering wind farms and sub-areas (horizon 0–48(120)hrs updated 2–4 times a day).
- Forecasted availability of the wind turbines.
- Other measurements/predictions (local wind speed, stability, etc. can be used).
System characteristics

The total system consisting of wind farms measured online, wind turbines not measured online and meteorological forecasts will inevitably change over time as:

- the population of wind turbines changes,
- changes in unmodelled or insufficiently modelled characteristics (important examples: roughness and dirty blades),
- changes in the NWP models.

A wind power prediction system must be able to handle these time-variations in model and system. WPPT employes **adaptive and recursive model estimation** to handle this issue.

Following the initial installation WPPT will automatically calibrate the models to the actual situation.
The power curve model

The wind turbine “power curve” model, \( p^{tur} = f(w^{tur}) \) is extended to a wind farm model, \( p^{wf} = f(w^{wf}, \theta^{wf}) \), by introducing wind direction dependency. By introducing a representative area wind speed and direction it can be further extended to cover all turbines in an entire region, \( p^{ar} = f(\bar{w}^{ar}, \bar{\theta}^{ar}) \).

The power curve model is defined as:

\[
\hat{p}_{t+k|t} = f(\bar{w}_{t+k|t}, \bar{\theta}_{t+k|t}, k)
\]

where
\( \bar{w}_{t+k|t} \) is forecasted wind speed, and
\( \bar{\theta}_{t+k|t} \) is forecasted wind direction.

The characteristics of the NWP change with the prediction horizon. Hence the dependency of prediction horizon \( k \) in the model.

Plots of the estimated power curve for the Hollandsbjerg wind farm (\( k = 0, 12, 24 \) and 36 hours).
The dynamical prediction model

The power curve models are used as input for an adaptively estimated dynamical model, which (as a simple example) leads to the following k-stop ahead forecasts:

\[
\hat{p}_{t+k|t} = a_1 p_t + a_2 p_{t-1} + b \hat{p}^{pc}_{t+k|t} + \sum_{i=1}^{3} \left[ c^c_i \cos \frac{2i\pi h^{24}_{t+k}}{24} + c^s_i \sin \frac{2i\pi h^{24}_{t+k}}{24} \right] + m + e_{t+k}
\]

where \( p_t \) is observed power production, \( k \in [1; 48] \) (hours) is prediction horizon, \( \hat{p}^{pc}_{t+k|t} \) is power curve prediction and \( h^{24}_{t+k} \) is time of day.

Model features include

- multi-step prediction model to handle non-linearities and unmodelled effects,
- the number of terms in the model depends on the prediction horizon,
- non-stationarity are handled by adaptive estimation of the model parameters,
- the deviation between observed and forecasted diurnal variation is model using Fourier expansions.
A model for upscaling

The dynamic upscaling model for a region is defined as:

\[
\hat{p}_{reg}^{t+k|t} = f( \bar{\omega}^{ar}_{t+k|t}, \bar{\theta}^{ar}_{t+k|t}, k ) \hat{p}_{loc}^{t+k|t}
\]

where
\[
\hat{p}_{loc}^{t+k|t} \quad \text{is a local (dynamic) power prediction within the region},
\]
\[
\bar{\omega}^{ar}_{t+k|t} \quad \text{is forecasted regional wind speed}, \quad \text{and}
\]
\[
\bar{\theta}^{ar}_{t+k|t} \quad \text{is forecasted regional wind direction}.
\]

The characteristics of the NWP and \( \hat{p}_{loc}^{t+k|t} \) change with the prediction horizon. Hence the dependency of prediction horizon \( k \) in the model.
This configuration of WPPT is used by a large TSO. Characteristics for the installation:

- A large number of wind farms and stand-alone wind turbines.
- Frequent changes in the wind turbine population.
- Offline production data with a resolution of 15 min. is available for more than 99% of the wind turbines in the area.
- Online data for a large number of wind farms are available. The number of online wind farms increases quite frequently.
Fluctuations at large offshore wind farms have a significant impact on the control and management strategies of their power output.

Focus is given to the minute scale. Thus, the effects related to the turbulent nature of the wind are smoothed out.

When looking at time-series of power production at Horns Rev (160MW) and Nysted (165 MW), one observes successive periods with fluctuations of larger and smaller magnitude.

We aim at building models:

- based on historical wind power measures only...
- ... but able to reproduce this observed behavior
- this calls for **regime-switching models**
Outline

- Regime-switching modeling
  - Regime-switching models relying on an observable process
    - SETAR
    - STAR
  - Regime-switching models relying on an unobservable process
    - MSAR: characteristics
    - MSAR: estimation issue
- Testing
  - Test-case: forecasting at Horns Rev and Nysted
  - Results
The SETAR model

- SETAR: Self-Exciting Threshold AutoRegressive
- The SETAR\((R; p_1, p_2, \ldots, p_R)\) model is given by:

\[
y_t = \theta_0^{(k_t)} + \sum_{i=1}^{p_{k_t}} \theta_i^{(k_t)} y_{t-i} + \sigma_{k_t} \varepsilon_t
\]

where the regime sequence is directly given by the value of the state variable \(z_t\):

\[
k_t = \begin{cases} 
1, & z_t \in ]-\infty; r_1] \\
2, & z_t \in ]r_1; r_2] \\
& \ldots \\
R, & z_t \in ]r_{R-1}; \infty[
\end{cases} \quad (\text{regime } R)
\]

- The variable \(z_t\) is commonly chosen as a linear combination of past measures of the process, i.e. \(z_t = \sum_{k>0} \alpha_k y_{t-k}\). Here, it is such that \(z_t = y_{t-1}\)
- The set of parameters resumes to \(\Theta_e = (\theta, r, \sigma)^\top\)
The STAR model

- STAR: Smooth Transition AutoRegressive
- A STAR \( R; p_1, p_2, \ldots, p_R \) model is given by:

\[
y_t = \sum_{k=1}^{R-1} \left( \left( \theta_0^{(k)} + \sum_{i=1}^{p_k} \theta_i^{(k)} y_{t-i} \right) \tilde{G}_k(z_t) + \left( \theta_0^{(k+1)} \sum_{j=1}^{p_{k+1}} \theta_j^{(k+1)} y_{t-j} \right) G_k(z_t) \right) + \varepsilon_t
\]

where \( G_k(z) \) is a smooth function that controls the transition from regime \( k \) to regime \( k + 1 \).
- \( G_k(z) \) is chosen to be a logistic function, which is defined by a location parameter \( c_k \) and a shape parameter \( \gamma_k \):

\[
G_k(z) = \left( 1 + \exp \left( -\gamma_k (z - c_k) \right) \right)^{-1}, \quad \gamma_k > 0
\]
- Again, the state variable \( z_t \) is such that \( z_t = y_{t-1} \)
- The set of parameters for that class of models is \( \Theta_s = (\theta, \Gamma, c, \sigma_\varepsilon)^\top \)
The basic idea of MSAR (Markov Switching AutoRegressive) models is that the regime sequence \( \{s_t\} \) is governed by an unobservable process.

An MSAR model for \( \{y_t\} \) then writes

\[
y_t = \theta_0^{(s_t)} + \sum_{i=1}^{p_{s_{t}}} \theta_i^{(s_t)} y_{t-i} + \sigma_s t \varepsilon_t
\]

\( \{s_t\} \) is assumed to follow a first order Markov chain:

\[
P(s_t = j | s_{t-1} = i, s_{t-2}, \ldots, s_0) = P(s_t = j | s_{t-1} = i), \quad \forall i, j, t
\]

such that all probabilities governing the switches are summarized by the so-called transition matrix \( P = \{p_{ij}\}_{i,j=1,...,R} \).

The set of model parameters is thus \( \Theta_m = (\theta^{(1)}, \ldots, \theta^{(R)}, \sigma, P)^\top \)
Tests - Setup

- Two offshore wind farms
  - Horns Rev: 160 MW (11 months of data)
  - Nysted: 165.5 MW (9 months of data)

- Raw power measurements at a 1-second rate

- Time-series of wind power averaged at a 1-, 5-, and 10-minute rate

- Each farm is modelled by a single wind turbine. Its power output is the average of available turbines’ output (2MW for Horns Rev and 2.3MW for Nysted)

- 4 competing models: ARMA as a benchmark, and SETAR, STAR and MSAR models with 3 regimes

- The exercise consists in 1-step ahead forecasting

- The model parameters are estimated on training sets of 1 month, and the remaining data are used as evaluation sets
The evaluation set is divided in 19 different periods of different lengths and characteristics.

The difference between ARMA, SETAR and STAR model is very small whatever the period.

MSAR models generally outperform the others.
Spatio-temporal modelling - Initial study

- 7 months of data for 22 wind farms in western Denmark:
  - wind power forecasts
  - wind power measurements
  - corresponding met. forecasts
- gathered in 5 groups

- Objectives:
  - analyse characteristics of errors
  - propose relevant local models

- Focus on 1-hour ahead forecast errors for group 5...
Effect of wind direction

- Study of the cross-correlation pattern between 1-hour ahead forecast error between group 1 and 5, for 4 wind direction regimes

Table 1: Directional correlation for groups 5 and 1, for lags ranging from 0 to 5 hours.

<table>
<thead>
<tr>
<th>lag</th>
<th>(0-90]</th>
<th>(90-180]</th>
<th>(180-270]</th>
<th>(270-360]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0457</td>
<td>0.1472</td>
<td>0.2240</td>
<td>0.1580</td>
</tr>
<tr>
<td>1</td>
<td>0.0499</td>
<td>0.2856</td>
<td>0.3597</td>
<td>0.2361</td>
</tr>
<tr>
<td>2</td>
<td>0.0672</td>
<td><strong>0.3103</strong></td>
<td><strong>0.4213</strong></td>
<td>0.2219</td>
</tr>
<tr>
<td>3</td>
<td>0.0358</td>
<td>0.1810</td>
<td>0.3218</td>
<td>0.1542</td>
</tr>
<tr>
<td>4</td>
<td>-0.0166</td>
<td>0.0985</td>
<td>0.2193</td>
<td>0.0519</td>
</tr>
<tr>
<td>5</td>
<td>0.0115</td>
<td>0.1130</td>
<td>0.1347</td>
<td>-0.0099</td>
</tr>
</tbody>
</table>
Effect of wind speed

- Study of the correlation pattern between 1-hour ahead forecast errors at group 1 and 5, in the (180-270] regime, as a function of wind speed

![Graph showing cross-correlation as a function of wind speed and lag]
Model proposals

- The observed cross-correlation pattern for various lags allows to think ARX models can be employed:

\[ y_t = \beta_0 + \sum_{l=1}^{7} \beta_l y_{t-l} + \sum_{i=1}^{3} \beta_{1,i} x_{1,t-i} + \sum_{j=1}^{2} \beta_{4,j} x_{4,t-j} + \epsilon_t \]

- The effect of wind direction may be captured with a regime-switching approach:

\[ y_t = \beta_0^{(st)} + \sum_{l \in L_y^{(st)}} \beta_l^{(st)} y_{t-l} + \sum_{i=1}^{4} \sum_{j \in L_x^{(st)}} \beta_{i,j}^{(st)} x_{i,t-j} + \epsilon_t \]

- The more complex effect of wind speed may be captured with conditional parametric models:

\[ y_t = \beta_0^{(st)}(\hat{u}_t) + \sum_{l \in L_y^{(st)}} \beta_l^{(st)}(\hat{u}_t) y_{t-l} + \sum_{i=1}^{4} \sum_{j \in L_x^{(st)}} \beta_{i,j}^{(st)}(\hat{u}_t) x_{i,t-j} + \epsilon_t \]
Spatio-temporal forecasting

Predictive improvement (measured in RMSE) of forecasts errors by adding the spatio-temperal module in WPPT.

- 23 months (2006-2007)
- 15 onshore groups
- Focus here on 1-hour forecast only
- Larger improvements for eastern part of the region
- Needed for reliable ramp forecasting.
- New EU project NORSEWinD will extend the region
Combined forecasting

- A number of power forecasts are weighted together to form a new improved power forecast.
- These could come from parallel configurations of WPPT using NWP inputs from different MET providers or they could come from other power prediction providers.
- In addition to the improved performance also the robustness of the system is increased.

The example shows results achieved for the Tunø Knob wind farms using combinations of up to 3 power forecasts.

If too many highly correlated forecasts are combined the performance may decrease compared to using fewer and less correlated forecasts. Typically an improvement on 10-15 pct is seen by including more than one MET provider.
Uncertainty estimation

In many applications it is crucial that a prediction tool delivers reliable estimates (probabilistic forecasts) of the expected uncertainty of the wind power prediction.

WPPT provides three methods for estimating the uncertainty of the forecasted wind power production:

- Adaptive variance estimation.
- Ensemble based - but corrected - quantiles.
- Quantile regression.

The plots show raw (top) and corrected (bottom) uncertainty intervals based on ECMWF ensembles for Tunø Knob (offshore park), 29/6, 8/10, 10/10 (2003). Shown are the 25%, 50%, 75%, quantiles.
Ensemble based uncertainties

Individual ECMWF wind speed ensembles and quantiles (10, 25, 50, 75, 90 percent) (Tuno Knob):
A significant adjustment of the raw ensemble based quantiles is needed. Statistical models are developed for adjusting the ensemble based quantiles. (H.Aa. Nielsen, et.al.: Wind Power Ensemble Forecasting, GW04, Chicago)
Network of synoptic stations in Europe

-20 -10 0 10 20 30 40
40 50 60 70

longitude [degrees East]

latitude [degrees North]

○ synoptic stations (1209)
General reliability evaluation

- Extremely good reliability if vs. analysis (>48-hour ahead)
General reliability evaluation

- Very poor if vs. observations (all horizons, especially <72-hour ahead)
MAE and RMSE against observations

- **MAE:**
  - generally below 3.4 m/s
  - higher values at coastal and montainous areas

- **RMSE:**
  - generally below 5.1 m/s

- Maps of MAE and RMSE against observations for 72-hour ahead forecasts
Quantile regression

A (additive) model for each quantile:

\[ Q(\tau) = \alpha(\tau) + f_1(x_1; \tau) + f_2(x_2; \tau) + \ldots + f_p(x_p; \tau) \]

- \( Q(\tau) \): Quantile of forecast error from an existing system.
- \( x_j \): Variables which influence the quantiles, e.g. the wind direction.
- \( \alpha(\tau) \): Intercept to be estimated from data.
- \( f_j(\cdot; \tau) \): Functions to be estimated from data.

Notes on quantile regression:

- Parameter estimates found by minimizing a dedicated function of the prediction errors.
- The variation of the uncertainty is (partly) explained by the independent variables.
Quantile regression - An example

Effect of variables (- the functions are approximated by Spline basis functions):

- Forecasted power has large influence.
- The effect of horizon is of less importance.
- Some increased uncertainty for Westerly winds.
Example: Probabilistic forecasts

- Notice how the confidence intervals vary ...
- But the correlation in forecasts errors is not described so far.
Correlation structure of forecast errors

- It is important to model the interdependence structure of the prediction errors.
- An example of interdependence covariance matrix:
Statistical scenarios

- The resulting reliable statistical scenarios can be used as input for economical optimizations etc.
Realistic development of the future – reflect the correctly calibrated quantiles and the observed auto correlation (on an appropriate scale).
Correct (top) and naive (bottom) scenarios
Types of forecasts required

- Basic operation: Point forecasts
- Operation which takes into account asymmetrical penalties on deviations from the bid: Quantile forecasts
- Stochastic optimisation taking into account start/stop costs, heat storage, and/or ’implicit’ storage by allowing the hydro power production to be changed with wind power production: Scenarios respecting correctly calibrated quantiles and auto correlation.
Wind power – asymmetrical penalties

The revenue from trading a specific hour on NordPool can be expressed as

\[ P_S \times \text{Bid} + \begin{cases} 
P_D \times (\text{Actual} - \text{Bid}) & \text{if } \text{Actual} > \text{Bid} \\
P_U \times (\text{Actual} - \text{Bid}) & \text{if } \text{Actual} < \text{Bid}
\end{cases} \]

\( P_S \) is the spot price and \( P_D/P_U \) is the down/up reg. price.

The bid maximising the expected revenue is the following quantile

\[ \frac{E[P_S] - E[P_D]}{E[P_U] - E[P_D]} \]

in the conditional distribution of the future wind power production.
Wind power – asymmetrical penalties

- It is difficult to know the regulation prices at the day ahead level – research into forecasting is ongoing.
- The expression for the quantile is concerned with expected values of the prices – just getting these somewhat right will increase the revenue.
- A simple tracking of $C_D$ and $C_U$ is a starting point.
- The bids maximizing the revenue during the period September 2009 to March 2010:
Value of wind power forecasts

- Case study: A 15 MW wind farm in the Dutch electricity market, prices and measurements from the entire year 2002.
- From a phd thesis by Pierre Pinson (2006)
- The costs are due to the imbalance penalties on the regulation market.
- Value of an advanced method for point forecasting: **The regulation costs are diminished by nearly 38 pct.** compared to the costs of using the persistence forecasts.
- Added value of reliable uncertainties: **A further decrease of regulation costs – up to 39 pct.**
Conclusions

Some conclusions from more than 15 years of wind power forecasting:

- The forecasting models must be adaptive (in order to account for changes of dust on blades, changes in roughness, etc. into account).
- Reliable estimates of the forecast accuracy are very important (check reliability by eg. reliability diagrams).
- Use more than a single MET provider for delivering the input to the wind power prediction tool – this improves the accuracy by 10-15 pct.
- The tool must be easy to calibrate to new wind farms, etc.
- It is advantageous if the same tool can be used for forecasting for a single wind farm, a collection of wind farms, a state/region, and the entire country.
- Use statistical methods for phase correcting the phase errors – this improves the accuracy by up to 20 pct.
- Estimates of the correlation in forecast errors important.
- Forecasts of cross correlations between (say) load forecasting and wind power forecasting might be of importance.

Almost the same conclusions hold for solar power forecasting.
Some references

Some references (Cont.)


For download of further information and details:
[www.enfor.dk](http://www.enfor.dk)