Work Package 3:
Candidate Prediction Models and Methods

Henrik Aalborg Nielsen (han@imm.dtu.dk)
Torben Skov Nielsen (tsn@imm.dtu.dk)
Henrik Madsen (hm@imm.dtu.dk)
Informatics and Mathematical Modelling
Technical University of Denmark
DK-2800 Kongens Lyngby

Gregor Giebel (gregor.giebel@risoe.dk)
Wind Energy Department
Risø National Laboratory
DK-4000 Roskilde

13th June 2005

Project title:	Intelligent wind power prediction systems
PSO Project number:	FU 4101
Ens. journal number:	79029-0001

Contents

1 Introduction **5**

2 Numerical Weather Predictions **5**

3 List of Models / Methods **5**

3.1 Zephyr/Prediktor 5

3.2 Zephyr/WPPT 6

3.3 New reference predictor 7

3.4 Prediction models not using meteorological input 7

3.5 Other systems 8

4 Conclusion **8**

References **9**

A Zephyr/Prediktor **11**

A.1 The height correction 11

A.2 The roughness and orography corrections – The WAsP matrix 14

A.3 The power curve 16

A.4 Wake effects 17

A.5 Conclusion on the full Zephyr/Prediktor 19

A.6 The most simple power curve model 20

B Zephyr/WPPT **22**

B.1 Prediction models 23

B.2 The upscaling model ($PP_{i,1}^{ar}$) 24

B.3 The area model ($PP_{i,2}^{ar}$) 24

B.4 The total model (PP^{to}) 25

C New models based on PPR **26**

Summary

This document lists candidate prediction models for Work Package 3 (WP3) of the PSO-project called “Intelligent wind power prediction systems” (FU4101). The main focus is on the models transforming numerical weather predictions into predictions of power production. The document also outlines the possibilities w.r.t. different numerical weather predictions actually available to the project.

1 Introduction

This document aims at listing candidate prediction models for Work Package 3 (WP3) of the PSO-project called "Intelligent wind power prediction systems" (FU4101). These models are going to be used in WP5 (combined forecasting) and they may potentially influence the methods considered in WP4 (adaptive estimation). It is also natural to expect some of the model structures to be considered in WP7 (estimation based on economic criteria).

Naturally, many candidate models and methods can be considered, including many Numerical Weather Prediction (NWP) models. To start with the NWP models; in this project we have for the Danish sites access to historic DMI-HIRLAM forecasts (Sass et al., 2002) and to forecasts from the Lokalmodell of Deutscher Wetterdienst (DWD) (Schrodin et al., 2001). For DWD the period covers full a year (1/12-2002 to 30/11-2003), for DMI the period is much longer. For the Spanish cite Alaiz we have access to one NWP only (also a HIRLAM model).

To some extend this limits the scope of the candidate models to the models relating the NWP model to the power production. According to Giebel et al. (2003) these models can be subdivided in to physical and statistical models, see also (Landberg et al., 2003). However, as also described by Giebel et al. (2003) models using both physical and statistical models can presumably perform well.

Note that candidate models / methods can also be e.g. different adaptive methods or just forgetting factors applied to the same model structure. For an example of this see e.g. (Nielsen and Madsen, 2000), although this reference considers heat consumption.

2 Numerical Weather Predictions

In the project we have access to Numerical Weather Predictions (NWPs) from DMI, specifically the HIRLAM system Sass et al. (2002) and from DWD, specifically the Lokalmodell system Schrodin et al. (2001). For Alaiz (Spain) we have access to the full ANEMOS test-case.

3 List of Models / Methods

3.1 Zephyr/Prediktor

Prediktor is system based on physical modelling, including (i) correction for height (NWP nominal height compared to hub height), (ii) correction for local effects (roughness and orography), and (iii) wind farm power curve, including wake effects. The physical models of Prediktor described in AppendixA are usually used together with Model Output Statistics (MOS) to further fine tune the predictions. The MOS module is a very simple linear correction of the local wind speed, to yield either minimum Mean Absolute Error or minimum Root Mean Square Error of the power forecasts.

For combined forecasting it makes sense to use Prediktor in the following setups:

- Prediktor in standard setup (including MOS)
- Prediktor without MOS, but with an adaptive calibration of the power output.
- Prediktor applied to different NWPs for the sites where this is available (DMI and DWD forecasts for the Danish sites Klim and Middelgrunden).

3.2 Zephyr/WPPT

In this project we consider WPPT (Wind Power Prediction Tool), version 4. This is a complex system which can be configured in many different ways to handle different aspects of operation in practice; including individual wind farms, up-scaling, and regional forecasts. In this project we focus on individual wind farms. A detailed description can be found in Appendix B.

Considering only individual wind farms WPPT employs a number of models where (i) the direction-dependent power curve model is the central part and (ii) the output from this model is corrected by a model taking autocorrelation and diurnal variation into account. The models are purely data-driven and are continuously updated using the actual observations. Both models are so-called (semi) non-parametric models and therefore the estimates are determined by a set of constants called *bandwidths* which must be selected. Furthermore the time-adaptiveness is determined constants called *forgetting factors*.[1] Knowledge of the site under consideration can in principle be used to guide the selection of these tuning parameters.

Combined forecasting can be applied to WPPT outputs resulting from different settings of bandwidths and forgetting factors. This can result in many inputs for the combined forecast. Experience shows that this can be a quite problem in real systems consisting of many wind farms as e.g. the system in the Western part of Denmark. In the project we aim at solving this problem by developing self-tuning methods for bandwidths and forgetting factors. Assuming this to be possible the setups of WPPT to be used in this project only concerns the NWP input:

- WPPT applied to different NWPs for the sites where this is available (DMI and DWD forecasts for the Danish sites Klim and Middelgrunden).

[1] Other parts of this project focus at making self-tuning systems which continuously make small changes to the bandwidths and/or forgetting factors if this improves the performance.

3.3 New reference predictor

The new reference predictor suggested by Nielsen et al. (1999) is a predictor only working on past production values. Assuming the correlation to be known the predictor finds an optimal combination between the persistent predictor and a global mean. The correlation must be estimated based on a training set or alternatively the correlation can be estimated adaptively.

3.4 Prediction models not using meteorological input

Prediction models not using meteorological input must be based on diurnal variation and autocorrelation in the power production. Actually the new reference predictor described in Section 3.3 belongs to this class.

An additive, but general form of this model is

$$P_t = d(t) + z_t \tag{1}$$

where $d(t)$ is a parameterization of the diurnal variation and z_t is an ARMA-process. Even when $d(t)$ is linear in the parameters adaptive estimation in the model is difficult to handle unless a two-stage approach is used and $\{z_t\}$ is approximated by an AR-process. The two stages to the estimation procedure is

1. Estimate the diurnal variation using OE-estimates Ljung (1987), i.e. LS-estimates ignoring the correlation of $\{z_t\}$.
2. Based on the residuals from the previous step; estimate the parameters of the AR-process, possibly separately for each horizon.

There is a number of possibilities with respect to modelling of the diurnal variation. These possibilities can be grouped in two:

1. Parametrization of $d(t)$ given in terms of harmonic expansions (as in WPPT), periodic spline bases, and possibly other parameterizations.
2. Measure of the time of day. The obvious approach is to use the time of day (excluding daylight savings). To model some of the seasonal variation more directly a measure based on the sun height may be developed, such a measure must include information about morning / afternoon also.

The model used in the original WPPT system ELSAM (1995) does not use meteorological input and hence it belongs to the class of models considered in this section. Besides on-line measurements of the power production it use on-line measurements of the wind speed (ELSAM, 1995, Sec. 6.5.4) and therefore we can not directly use the model in this project.[2]

3.5 Other systems

Both Prediktor and WPPT may have some limitations w.r.t. complex terrain. In Appendix C some ideas for new model developments are listed, but this is somewhat beyond the focus of this project. Other possibilities would be to use model outputs from ANEMOS and develop routines for combined forecasting based on this. This is of special interest for the Alaiz and Klim cases which is included in both ANEMOS and in this project.

4 Conclusion

A number of candidates for inclusion into a forecasting systems based on combined forecasting is described in this report. The most obvious are the two Zephyr systems Prediktor and WPPT since these are well known to the participants in the project. However, historically, none of these systems has had very much focus on horizons below 12 hours. It therefore also makes sense to include methods dedicated to short term prediction as described in Sections 3.3 and 3.4.

An other approach to the investigation into combined forecasting may be to use the results from ANEMOS and combine forecasts from these quite different systems into a forecast. Hopefully, an adaptive procedure can be developed, whereby the combined forecast will be near-optimal for any of the test-cases considered.

[2] It has been decided not to base the prediction methods on measurements of wind speed since experience indicate that these measurements are unreliable.

References

Bergmann J.C. A physical interpretation of von karman's constant based on asymptotic considerations – a new value. *Journal of Atmospheric Science*, **55**:3403–3405, 1998.

ELSAM. *Final Report, Wind Power Prediction Tool in Central Dispatch Centres, JOU2-CT92-0083*. ELSAM, 1995.

Friedman J.H. SMART¡ user's guide. Technical Report 1, Laboratory for Computational Statistics, Department of Statistics, Stanford University, California, 1984.

Friedman J.H. and Stuetzle W. Projection pursuit regression. *Journal of the American Statistical Association*, **76**:817–823, 1981.

Giebel G., Brownsword R., and Kariniotakis G. The state-of-the-art in short-term prediction of wind power; a literature overview. Technical report, Project ANEMOS public document, 2003. `http://anemos.cma.fr/download/ANEMOS_D1.1_StateOfTheArt_v1.1.pdf`.

Hastie T., Tibshirani R., and Friedman J.H. *The elements of statistical learning: data mining, inference, and prediction*. Springer-Verlag Inc, 2001.

Hoerl A.E. and Kennard R.W. Ridge regression: biased estimation of nonorthogonal problems. *Technometrics*, **8**:27–51, 1970.

Landberg L., Giebel G., Nielsen H.A., Nielsen T., and Madsen H. Short-term prediction – an overview. *Wind Energy*, **6**:273–280, 2003.

Ljung L. *System Identification, Theory for the User*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

Ljung L. and Söderström T. *Theory and Practice of Recursive Identification*. MIT Press, Cambridge, Massachusetts, 1983.

Marti I., Nielsen T.S., Madsen H., Navarro J., and Barquero C.G. Prediction models in complex terrain. In *Proceedings of the European Wind Energy Conference*. Copenhagen, Denmark, 2001.

Nielsen H.A. and Madsen H. *Predicting the Heat Consumption in District Heating Systems using Meteorological Forecasts*. Department of Mathematical Modelling, Technical University of Denmark, DK-2800 Lyngby, Denmark, 2000. URL `http://www.imm.dtu.dk/~han/pub/efp98.pdf`. ENS. J. Nr. 1323/98-0025.

Nielsen T. (editor). *Using Meteorological Forecasts in On-line Predictions of Wind Power*. Eltra, DK-7000 Fredericia, Denmark, 1999. ISBN: 87-90707-18-4.

Nielsen T.S., Joensen A., Madsen H., Landberg L., and Giebel G. A new reference for predicting wind power. *Wind Energy*, **1**:29–34, 1999.

Nielsen T.S., Madsen H., Nielsen H.A., Giebel G., and Landberg L. Zephyr – the prediction models. In *Proceedings of the European Wind Energy Conference*. Copenhagen, Denmark, 2001.

S-PLUS. *S-PLUS 6; Guide to Statistics, Vol. 1*. Data Analysis Division, MathSoft, Seattle, WA, 2000.

Sass B.H., Nielsen N.W., Jørgensen J.U., Amstrup B., Kmit M., and Mogensen K.S. The operational DMI-HIRLAM system 2002-version. Technical Report 02-05, Danish Meteorological Institute, 2002. `http://www.dmi.dk`.

Schrodin R., Link A., and Anlauf A. Quarterly Report of the Operational NWP-

Models of the Deutscher Wetterdienst, No. 28. Deutscher Wetterdienst, Meteorologische Analyse und Modellierung, Postfach 10 04 65, 63004 Offenbach, 2001. `http://www.dwd.de/en/FundE/Veroeffentlichung/Quarterly/qr_no_28.pdf`.

Venables W.N. and Ripley B.D. *Modern applied statistics with S-PLUS*. Springer-Verlag Inc, 1999.

A Zephyr/Prediktor

Prediktor consists of three separate physical parts, which are applied sequentially:

- The height correction.
- The local corrections due to roughness and orography (terrain height profile).
- The wind farm power curve including wake effects.

This text will give a short overview of all three mechanisms, and the physical considerations behind each one of them. Additional to the physical considerations mentioned here, Prediktor usually also employs the use of a (relatively simple) Model Output Statistics (MOS) step, if such data is available. The physics behind Zephyr/Prediktor is based in part on the knowledge gained in the development of the WAsP model (Wind Atlas Analysis and Application Program) developed in the late 1980s at Risø National Laboratory. WAsP was developed to be able to “translate” wind speed measurements made at one site to wind climates at a (nearby) site. It does this by calculating a wind rose and distribution, then taking out the local effects that have influenced that particular measurement, thereby creating a regional wind climate, and then reintroducing the local effects at the site of interest (e.g. a wind turbine). The one effect WAsP can calculate, but Prediktor usually ignores, is the effect of obstacles, as in any proper wind farm, there are none.

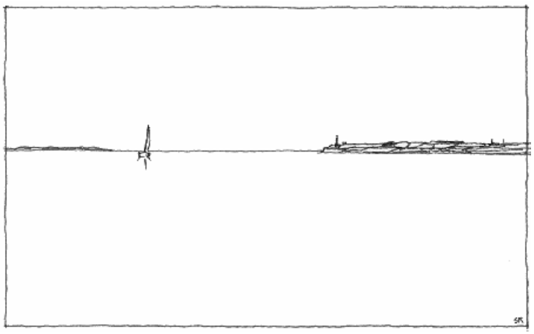
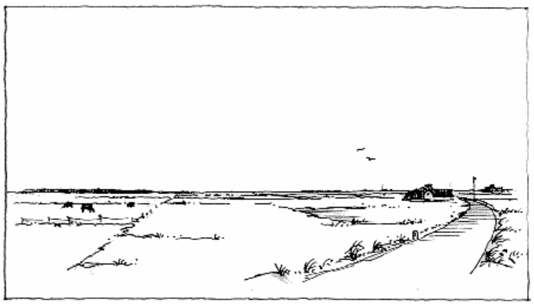
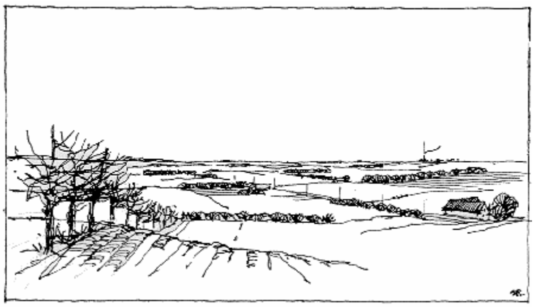
A.1 The height correction

The wind speed above the surface corresponds to the following formula, the so-called logarithmic height profile:

$$u(z) = \frac{u_*}{\kappa} \ln\left(\frac{z}{z_0}\right) \tag{2}$$

u_* is the friction velocity, κ is the von Kármán constant (according to some theoretical estimates 1/e, i.e. ca. 0.37) (Bergmann, 1998), z is the height above ground level (a.g.l.), and z_0 is the so-called roughness length. Conceptually, the roughness length is the length where the line drawn by the logarithmic profile in a log-linear plot crosses zero velocity (it does not do so in practice, since there are additional surface effects very close to the ground). The influence of the roughness on the vertical profile can be seen in Illustration 1, where all wind speeds converge for large height outside of the planetary boundary layer to a geostrophic wind of 9 m/s.

Here is an example for the different roughness classes:

 *Example of terrain corresponding to roughness class 0:* water areas. This class comprises the sea, fjords, and lakes. The roughness length is $z_0 = 0.0002m$.	 *Example of terrain corresponding to roughness class 1:* open areas with few windbreaks. The terrain appears to be very open and is flat or gently undulating. Single farms and stands of trees and bushes can be found. The roughness length is $z_0 = 0.03m$.
 *Example of terrain corresponding to roughness class 2:* farmland with windbreaks, the mean separation of which exceeds 1000 m, and some scattered built-up areas. The terrain is characterised by large open areas between the many windbreaks, giving the landscape an open appearance. The terrain may be flat or undulating. There are many trees and buildings. The roughness length is $z_0 = 0.10m$.	 *Example of terrain corresponding to roughness class 3:* urban districts, forests, and farmland with many windbreaks. The farmland is characterized by the many closely spaced windbreaks, the average separation being a few hundred meters. Forest and urban areas also belong to this class. The roughness length is $z_0 = 0.40m$.

Therefore, if one has NWP (Numerical Weather Prediction) input from typically 10 m a.g.l., but wants to scale that up to hub height, then the logarithmic profile comes up with the following translation for the wind speeds: $u(z_2) = u(z_1) \times ln(z_2/z_0)/ln(z_1/z_0)$. With z_1 and z_2 being the NWP level height and the hub height, respectively, this relationship comes down to one factor depending on z_0. However, the roughness length can have a fairly complex behaviour, depending on the direction. One just has to think of coastal sites. WAsP gets a roughness map

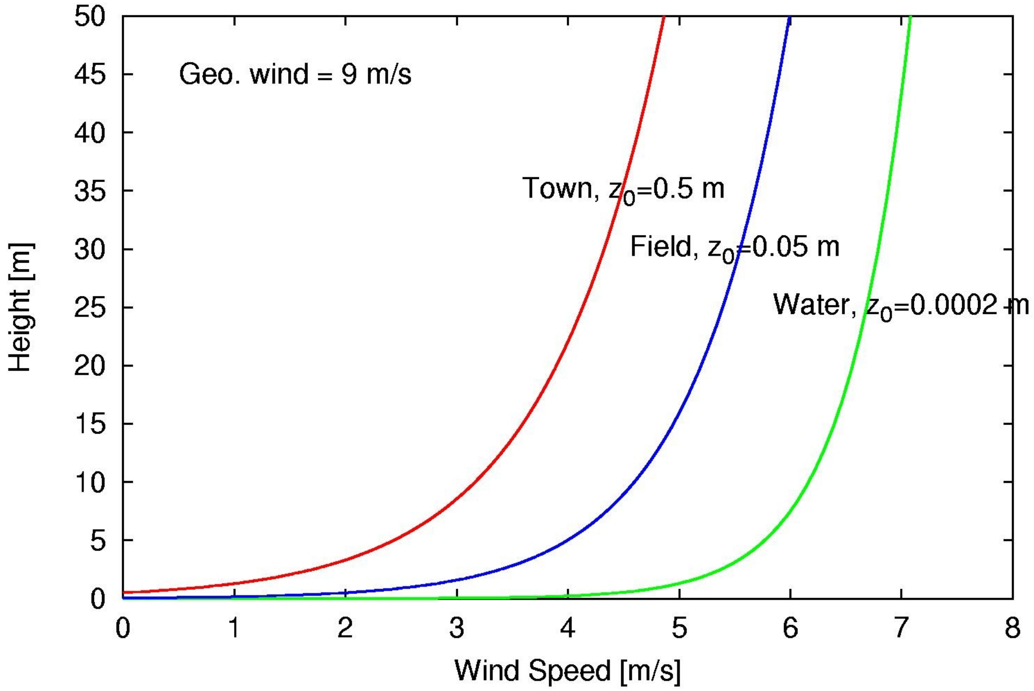


Illustration 1: The wind speed profile for different roughness, keeping the geostrophic wind constant.

or roughness rose as an input, and calculates from that an ”effective roughness length”, also taking into account the internal boundary layers developing at roughness changes. The rule of thumb for the upward travel of information on a roughness change is 1:100, which means that for a 100 m turbine, roughness information from as far as 10 km away can be important. For this calculation, WAsP partitions the directions into (typically) 12 sectors.

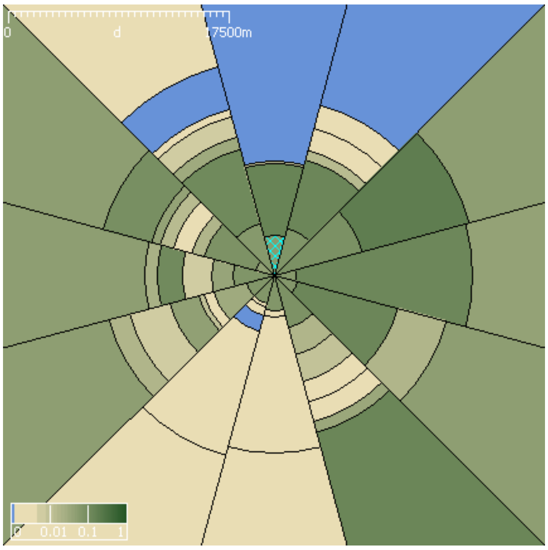


Illustration 2: The roughness rose for the Klim wind farm.

That all means that, in effect, the wind speed at a certain height $u(z_2) = u(z_1) \times \text{const(sect)}$. In other words, the variation of the wind speed for scaling the wind from one level to hub height is linear, and only depends on the direction.

A.2 The roughness and orography corrections – The WAsP matrix

The WAsP matrix introducing the local effects also contains a roughness correction. This one reflects the changes due to the roughness deviating from a uniform 3-cm roughness value.

The third effect of the WAsP matrix is due to the orography. Wind shows a speedup if it flows over a hill, and also is faster offshore (ie in very low roughness). Illustration 3 shows this principle for water to the left, and land to the right. On the land there is a mountain, and in the sea there is an island. The graphs show the wind resource.

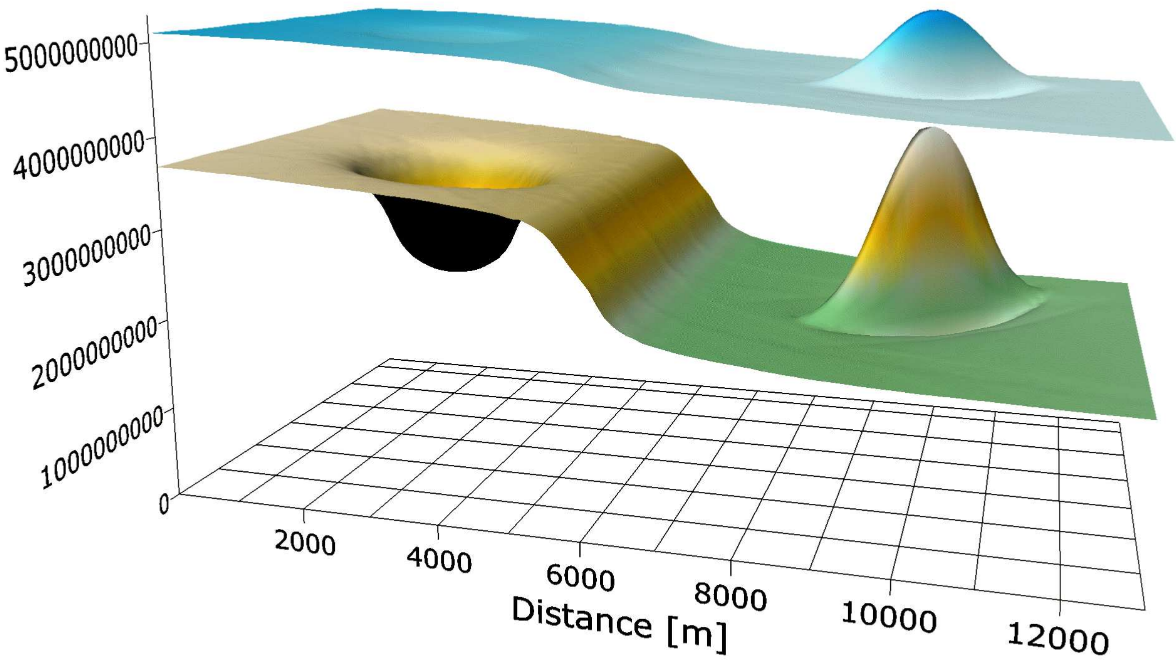


Illustration 3: Power production @ 10 and 100 m a.g.l. Uniform wind rose, hill is 100 m high. The shore is at the 6000m-line. There is an island where the locally low resource is, and a mountain at the high resource.

DIR	INPUT		OBSTACLES		ROUGHNESS		OROGRAPHY		Z0
0	0.000	0.000	0.000	0.000	-8.000	0.000	-0.135	0.053	0.007952
30	0.000	0.000	0.000	0.000	-7.552	0.000	-0.038	0.050	0.012074
60	0.000	0.000	0.000	0.000	0.055	0.000	-0.040	-0.014	0.092192
90	0.000	0.000	0.000	0.000	-1.814	0.000	-0.099	-0.051	0.072815
120	0.000	0.000	0.000	0.000	-1.478	0.000	-0.178	-0.040	0.073059
150	0.000	0.000	0.000	0.000	-3.428	0.000	-0.193	0.012	0.028615
180	0.000	0.000	0.000	0.000	-7.282	0.000	-0.125	0.048	0.002046
210	0.000	0.000	0.000	0.000	-6.222	0.000	-0.036	0.039	0.001180
240	0.000	0.000	0.000	0.000	-5.219	0.000	-0.007	-0.008	0.011303
270	0.000	0.000	0.000	0.000	-1.579	0.000	-0.099	-0.048	0.039645
300	0.000	0.000	0.000	0.000	-5.063	0.000	-0.174	-0.038	0.017922
330	0.000	0.000	0.000	0.000	-6.814	0.000	-0.205	0.005	0.015095

This matrix contains everything Prediktor uses to calculate the wind speed at hub height from the incoming NWP wind speed: roughness corrections (only speed), orography speed-ups and turnings, the roughness length for the height profile, and additionally the user corrections (not used) and the corrections due to obstacles (there are no obstacles near a wind farm).

A.3 The power curve

WAsP uses a standard power curve from a manufacturer, an example is shown in Illustration 4. This power curve can be different according to the noise level restriction or air density. However, compared with all the other uncertainties involved in short-term predictions, these are smaller error sources. A very practical feature of WAsP is that it comes with an extensive library of (Danish) power curves, which can be used for most projects, at least as start-up. The differences between different companies turbines of same class and same rating are relatively small. The most important feature to look for is the rotor diameter and the rating. Larger rotor diameters mean that even for lower wind, there already is some production, and that the amount of full load hours is probably higher.

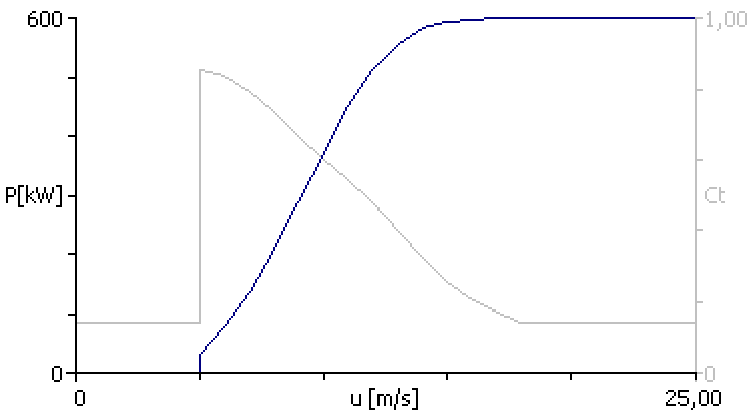
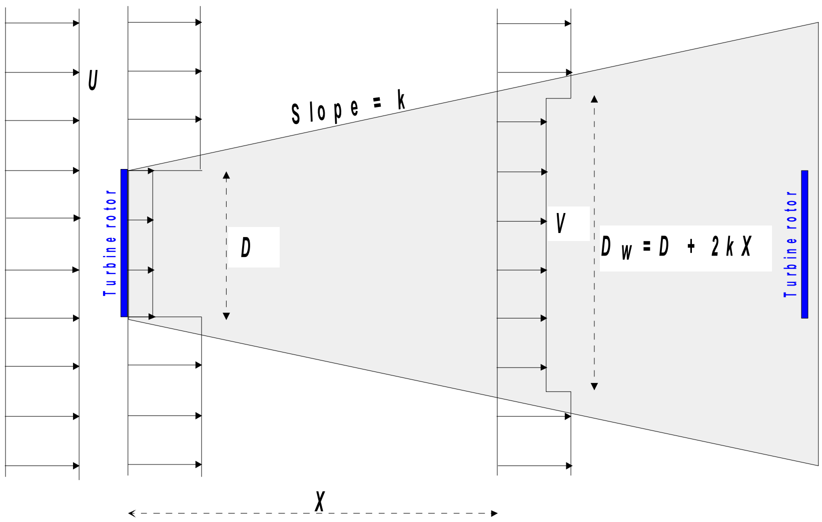


Illustration 4: The power and Cp curve of the Vestas V44 used in Klim.

A.4 Wake effects

The principle of wake effects in wind farms is shown in Illustration 5. Behind the turbine, a wake develops, where the wind speed is reduced. This is logical, since the whole point of a wind turbine is to draw kinetic energy (read: speed) out of the wind. This wake is getting successively larger, and thereby gets diluted, up to the point where it vanishes. One of the important factors here is the ambient turbulence, since this determines how coherent the wake stays behind the turbine. The lower the ambient turbulence (eg offshore), the longer the wake is measurable. In a typical wind farm, the wake loss is 5% or less, but in some cases, especially in larger wind farms and in wind farms with small horizontal separation (smaller than 5 rotor diameters between turbines), the losses can be higher.



$$V = U\left[1 - (1 - \sqrt{1 - C_t})\left(\frac{D}{D + 2kX}\right)^2\right]$$

Illustration 5: Schematic idea of wake effect model used in WAsP, and the relevant formula. k = Wake decay constant.

From the standard power curves and the wake effects, WAsP calculates a park power curve (see this extract of the Klim park power curve):

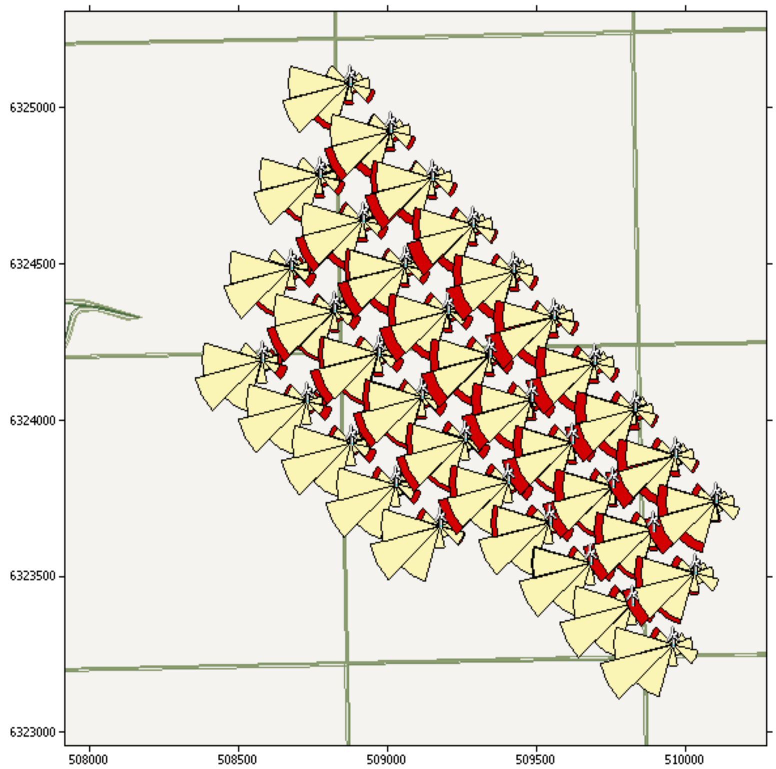


Illustration 6: The Klim wind farms annual energy production and the wake losses (red) per sector for the individual turbines.

U/Dir	0.0	6.0	12.0	18.0	24.0	30.0	36.0	42.0	48.0	54.0	60.0	66.0	72.0	78.0	84.0	90.0
4.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5.0	0.0089	0.0075	0.0075	0.0000	0.0000	0.0000	0.0000	0.0000	0.0045	0.0045	0.0045	0.0045	0.0060	0.0000	0.0000	0.0000
6.0	0.1117	0.1129	0.0711	0.0650	0.0821	0.1184	0.1099	0.1097	0.1131	0.0768	0.0609	0.0774	0.1040	0.1126	0.1050	0.0899
7.0	0.2022	0.2038	0.1528	0.1420	0.1634	0.2108	0.1996	0.1993	0.2034	0.1569	0.1406	0.1560	0.1917	0.2031	0.1928	0.1721
8.0	0.3144	0.3165	0.2512	0.2367	0.2639	0.3247	0.3107	0.3103	0.3148	0.2551	0.2339	0.2539	0.2998	0.3149	0.3016	0.2750
9.0	0.4452	0.4476	0.3714	0.3531	0.3864	0.4562	0.4403	0.4398	0.4440	0.3742	0.3482	0.3727	0.4267	0.4448	0.4296	0.3990
10.0	0.5847	0.5870	0.5081	0.4871	0.5238	0.5947	0.5788	0.5784	0.5813	0.5086	0.4799	0.5069	0.5636	0.5830	0.5676	0.5365
11.0	0.7189	0.7211	0.6482	0.6270	0.6628	0.7271	0.7127	0.7123	0.7139	0.6463	0.6180	0.6443	0.6975	0.7163	0.7024	0.6742
12.0	0.8324	0.8340	0.7762	0.7575	0.7879	0.8378	0.8267	0.8264	0.8266	0.7722	0.7477	0.7703	0.8134	0.8292	0.8185	0.7966
13.0	0.9148	0.9158	0.8792	0.8654	0.8859	0.9171	0.9104	0.9103	0.9092	0.8738	0.8565	0.8722	0.9007	0.9117	0.9050	0.8912
14.0	0.9626	0.9631	0.9462	0.9389	0.9486	0.9631	0.9600	0.9599	0.9587	0.9421	0.9326	0.9406	0.9546	0.9604	0.9573	0.9508
15.0	0.9867	0.9869	0.9802	0.9772	0.9806	0.9866	0.9853	0.9853	0.9843	0.9778	0.9740	0.9770	0.9827	0.9854	0.9841	0.9814
16.0	0.9957	0.9957	0.9939	0.9930	0.9939	0.9955	0.9952	0.9952	0.9948	0.9930	0.9920	0.9927	0.9943	0.9952	0.9948	0.9941
17.0	0.9995	0.9995	0.9988	0.9983	0.9987	0.9994	0.9993	0.9993	0.9990	0.9983	0.9979	0.9982	0.9988	0.9992	0.9991	0.9988
18.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
19.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
21.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

A note on the park power curve: it uses as input the wind speed and direction for turbine 1. For the other turbines, the WAsP matrix (or rather, the difference between the relative speed-ups and other changes) is used accordingly. So the proper procedure to use this is to get the WAsP matrix for turbine 1, and change the NWP hub height wind accordingly.

Also, take care when parsing the.dmp file: it is not always the same length. So better parse the wind speeds with it.

A.5 Conclusion on the full Zephyr/Prediktor

Prediktor uses three mechanisms in sequence:

•Height correction: $u(z_{HH}) = u(z_{NWP}) \times$ const(sector)

•Local corrections: $u_{loc} = u(z_{HH}) \times$ mtx(sector)

•Park power curve: $P(u, d) =$ const$(u_{loc},$ sector$)$

This can be used in this project to yield a power curve analog to the one inserted above. For this, a Prediktor module is set up, and winds from 0 to 29 m/s and from 0 to 354° are put through it.

In the future (Illustration 7), more advanced flow models like KAMM (Karlsruhe Atmospheric Mesoscale Model) could be substituted for WAsP, taking into account a far wider range of effects on a larger scale, including channeling effects, orographic induced winds and stability effects. In essence, there is always a step for the physical considerations, and optionally (depending on the availability of measurements) some Model Output Statistics to correct for biases

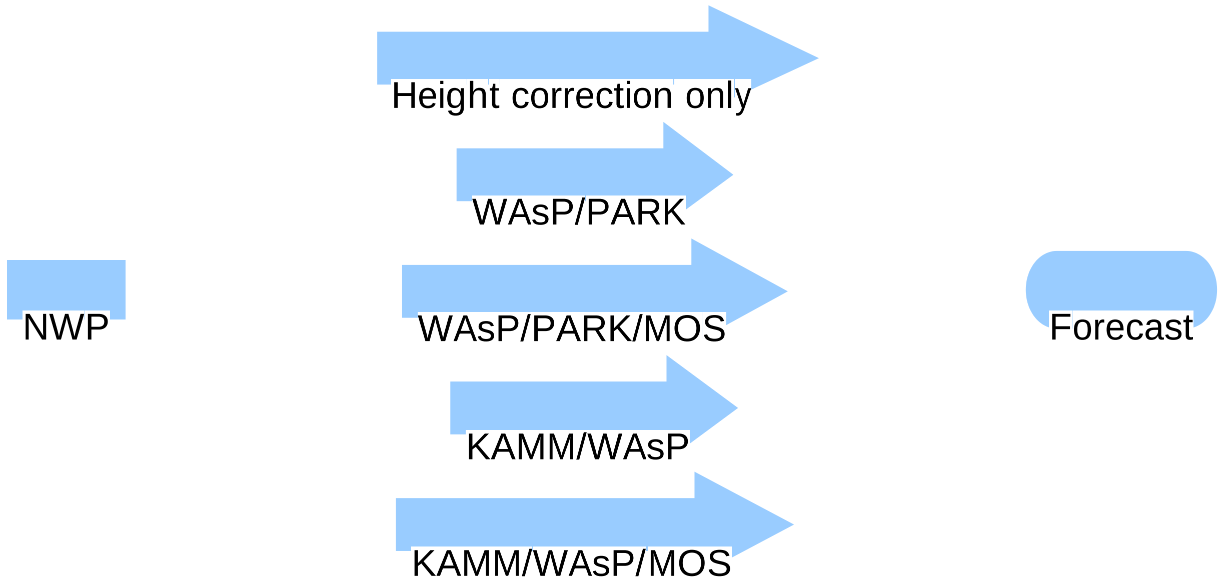


Illustration 7: Future plans for Prediktor

etc.

A.6 The most simple power curve model

Of all these steps, most require to have access to WAsP and/or digitized maps of the area. However, this is not always available, so another model is proposed, building on the same principles, but using educated guesses for most steps.

First, the height correction could be done with an estimate of the background roughness of the site. This would not be sector dependent, and would just entail some generic background roughness yielding a factor to be multiplied to the NWP result. As a typical default, the value of 3cm can be used. For a more thorough analysis without actually visiting the site, there is NASA's WorldWind tool (`http://worldwind.arc.nasa.gov/`), which allows to see a wind farm site in 3D including a satellite background photographed from Landsat.

Next, the step with the local corrections would be omitted.

Finally, instead of a proper park power curve, the manufacturers power curve would be used and just multiplied by the number of turbines. The estimated difference between this and the full park power curve is for most farms in the order of some percent, but the difference between the manufacturers power curve and the real-life wind turbine is often in the same order of magnitude.

All this model needs is access to the manufacturers power curve, which in some cases can even be replaced by a similar model (similar hub height, same rating, same rotor size) without introducing too much additional error.

B Zephyr/WPPT

The Zephyr/WPPT modelling system described in the following calculates predictions of the available wind power from wind turbines in a region. For a larger region this is done by separating the region into a number of sub-areas. Wind power predictions are then calculated for each sub-area and hereafter summarized to get a prediction for the total region.

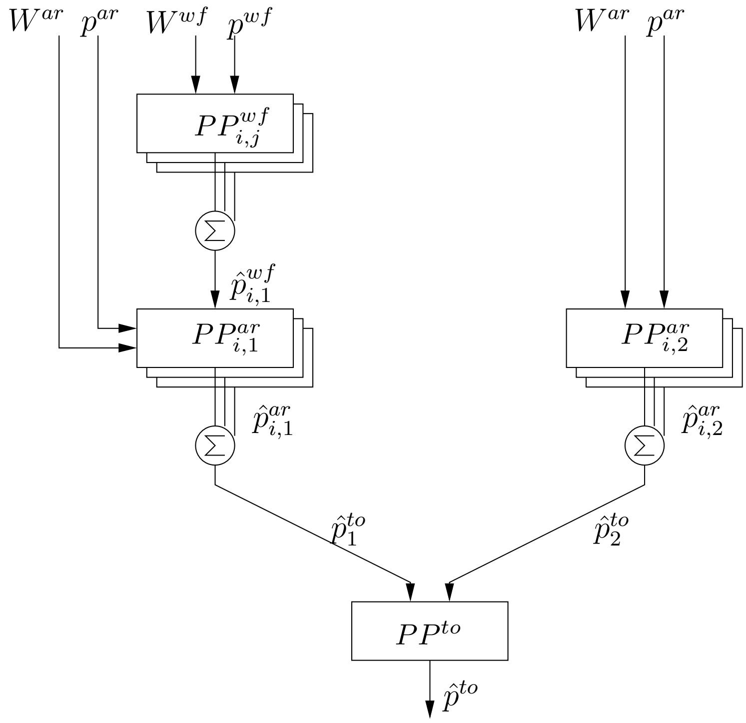


Figure 8: Overview of the model structure in Zephyr/WPPT. Two different predictions are calculated for the wind power production in a region: In the left model branch the wind farm models, $PP^{wf}_{i,j}$, are used to calculate power predictions for the reference wind farms in sub-area i. The predictions for the reference wind farms in sub-area i are summarized to $\hat{p}^{wf}_{i,1}$, which hereafter is upscaled by the upscaling model $PP^{ar}_{i,1}$ to a power prediction, $\hat{p}^{ar}_{i,1}$, for all wind turbines in the sub-area. The predictions for the sub-areas are then summarized to get the power prediction of the left model branch for the total region, $\hat{p}^{to}_{1}$. In the right model branch power predictions of the power production in sub-area i, $\hat{p}^{ar}_{i,2}$, are calculated directly by the area model $PP^{ar}_{i,2}$. The predictions for the sub-areas are then summarized to get the power prediction of the right model branch for the total region, $\hat{p}^{to}_{2}$. The final power prediction for the region, $\hat{p}^{to}$, is calculated by model $\hat{p}^{to}$ as a weighted average of the predictions from the two model branches.

The predictions are calculated using on-line production data from a number of wind farms in the area (reference wind farms), off-line production data for the remaining wind turbines in the area and numerical weather predictions of wind speed and wind direction covering the area. The predictions covers a horizon corresponding to the prediction horizon of the numerical weather predictions hours – typical from 0 to approximately 48 hours ahead in time. The time resolution of the predictions can be chosen freely but a reasonable choice for the longer prediction horizons is to use the same time resolution as the numerical weather predictions.

The predictions for the total region are calculated using a two branch approach as illustrated in figure 8.

- In the left model branch predictions of wind power are calculated for a number of reference wind farm using on-line measurements of power production as well as numerical weather predictions as input (see Appendix B.1). The predictions from the reference wind farms in a sub-area are summarized and hereafter upscaled to get the prediction of power production of all wind turbines in the sub-area (see Appendix B.2). This model branch takes advantage of the auto-correlation which is present in the power production for prediction horizons less than approximately 12 hours.
- The right model branch predicts the power production in a sub-area explicitly by using a model linking off-line measurements of total power production in the sub-area to the numerical weather predictions (see Appendix B.3). This model branch takes advantage of the smooth properties of the total production as well as the fact that the numerical weather models perform well in predicting the weather patterns but less well in predicting the local weather at a particular wind farm.

For both model branches the power prediction for the total region is calculated as a sum of the predictions for the sub-areas. The final prediction of the wind power production for the total region is then calculated as a weighted average of the predictions from the model two branches (see Appendix B.4).

B.1 Prediction models

Conditional parametric models are used to describe the relationship between observed power production in wind farms or areas and meteorological forecasts of wind speed and wind direction (the power curve). These relationships are difficult to parameterize explicitly, but can, as it is shown in Nielsen et al. (2001), readily be captured by conditional parametric models. The dynamic relationsship between observed production and predicted production from the (static) power curve models are described using a set of linear k-step predictions models, which are estimated recursively and adaptively as described in Ljung and Söderström (1983), whereas the model structure in the k-step models is identified in Nielsen (1999).

The wind farm model ($PP^{wf}_{i,j}$) The wind farm model uses wind direction dependent power curves in the transformation of forecasted wind speed and wind direction to power. The model for the jth wind farm in the ith sub-area is given as

$$\begin{aligned}\hat{p}^{pc}_{i,j}(t+k) &= f(w^{wf}_{i,j}(t+k), \theta^{wf}_{i,j}(t+k), k) \\ \hat{p}^{wf}_{i,j}(t+k) &= a_1 p^{wf}_{i,j}(t) + a_2 p^{vm}_{i,j}(t-1) + b\hat{p}^{pc}_{i,j}(t+k) + \\ & \sum_{i=1}^{3}[c^c_i \cos\frac{2i\pi h^{24}(t+k)}{24} + c^s_i \sin\frac{2i\pi h^{24}(t+k)}{24}] + m + e(t+k) \qquad (3)\end{aligned}$$

where $p^{wf}_{i,j}(t)$ is the observed power at time t, $w^{wf}_{i,j}(t+k)$ and $\theta^{wf}_{i,j}(t+k)$ are local forecasts of wind speed and wind direction, respectively, and f, a, b, and h^{24} are time-varying model parameters to be estimated. The difference between observed and forecasted diurnal variation of wind speed is contain in the h^{24} term.

The wind farm model takes advantage of the auto-correlation which is present in the power production for prediction horizons less than approximately 12 hours.

The choice of model order and input variables for each prediction horizon is described in Nielsen (1999).

B.2 The upscaling model ($PP^{ar}_{i,1}$)

The predicted power production in sub-area i is calculated by multiplying the summarized power predictions for the wind farms in the sub-area by a upscaling function, which depends on area forecasts of wind speed and wind direction. The model is given as

$$\begin{aligned}\hat{p}^{ar}_{i,1}(t+k) &= \\ & b(w^{ar}_i(t+k), \theta^{ar}_i(t+k), k)\sum_j \hat{p}^{wf}_{i,j}(t+k) \qquad (4)\end{aligned}$$

where $w^{ar}_i(t+k)$ and $\theta^{ar}_i(t+k)$ are area forecasts of wind speed and wind direction, respectively, and b is a smooth time-varying function to be estimated.

B.3 The area model ($PP^{ar}_{i,2}$)

The area model transforms area forecasts of wind speed and wind direction to power in a way similar to the wind farm power curve model by explicitly linking weather forecasts for the area to off-line observations of the power production in the area. For sub-area i the model is given as

$$\hat{p}^{ar}_{i,2}(t+k) = f(w^{ar}_i(t+k), \theta^{ar}_i(t+k), k). \qquad (5)$$

where f is a smooth time-varying function to be estimated.

This model takes advantage of the smooth properties of summarized power productions and the fact that the numerical weather models perform well in predicting the weather patterns but less well in predicting the local weather at a particular wind farm.

B.4 The total model (PP^{to})

The prediction of the total power production in the region is calculated using the total predictions from the two model branches in figure 8. The prediction is calculated as a prediction horizon dependent weighted average of the power predictions for the two model branches using Root Mean Square (RMS) as weighting criterion. The model is given as

$$\hat{p}^{to}_{t+k} = b_1(k)\hat{p}^{ar}_1(t+k) + b_2(k)\hat{p}^{ar}_2(t+k) \tag{6}$$

where $\hat{p}^{ar}_1(t+k)$ and $\hat{p}^{ar}_1(t+k)$ are the power predictions for model branch 1 and 2, respectively, and b_1 and b_2 are smooth time-varying functions to be estimated.

The predictions from the two model branches are closely correlated especially for the longer prediction horizons. Thus a regularized estimation procedure must be used to ensure stable estimates of the b_1 and b_2 functions. Here Ridge Regression Hoerl and Kennard (1970) has been used. The weighting scheme applied here might have to be changed following the results of WP5 “Combined Forecasting”.

C New models based on PPR

Marti et al. (2001) use Principal Component Regresion (PPR) (Hastie et al., 2001) to find linear combinations of output from a NWP model wich are good at predicting the wind speed measured locally. A power curve model is then used to relate the local wind speed to the power output.

Inspired by this procedure we propose to use Projection Pursuit Regression (PPR) (Friedman and Stuetzle, 1981; Hastie et al., 2001) instead of PCR followed by power curve modelling. As will be seen from the following this eliminates the need for a local wind speed measurement and the projection directions are chosen optimal with respect to a squared loss criterion on the power scale.

Let P_t denote the power production at time t and let $\mathbf{x}_t$ be a vector containing all the avilable meteorological forecasted variable at time t. In practice $\mathbf{x}$ will contain a number of forecasted variables at different model levels in a grid around the farm. The PPR model for this setup is

$$P_t = \mu + \sum_{m=1}^{M} \beta_m \phi_m(\mathbf{a}_m^T \mathbf{x}_t) + e_t \ , \tag{7}$$

where e_t is the model error at time t and $\phi_m(\cdot);\ m = 1, \ldots, M$ are functions to be estimated from data. These functions are standadized to have mean zero and unity variance over the data used for fitting the model. Finally, $\mathbf{a}_m;\ m = 1, \ldots, M$ are unit vectors to be estimated from data. These vectors are directions onto which the meteorological forecasts are projected. Above we have followed a terminology similar to the one used in (S-PLUS, 2000).

Under the restrictions outlined, given a scatter plot smoother for estimation of the functions $\phi_m(\cdot)$, and given a fititng procedure Friedman (1984) the estimates are unique if μ is fixed to e.g. the overall mean of the power production. Using the function `ppr` in the MASS library of S-PLUS or R (Venables and Ripley, 1999) ads the possibility of using smoothing splines for estimation of the functions $\phi_m(\cdot)$.

Note that since the functions $\phi_m(\cdot)$ are standardized the estimates of $\beta_m;\ m = 1, \ldots, M$ can be used as an initial guide on how many terms to included.

The procedure can also be applied to principal components of $\mathbf{x}$. However, this differs only from the above if some of the principal components are excluded.

Handling of the wind direction requires special consideration. Marti et al. (2001) models measured u- and v-components of the wind separately based on u- and v-components from HIRLAM. In this case this is not possible since we assume that only the power output from the wind farm is measured. Experimentation is needed in order to identify the most appropriate method. Below a number of observations are listed:

1. If the number of non-linear terms is high PPR can model complex interactions.

2. If both the wind speed and its u- and v-components are included then the projection part of PPR can adjust the wind speed by subtracting a plane, e.g. $\sqrt{u^2+v^2} - 0.1u + 0.05v$.

3. To include the model of Marti et al. (2001) directly in the PPR-model note that

$$P = f(\sqrt{u^2+v^2}) = g(u^2+v^2) = g(\sum_i a_i x_i + \sum_{i,j} b_{ij} x_i x_j)\ ,$$

   i.e. if the data is extended with quadratic terms the model of Marti et al. (2001) is included in the PPR approach. However, note that $x_1 x_2 = (x_1+x_2)^2 - (x_1-x_2)^2$ and for this reason even when not extending the data the model of Marti et al. (2001) is included in the PPR approach (if the number of terms is high enough).

4. A possible natural model is a power-curve multiplied by a direction dependence:

$$P = f(a^T x) \times \text{direction dep.}$$

   log-transformation will make this model additive. However, to avoid removing the weight from situations with high power output we must use weights when fitting the model.