02418 Project 1: Wind power forecast

Jan K. Møller

September 2018

Formalities

The data can be found in the file tuno.txt together with the assignment on CampusNet.

In this assignment you will be analysing a data set from Tunø Knob wind power plant. It is important to describe your results and conclusions, not only ny numbers, but also in words interpreting your results.

Motivation

Wind energy currently accounts for more than 30 % of the total electricity production in Denmark and this percentage is increasing. In 2050 100% of the total electricity production in Denmark should come from renewable sources. The big and increasing proportion of wind power imply that accurate predictions of production is essential; e.g. to balance the grid and to decide optimal trading strategies on the Nordic power exchange (Nord Pool). In the assignment you will model the average daily wind power production for a wind power plant in Denmark.

Data

The data set consists average daily values for wind energy production for Tunø-Knob wind power plant, the plant is a small off shore wind power plant located north of Samsø. The installed capacity is 5.000kw (maximum production).

Data prepossessing

This section gives a brief description of how data was treated before you use them for modelling of daily production. The original data set consists of average hourly production between midnight and 18.00, these values are in your data set translated into average production for each day. In addition to production the original data set contained corresponding meteorological predictions of a large number of variables, these are also converted to the daily averages (also based of the time period 00:00 to 18:00).

A wind power plant where all turbines stand still will actually have a negative production such numbers are set equal to zero in the original data set. Hence the data set consisted of numbers between 0 and $5,000$ kW. Average daily values that are exactly zero are thought of as special phenomena where, for one reason or another the plant is down, these observations are removed from the data set.

The total data set can be divided into three categories; 3 variables to describe the time (days since 1/1-2003, month, and day of month), 1 variable describing the observed output (power) and 2 meteorological variable. The individual variables are described in Table 1.

The data set contains meteorological variables related to wind speed and direction.

Table 1: Overview of all variables and their meaning data set `tuno.txt` .

Variable	Meaning	Unit
`r.day:`	Days since 1/1 2003	days
`month:`	Month in year	
`day:`	Day in month	
`pow.obs:`	Average daily wind power production	kW
`ws30:`	Predicted wind speed 30 meters above ground level	m/s
`wd30:`	Predicted wind direction 30 meters above ground level	rad

1 The project:

Descriptive statistics

1. Read the data `tuno.txt` into R.

2. Make a graphical presentation of data or parts of the data, and present some summary statistics.

Before you start the modelling it will make some of the questions simpler if you normalise (to numbers strictly between 0 and 1) the power production. Be careful to do this in a meaningfull way and describe precisely how it is done.

Simple models:

1. Fit different probability density models to the wind power and wind speed data. You might consider different models e.g. beta, gamma, log normal, and different transformations e.g. (for wind power)

$$y^{(\lambda)} = \frac{1}{\lambda}\log\left(\frac{y^{\lambda}}{1-y^{\lambda}}\right); \quad \lambda > 0 \tag{1}$$

$$y^{(\lambda)} = 2\log\left(\frac{y^{\lambda}}{(1-y)^{1-\lambda}}\right); \quad \lambda \in (0,1) \tag{2}$$

2. Conclude on the most appropriate model, for each of the variables.

Regression models:

1. Formulate an initial model,

$$\hat{y} = f(w_s) \tag{3}$$

   where y is observed power, and f, is a function of wind speed (e.g. $f = \beta_0 + \beta_1 w_s + \beta_2 ws^2$).

2. You might consider non-normal models and/or normal model with data transformation. Further you might consider including wind direction. You should devellop a suited model for prediction of daily power production.

3. Present the parameters of the final model, this include quantification of the unceartainty of the parameters.

4. Give an interpretation of the parameter values in particular this should include presentation of any nonlinear functions (series expansions) of the explanatory variables.

5. Present the final model, e.g. some graphical presentation of predictions under different scenarios of wind speed and wind direction.