Wind Power Prediction by using Ensemble Models

René Jursa

1 Institut für Solare Energieversorgungstechnik e.V., Königstor 59, 34119 Kassel, Germany
rjursa@iset.uni-kassel.de

Abstract. We compare structural different methods of the artificial intelligence for wind power prediction modeling and build additionally ensembles of the models. As input variables for these prediction methods weather data of a numerical weather prediction model are used. The performance of the presented methods is compared to the predictions of the neural network based model.

Keywords: Wind power prediction, comparative study, ensemble models.

1 Introduction

For wind power prediction physical models, statistical or artificial intelligence methods are used [1], [2]. We compare four methods for wind power prediction namely artificial neural networks (ANN), mixture of experts (ME), support vector machine (SVM) and nearest neighbour search (NNS) with a superior particle swarm optimization (PSO). In a similar recent study for wind speed prediction SVM was compared to ANN [3] and for wind power prediction a comparison of ANN with NNS [4] was studied. We use additionally ensembles of the structural differing models by averaging the model outputs.

2 Prediction Methods

Artificial Neural Network. An artificial neural network (ANN) consists of nonlinear functions, which are combined by a weighted series of linear filters [5]. As input values for every prediction time \( t \) NWP data for the location of the wind farm, i.e. six values of the wind speed \( ws(t - 1h), \; ws(t), \; ws(t + 1h) \) for a height of 10m and 100m above sea level and the sine and cosine of the wind direction \( wd(t) \) of 10m height are used. For the training of the ANN we use historical NWP data and historical measured power data and train the ANN with the back-propagation algorithm.

Mixture of Experts. The mixture of experts (ME) model [5] consists of some expert neural networks with outputs \( \hat{P}_i(t) \) and a gating network with outputs \( g_i(t) \). The output of the whole network is a linear combination of the experts, which can be regarded as an ensemble model of structural equal neural networks:

\[
\hat{P}(t) = \sum_{i=1}^{k} g_i(t) \hat{P}_i(t).
\] (1)
The experts and the gating network are trained with the expectation-maximization (EM) algorithm with the same input variables as used for the ANN model.

**Support Vector Machine.** The Support Vector Machine (SVM) maps the input data vectors into a high-dimensional feature space and constructs there a linear regression function [6]. The SVM uses so-called kernel functions for mapping non-linear relationships in the original data to linear relationships in the feature space. The SVM uses the same input data as the ANN model.

**Nearest Neighbour Search.** For the Nearest Neighbour Search (NNS) we construct a state space representation $\tilde{S}(t)$ from the wind speed $w_{si}$ and wind direction $wd_i$ data of several wind farm locations $i$ in a considered spread area. The part $\tilde{S}_i(t)$ for the $i$-th location of $\tilde{S}(t)$ has the following representation:

$$
\tilde{S}_i(t) = (w_{si}(t-s_{i,ws}), w_{si}(t-(s_{i,ws}+\tau_{i,ws})), \ldots, w_{si}(t-(s_{i,ws}+(d_{i,ws}-1)\tau_{i,ws})),
wd_i(t-s_{i,wd}), wd_i(t-(s_{i,wd}+\tau_{i,wd})), \ldots, wd_i(t-(s_{i,wd}+(d_{i,wd}-1)\tau_{i,wd})).
$$ (2)

The construction parameters $\tau_{i,ws}$, $s_{i,ws}$ and $d_{i,ws}$ denote the time delay, the time shift and the embedding dimension for $w_{si}$, respectively. Analog $\tau_{i,wd}$, $s_{i,wd}$, $d_{i,wd}$ denote the construction parameters for $wd_i$.

The state space representation of the whole system $\tilde{S}(t)$ with $M$ locations is a common time delay vector built from the $\tilde{S}_i(t)$:

$$\tilde{S}(t) = (\tilde{S}_1(t), \tilde{S}_2(t), \ldots, \tilde{S}_M(t)).$$ (3)

The corresponding power values of the neighbourhood state $\tilde{S}_n$ at time $t_n$ are used for the prediction model:

$$
\hat{P}(t) = \frac{1}{\sum w_n \sum_{s_i \in G(t)} w_n P(t_n)}.
$$ (4)

As weighting coefficients $w_n$ for the neighbours we take the bi-weight function to take the distances of the neighbours $\tilde{S}_n$ into account. With the construction parameters of $\tilde{S}(t)$ we select the input variables of the NNS. For this variable selection an optimization method namely the particle swarm optimization is used, which is a population-based non-linear optimization algorithm [7].

**Ensembles of the Models.** The presented models are used for the construction of ensembles, which have in general a better performance than the individual models [8]. In our study we build equal weighted ensembles by using the simple mean model output of the presented models $m_k$:

$$
M(m_1(t), m_2(t), \ldots, m_k(t)) = \frac{1}{K} \sum_{k=1}^K m_k(t).
$$ (5)
3 Results

Wind power predictions for 10 wind farms of the German E.ON control zone have been made by using NWP data from the German weather service DWD. For the evaluation of the predictions we choose by random 10000 test data points of the whole data set of about 25000 data points of the time period February 2004 to December 2006. The performance of our prediction methods is estimated by the root mean squared error (RMSE) scaled by the rated power $P_{rated}$ of the wind farms:

$$RMSE = \sqrt{\frac{1}{N_T} \sum_{n=1}^{N_T} \left( \frac{P_n - \hat{P}_n}{P_{rated}} \right)^2}.$$  

(6)

$P_n$ and $\hat{P}_n$ are the measured and predicted power respectively for every time step $n$ and $N_T$ is the number of test data points. To compare the results to our benchmark model, i.e. the ANN, we present the improvement (IMP) on the test data

$$IMP = \frac{RMSE_{ANN} - RMSE_{model}}{RMSE_{ANN}}.$$  

(7)

In figure 1 the improvements of the individual models compared to ANN are presented. The ME yields the best results for all wind farms. In figure 2 the improvements of several ensemble models are presented, and one can see that the best combination for nearly all wind farms is M(ME, NNS, SVM). As mean improvement over all 10 wind farms, we receive nearly 8.8% improvement in relation to the best single model, i.e. the ME model, and with the best combination of ME, NNS and SVM we receive a mean improvement of 14.8%.

Fig. 1. Comparison of the improvement of ME, NNS and SVM.
Fig. 2. Improvement of several Ensembles of the prediction models ANN, ME, NNS and SVM.

4 Conclusion

We can conclude that there is an advantage to use ensembles of structural different models for wind power predictions. In comparison to the benchmark model, i.e. the neural network model, we get the best improvement of almost 15% by using an ensemble of mixture of experts, nearest neighbour search and support vector machine. The best individual model with about 8.8% improvement are the mixture of experts as an ensemble of structural equal neural networks.

References