Hindawi Publishing Corporation
Journal of Renewable Energy
Volume 2014, Article ID 986830, 15 pages
http://dx.doi.org/10.1155/2014/986830

Research Article

Comparison and Optimization of Neural Networks and Network Ensembles for Gap Filling of Wind Energy Data

Andres Schmidt[1] and Maya Suchaneck[2]

[1] *Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA*
[2] *Department of Geography, Ruhr University Bochum, 44780 Bochum, Germany*

Correspondence should be addressed to Andres Schmidt; andres.schmidt@oregonstate.edu

Received 28 January 2014; Revised 15 April 2014; Accepted 16 April 2014; Published 26 May 2014

Academic Editor: Shuhui Li

Wind turbines play an important role in providing electrical energy for an ever-growing demand. Due to climate change driven by anthropogenic emissions of greenhouse gases, the exploration and use of sustainable energy sources is essential with wind energy covering a significant portion. Data of existing wind turbines is needed to reduce the uncertainty of model predictions of future energy yields for planned wind farms. Due to maintenance routines and technical issues, data gaps of reference wind parks are unavoidable. Here, we present real-world case studies using multilayer perceptron networks and radial basis function networks to reproduce electrical energy outputs of wind turbines at 3 different locations in Germany covering a range of landscapes with varying topographic complexity. The results show that the energy output values of the turbines could be modeled with high correlations ranging from 0.90 to 0.99. In complex terrain, the RBF networks outperformed the MLP networks. In addition, rare extreme values were better captured by the RBF networks in most cases. By using wind meteorological variables and operating data recorded by the wind turbines in addition to the daily energy output values, the error could be further reduced to more than 20%.

1. Introduction

The Combination of climate change and the dependence on fossil fuels slowly cause changes in energy policy and trigger an increasing demand for sustainable energy sources. Global carbon dioxide emissions are ever increasing and the associated consequences for the climate are widely scientifically recognized [1–3]. Over the last decade and in particular since the release of the report of the Intergovernmental Panel on Climate Change (IPCC) in 2007, public and political awareness of renewable energy technologies has increased considerably. This is not at least due to the large and fast growing economies and the associated increase of numbers of cars and energy consumption, and therefore of CO_2 emissions [4]. Wind energy has the potential to be a vital contributor to renewable energy technologies that will substitute more and more for gas and coal [5].

In order to decrease the uncertainty of wind energy yield predictions during the planning of a single turbine or wind farm, data of nearby existing wind turbines are often used as reference for model evaluation.

In Germany, the legislation that grants priority to renewable energy sources (Renewable Energy Resources Act, EEG) states that only wind turbines in areas with sufficient wind energy potential are qualified to receive compensation for the electrical power provided for the power grid [6].

Moreover, according to the EEG, power grid owners are not required to connect turbines to the grid that does not meet or exceed 60% of the turbine type-specific reference value calculated based on reference wind conditions.

Because of these restrictions, a correct prediction for expected wind energy yield is an indispensable economic criterion for most wind farm projects. However, because of technical malfunctions, maintenance routines, or other problems, the availability of valuable comparison data from existing turbines nearby is limited. Such data limitation affects the statistical safety of model predictions of future energy yields.

Artificial neural networks are able to approximate nonlinear relationships between individual data series by adjusting network parameters in a purely data-driven, and, in our case, supervised learning process.

Here, we present a method to model the data of nearby wind turbines using different types of neural networks and network ensembles to fill in gaps in time series of wind energy outputs. Real-world operating data of six exemplary wind farms in Germany were available for this purpose.

2. Data and Methods

To provide a high planning dependability, the potential annual wind energy yields have to be estimated carefully during the planning process of a wind turbine. The accuracy of predictions becomes even more crucial when wind farms are planned due to the substantially increased financial risk. Several microscale and mesoscale wind flow models based on computational fluid dynamics are available to calculate future wind energy yields of single turbines and wind farms. Typically, terrain data such as surface roughness, orography, and existing wind obstacles for a radius of 20 km around the proposed location of the wind turbine is considered for the model of computations in combination with long-term wind statistics.

One widely used model, for instance, that has passed several stages of development over the last decades is *WAsP* (Wind Atlas Analysis and Application Program) [7]. *WAsP* was developed by the *RISØ* National Laboratory, Roskilde, Denmark, and is approved among others by the German Federal Research Ministry.

The various specific models that account for wind obstacles, orography, wake effects, and surface roughness were assembled in the model suite of *WAsP*. A detailed description of the model and its algorithms can be found in the respective literature [8].

In order to validate the predicted long-term average wind energy yields based on the model results considering the environmental conditions, the geometry of the turbine and its power curve [9] data of existing, nearby turbines are used as a reference. For that purpose, energy output values of existing wind turbines are corrected for technical availability. In addition, the wind conditions and the corresponding observed electrical energy output values of a certain year are compared and scaled to long-term data for the area of interest [10]. After correcting the observed values through linear regression to 100% long-term averages, the values can be compared to the model results that are based on long-term wind statistics and therefore also represent long-term averages [11].

2.1. Artificial Neural Networks. Artificial neural networks provide a method to map input variables on target variables by using a combination of nonlinear functions and a learning procedure that can be supervised or unsupervised [12, 13]. The ability to find mathematical functions without prior knowledge of the functional relationship makes neural networks a powerful tool to solve problems for which no analytical solution exists or if the function that relates variables to each other is unknown [14–16]. Artificial neural networks have found a growing range of applications in recent years including the field of wind energy research [17, 18]. The ability to self-adjust its parameters makes neural networks a superior fitting method compared to classical data-fitting and prediction methods [19–21].

The most popular network architecture currently used is the so-called multilayer perceptron (MLP) topology, which was presented comprehensively first by Rumelhart et al. [22].

Within this study we compare the performance of MLP networks to the performance of radial basis function (RBF) neural networks. Compared to MLP networks, the parameters of RBF networks can be adjusted faster to the data presented to the network during the training process. In addition, RBF networks are less affected by the problem of local minima [23]. MLP networks use the scalar product of the input data vector consisting of n input variables and a weight vector to calculate the neuron output usually applying hyperbolic activation functions. In contrast to that, RBF networks use the distance between the input vector and the center of the radial basis function to determine the activation value within an RBF neuron. The most commonly used activation function type with a radial basis is the Gaussian function that was also deployed in the networks used for this study. The n-dimensional vectors that determine the shape of each neuron's n-dimensional Gauss function in the input space are defined by the center vector μ and the n-dimensional variance vector σ.

The central values of μ have been optimized within this work using the k-means clustering algorithm. The input vectors were separated into K clusters and the center values of the RBF neurons were set to represent all considered input vectors while minimizing the number of clusters by finding the centers for each cluster with the smallest mean squared distance to all points in the cluster. According to Hestenes [24], this can be calculated by minimizing the function J as given in ((1) and (2)). Consider

$$J = \sum_{j=1}^{K} \sum_{i \varepsilon S_j} \left\| x_{ij} - \mu_j \right\|^2, \tag{1}$$

with

$$\mu_j = \frac{1}{N_j} \sum_{i \varepsilon S_j} x_j. \tag{2}$$

Here, μ_j is the mean of subsample S_j which the cluster j is composed of. The index i is the index over the subsample S_j. At the end of the procedure, each input vector x is assigned to the cluster center (i.e., RBF network node) to which it has the least Euclidean distance. The closer an input vector to the RBF center of a neuron is, the higher the activation value of that neuron is. Hence, the parameters of all RBF functions are adjusted during the network learning phase so that every input can be assigned to one of the RBF neurons in the hidden layer and the weighted sum of the activation functions can be transformed to satisfyingly match the target values. More details about the large field of learning algorithms for neural networks can be found in the literature [12, 13, 16].

For the network training and evaluation process, each available learning dataset was divided into 3 different dataset sections: training data, test data, and validation data. The training data is transformed by the network functions. After the output of the network is compared to the real available measured outputs, the error is determined. The error value used for this performance check is the final RMSE (root of the mean squared error) summing up the differences between modeled output and known measured output according to

$$E = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(m_j - y_j\right)^2}. \quad (3)$$

Here, m_j is the measured result within the available learning dataset (i.e., data that contains all values for all input variables and measured samples of the corresponding target value) and y_i is the value estimated by the neural network. N is the number of data records used to determine the error E for the training data and test or validation data, respectively.

The parameters of the network are, then, iteratively adjusted applying the conjugate gradient descent method for the training algorithm [12, 24] until an acceptable error is reached.

In a second step the test data is used. Test data was not used for the network parameter adjustment during the training. The trained network is now applied to the test dataset and the RMSE is determined again.

If the error for the training data is small and the error for the test data is large, this indicates that the network parameters have been over adjusted to the training data. Hence, the training has to be started again until minima for both, the training error and the test error, have been reached.

Once the optimal balance between training data error and test data error is reached, the network has the ability to sufficiently generalize the functional relationship between input and target variable well and is not over-fitted to a specific training dataset.

The third dataset section, the validation data, has not been used for the learning process of the network at all and is processed with the adjusted neural network in a final step. Hence, the validation error is the only of the three error values that is admissible to assess the goodness of a neural network model. The division of the data into the training data, test data, and validation data was conducted according to values presented in the literature [25, 26]. The validation error indicates the final performance of the trained network when applied on new data and is therefore given in the results presented.

In order to scale all input variables to values between 0 and 1 in a preliminary step we applied min-max normalization to prevent the network results from being biased by the stronger numerical influence of a variable measured in units with larger numbers than a variable that is limited to a smaller scale by default [27]. For each dataset, we trained specific neural networks and determined the three best networks. For that purpose, a batch algorithm was conducted to test 20,000 networks for each dataset while varying the number of neurons, number of input variables, network type, and activation functions in the nodes.

To demonstrate the potential of the method, three pairs of turbine sites were selected with different large-scale wind conditions, whereupon the wind conditions amongst the pairs are similar due to the relative spatial adjacency.

2.2. Real-World Case Studies with Operating Data of Exemplary Wind Farms. The daily energy output W of existing wind turbines provide important information and are used as input data for the neural networks. Also, the mean values, minima and maxima of wind speed, wind direction, instantaneous power, and number of rotor revolutions recorded by the operating and surveillance software were taken into account for the neural network calculation. The three exemplary test areas are located in Germany where the prevailing wind conditions are dominated by the global West wind drift of the temperate latitudes in the Northern hemisphere.

The locations were selected with respect to their different orographic features to validate the robustness of the method presented.

It is a coastal location, a site in the *Muensterland* lowland area in Northwest Germany, and an area located in the mountainous forested uplands of West central Germany (Figure 1).

Overall, the data from 21 modern gearless wind turbines of the manufacturer Enercon (ENERCON GmbH, 26605 Aurich, Germany) recorded over a period of four years were available for the calculations.

3. Results and Discussion

In the following section the results for 3 case studies are presented, consecutively with increasing topographic complexity. In each case, we provide detailed information about turbine types, wind conditions, and topographic circumstances that affect the local wind fields for the distinct types of terrain.

In each case, specific neural networks were trained to reproduce the electrical energy output using the data of existing wind turbines with various distances to the respective target turbines and wind farms. The performances of the best trained networks are presented and errors values are given and compared.

3.1. The Coastal Sites. Due to the flat landscape with no significant changes in elevation, orographic effects on the wind speed are negligible around the coastal sites of *Hinte* and *Jennelt*.

The site belongs to the district of the city of Aurichin the East Frisia region in Lower Saxony, Germany. The North Sea is located in a distance of about 12 km to the West of the site *Hinte*. The nearest city Emden is located at a distance of 5.8 km South of the wind farm *Hinte*. The immediate vicinity of the site is characterized by meadows and agriculturally used land. The terrain is flat and mostly free of wind obstacles.

The wind farm is located 1.7 km Northwest of *Hinte*. The wind farm consists of a total of 15 wind turbines operated by different companies with an average annual energy output of

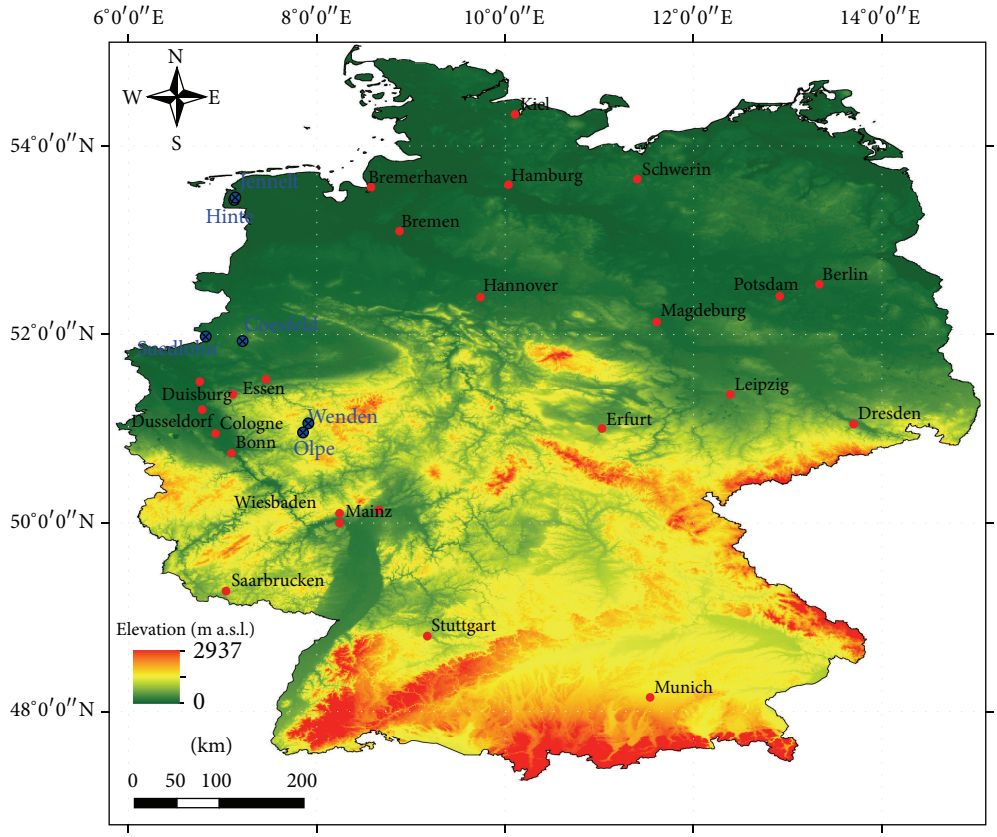

Figure 1: Shaded relief overview map of the incorporated wind farm sites in Germany. The elevations shown are based on the NASA Shuttle Radar Topography Mission data.

Table 1: Coordinates and data of the wind turbines at the site *Hinte*.

Turbine	Type	Hub height [m]	Rotor Ø [m]	Nominal power [kW]	Longitude (WGS 84)	Latitude (WGS 84)	Elevation (m a.s.l.)
HI 1	E 66/18.70	65.0	70.0	1800	7.170441E	53.420212N	0
HI 2	E 66/18.70	65.0	70.0	1800	7.166404E	53.424987N	0
HI 3	E 66/18.70	65.0	70.0	1800	7.166359E	53.422247N	0

Table 2: Coordinates and data of the wind turbines at the site *Jennelt*.

Turbine	Type	Hub height [m]	Rotor Ø [m]	Nominal power [kW]	Longitude (WGS 84)	Latitude (WGS 84)	Elevation (m a.s.l.)
JE 1	E 66/18.70	65.0	70.0	1800	7.137170E	53.448060N	2
JE 2	E 66/18.70	65.0	70.0	1800	7.132560E	53.456434N	2
JE 3	E 66/18.70	65.0	70.0	1800	7.139583E	53.456116N	1
JE 5	E 66/18.70	65.0	70.0	1800	7.139436E	53.452882N	1
JE 5	E 66/18.70	65.0	70.0	1800	7.125192E	53.452810N	2

47 GWh. The specifications and geographical coordinates of the turbines used for the calculations are given in Tables 1 and 2, respectively.

Due to ongoing expansions and changes, the wind farm information given in this study always refer to the turbines with data incorporated in the analyses and do not necessarily represent the current total number of installed turbines.

The wind farm *Jennelt* is located at 4.1 km to the North-west of the wind farm *Hinte*. The North Sea lies West of the site in a distance of 7.4 km.

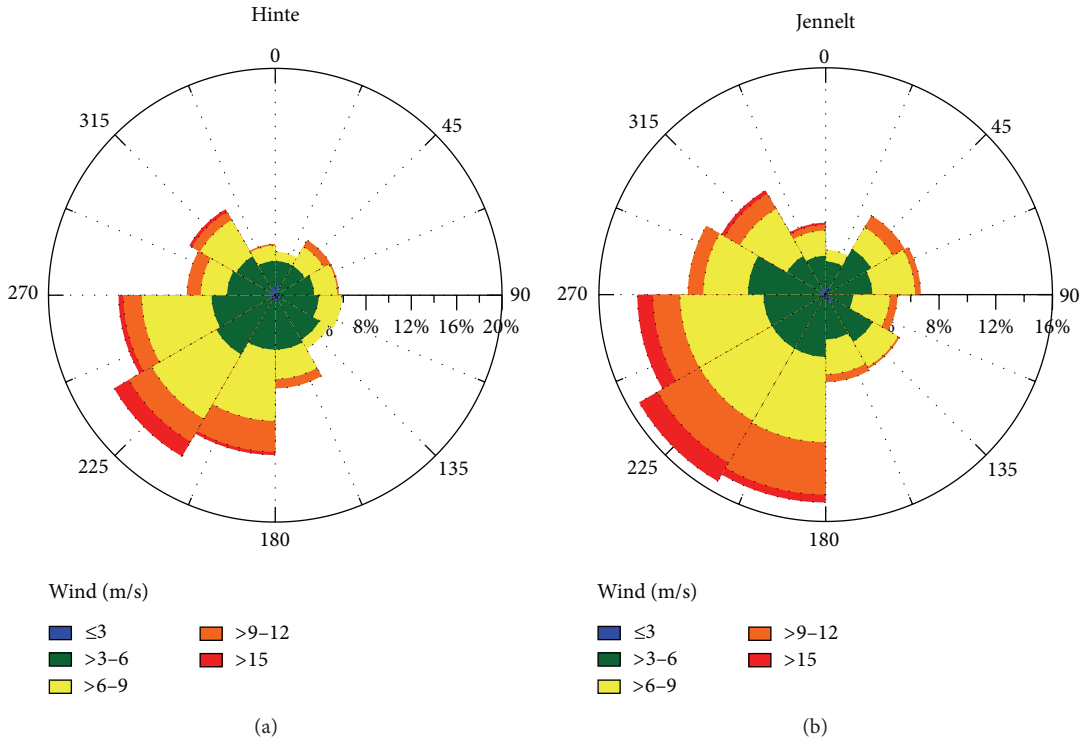

Figure 2: Average wind conditions at the coastal sites in *Hinte* (target site) and *Jennelt* (input site) during the observation period from January 2006 to January 2008.

The data collection period for the site pair *Hinte*/*Jennelt* is two years, from January 1, 2006, to December 31, 2008. Thus, 732 data records were available for the two sites.

The wind conditions at the sites *Hinte* and *Jennelt* measured by anemometers on the nacelle of the wind turbines (i.e., 67 m above ground level) were similar with regard to the wind speeds and the distribution of wind directions during the data collection period (Figure 2). Even though the wind measurements on the nacelle of the plants are subjected to certain errors by flow distortions, they are still sufficiently accurate to capture the speed and direction for most applications [28].

Comparisons between the wind speeds measured on the nacelle anemometer and an undisturbed measurement of wind in front of the rotor show only minor deviations of 2% and less [29].

600 daily data records were used for the training of the neural networks. The remaining records were used to the same parts for the test dataset and the validation dataset.

The nonparametric Spearman rank correlation coefficient r_s was used for the comparison of the measured electrical energy output W and the results reproduced by the neural networks. The RMSE values given refer to the validation data set that was composed of values that were randomly distributed over the entire dataset. In Table 3, the results for the three best networks are shown, that is, the networks with the lowest validation errors and highest correlation coefficients.

Table 3: Summary of results for the energy output of the target turbines at the coastal site *Hinte*.

Target variable	Input variables	Network type	Network topology	RMSE (kWh)	r_S
$W_{\text{HI 1}}$	$W_{(\text{JE 1–JE 5})}$	MLP	5-6-1	1326	0.98
$W_{\text{HI 1}}$	$W_{(\text{JE 1–JE 5})}$	RBF	5-44-1	1312	0.99
$W_{\text{HI 1}}$	$W_{(\text{JE 1–JE 5})}$	RBF	5-80-1	1458	0.99
$W_{\text{HI 2}}$	$W_{(\text{JE 1–JE 5})}$	MLP	5-10-1	1794	0.99
$W_{\text{HI 2}}$	$W_{(\text{JE 1–JE 5})}$	RBF	5-51-1	1911	0.99
$W_{\text{HI 2}}$	$W_{(\text{JE 1–JE 5})}$	RBF	5-53-1	2208	0.98
$W_{\text{HI 3}}$	$W_{(\text{JE 1–JE 5})}$	MLP	5-6-1	2256	0.98
$W_{\text{HI 3}}$	$W_{(\text{JE 1–JE 5})}$	RBF	5-42-1	2244	0.98
$W_{\text{HI 3}}$	$W_{(\text{JE 1–JE 5})}$	RBF	5-80-1	2330	0.98

All correlations given are significant on a 95% confidence level ($P \leq 0.05$). The network topology refers to the number of input variables, the number of hidden nodes, and the number of output nodes of the neural networks used. The number of hidden neurons within RBF networks is usually higher than for the MLP type networks to accomplish the same ability of generalization [13, 30].

The differences in network performance are overall small. All networks achieved very high correlations of 0.98 and 0.99 and the statistics for the measured and reproduced power output data are similar (Tables 3 and 4).

Figure 3 shows the observed and network reproduced frequency distributions of the daily values of W for the

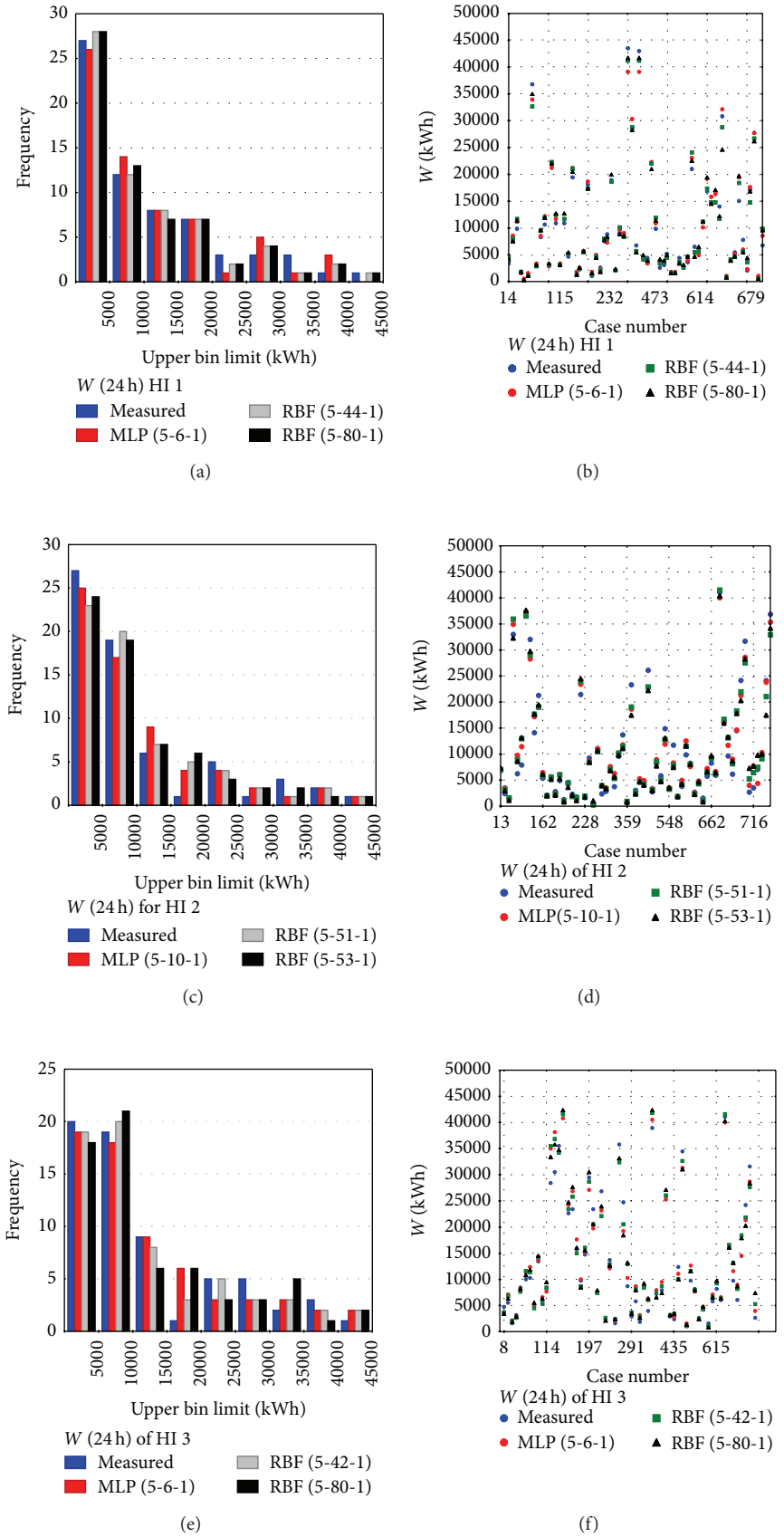

FIGURE 3: Results for the 3 best networks for the dataset of the coastal wind farms *Hinte* and *Jennelt*. The case numbers in the right panels represent individual data records of the complete time series available that were chosen randomly and assigned to the validation dataset.

Table 4: Descriptive statistics of the measured MES time series and the time series modeled by the best artificial neural network ANN calculated over all available daily values.

Target variable	Average		Standard deviation		Maximum		Minimum	
	Measured	ANN	Measured	ANN	Measured	ANN	Measured	ANN
$W_{\text{HI 1 (kWh)}}$	11137	11104	9918	9743	43627	42695	186	90
$W_{\text{HI 2 (kWh)}}$	11477	11452	10358	10138	44491	43017	168	205
$W_{\text{HI 3 (kWh)}}$	11220	11213	10123	9895	43589	41752	216	251

validation data set by the three best neural networks for the three target turbines H1 (a), H2 (c), and H3 (e). The frequency distributions show that all networks produce a similar value range of the energy output for the three turbines indicating that the functional relationships between the input and output variables are well captured by all 3 best networks (Table 3). It is noteworthy that, for the rare peak values of over 40 MWh of turbine HI 1, the RBF networks exhibit smaller differences to the measured data than the MLP networks (Figures 6(a) and 6(b)). These extreme values are not outliers in the statistical sense but rare, yet physically meaningful, values.

The fact that RBF networks outperform MLP networks while generalizing the underlying functional relation between input and output also covering extreme values has also been observed in other studies [31, 32].

Also the direct comparisons of individual nonconsecutive samples of the validation dataset (Figures 3(b), 3(d), and 3(f)) show that the two RBF networks estimate values on either side of the end of the scale are the best. The first and second order statistic moments of the measured and modeled data are in good agreement (Table 4).

Prior to the learning process, three continuous periods of approximately three weeks each were extracted from the dataset. This data was then calculated with the best neural network for each turbine of the target wind farm (Figure 4). In contrast to the actual validation data (Figure 3) that were also excluded from the training processes, this artificially created data gap simulates real-world situations when continuous gaps in the records for several days due to technical problems or for several hours due to machine maintenance may occur. The energy output values of the system of wind turbines could be reproduced closely with the neural networks.

One reason for the very reliable and accurate calculation ($0.98 \leq r_S \leq 0.99$) is given by the similarity of the data. The wind conditions at the locations of the wind turbines are not compromised by a complex terrain or large topographic elements. Furthermore, the turbines of both wind farms are of the same type with the same hub height. This assumption about the initial similarity of the data of the two wind farms is supported by the linear regression matrix for the measured data of the two wind farms (Figure 5).

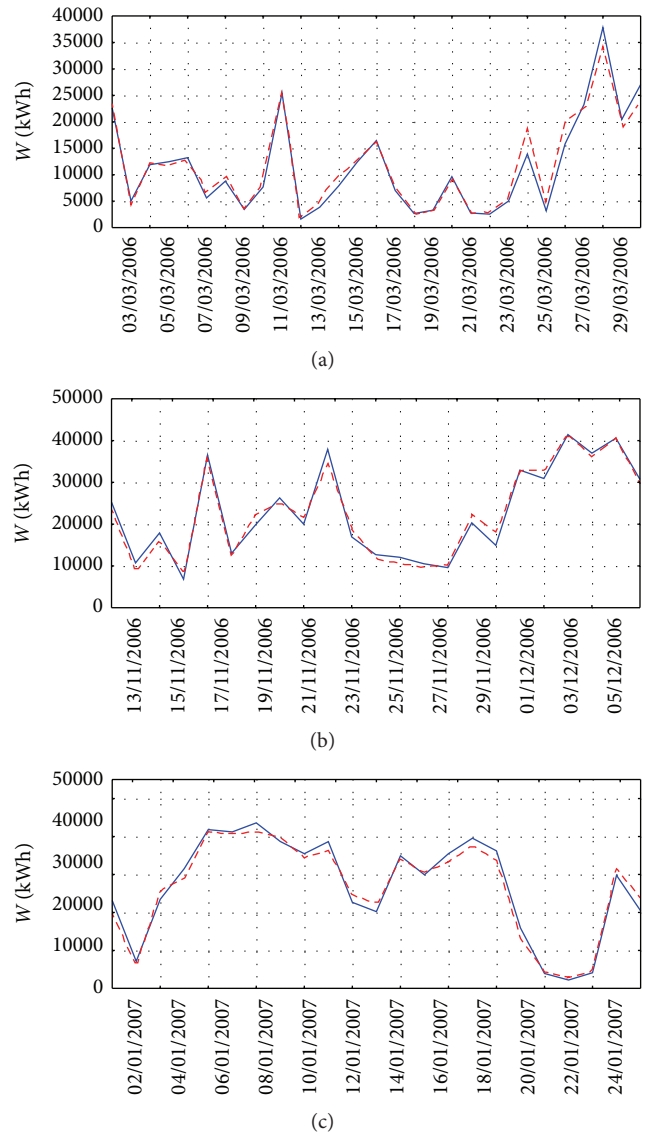

Figure 4: Comparison of the measured (blue continuous line) and network reproduced (red dashed line) daily energy values of the turbines at the coastal site.

3.2. Sites in the Westphalian Basin. With regard to the orography, the *Muensterland* lowland area around the sites of *Coesfeld* and *Suedlohn* can be classified as a transitional landscape between the flat area in the very Northwest of Germany close to the North Sea and the low mountain range to the Southeast.

Despite some slight elevations with weak slopes, the landscape is relatively flat as typical for the Westphalian Basin in the Southern marginal area of the North German

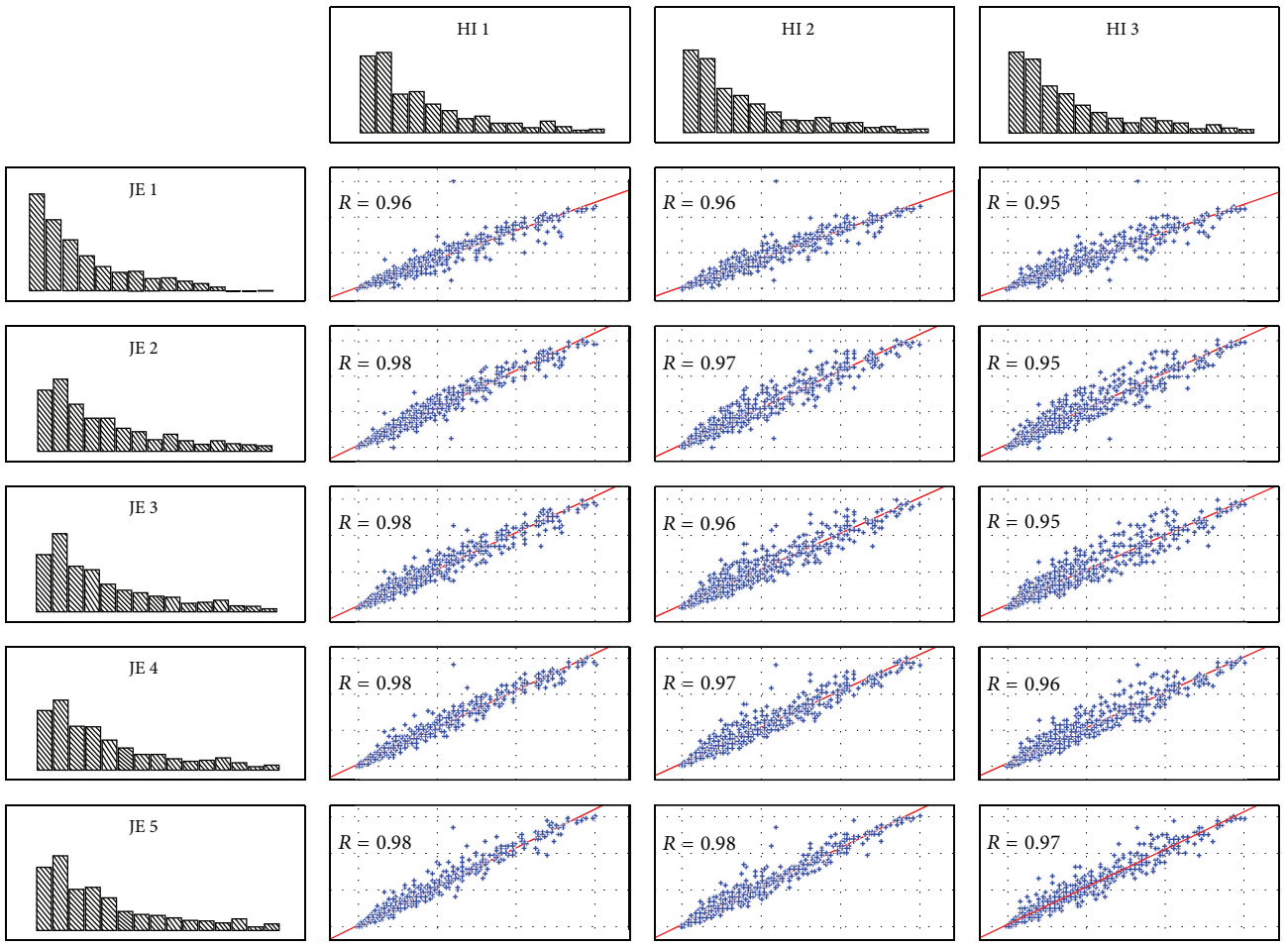

Figure 5: Graphic correlation matrix with linear regression coefficients R for the daily energy yields at the coastal sites for the entire data period from the beginning of 2006 through the end of 2008. The frequency distributions are equally divided into 15 classes for all distributions.

Plain. The surface roughness is generally higher compared to the coastal sites *Hinte* and *Jennelt* due to many small towns and some small forests that disrupt the predominantly agricultural region.

The wind farm in *Coesfeld* consisting of five wind turbines is located at a distance of 24 km South-East of the town *Coesfeld*. The farm is composed of two different types of wind turbines, two turbines of the type E-58/10.58, and three larger turbines of the type E-66/10.70. The topography in the immediate vicinity around the wind farm is slightly undulating. The specifications and geographical coordinates of the turbines in the Westphalian Basin area used for the calculations are given in Tables 5 and 6, respectively.

The target wind farm *Suedlohn* is located at a distance of about 1.5 km East of the border to the Netherlands. The wind farm is located Southwest of the community of *Suedlohn*, 24 km away from the wind farm *Coesfeld*. As for the wind farm *Coesfeld,* the topography in *Suedlohn* is expected to have some effects on the wind regime due to upwind obstruction effects at hills [8, 33]. The near surrounding of the site is dominated by arable land, pasture, and widespread farm buildings.

At the site *Suedlohn,* noise emissions restrictions produce legal reasons to reduce the sound level of the turbines affecting the farm houses in the vicinity of the turbines during nighttime. Therefore, the turbines SU 2 and SU 3 are operated with sound reduced performance characteristics from 22:00 to 6:00 which reduces the nominal power output from 800 kW to 600 kW.

The data collection period for the two sites in the Westphalian Basin covers two years from January 1, 2006, to January 1, 2008, providing 732 daily data records. The wind is dominated by winds from West and Southwest at both wind farms because no larger topographic structures are affecting the superimposed West wind drift.

However, the Southwesterly wind directions exhibit a slightly higher frequency at the site *Coesfeld* (Figure 6). In order to achieve the best comparability anemometer data from turbine CO 1 was used for the comparison of the wind conditions as the hub height of 70.5 m is closest to the hub height of 75.6 of the turbines at the target site *Suedlohn*.

The correlation coefficients (Table 7) indicate that the combination of increased topographic structure and the distance of 25 km between the wind farm delivering the input

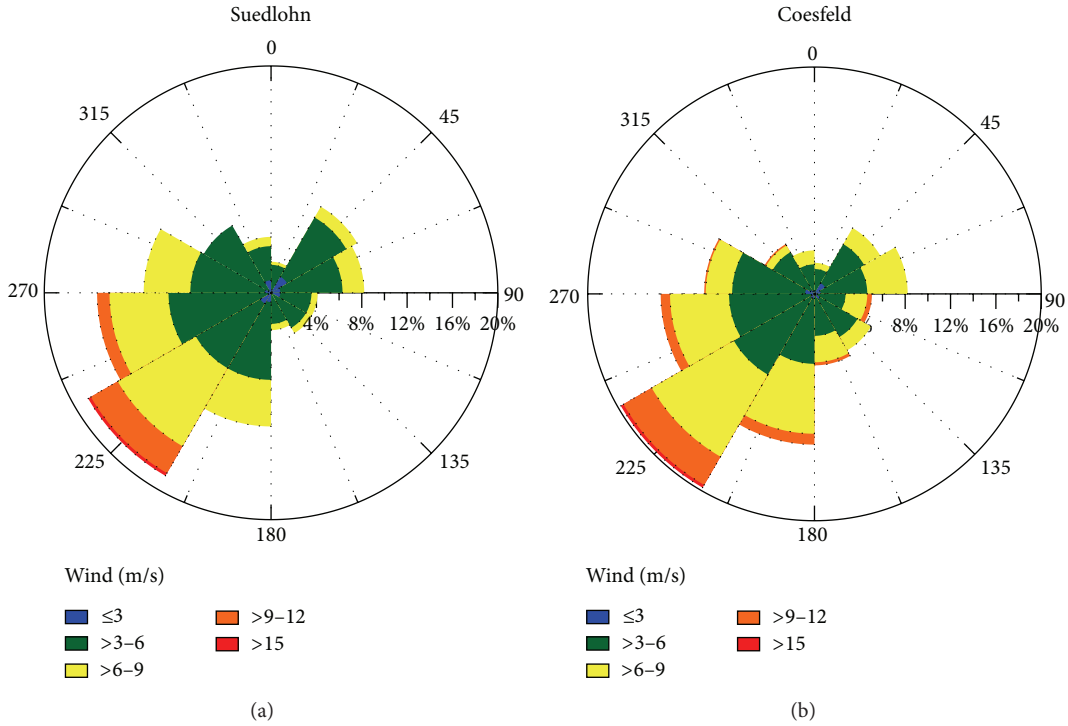

Figure 6: Average wind conditions at the locations *Suedlohn* (target site) and *Coesfeld* (input site) during the data collection period from January 2006 to January 2008.

Table 5: Coordinates and data of the wind turbines at the site *Coesfeld*.

Turbine	Type	Hub height [m]	Rotor Ø [m]	Nominal power [kW]	Longitude (WGS 84)	Latitude (WGS 84)	Elevation (m a.s.l.)
CO 1	E-58/10.58	70.5	58.6	1000	7.2321950E	51.9316590N	136
CO 2	E 66/18.70	98.0	70.0	1800	7.2198140E	51.9311870N	101
CO 3	E 66/18.70	86.0	70.0	1800	7.2113600E	51.9261410N	121
CO 4	E 66/18.70	98.0	70.0	1800	7.2247980E	51.9358810N	114
CO 5	E-58/10.58	89.0	58.6	1000	7.1940630E	51.9213830N	91

Table 6: Coordinates and data of the wind turbines at the site *Suedlohn*.

Turbine	Type	Hub height [m]	Rotor Ø [m]	Nominal power [kW]	Longitude (WGS 84)	Latitude (WGS 84)	Elevation (m a.s.l.)
SU 1	E 48/8.48	75.6	48.0	800	6.828237E	51.950485N	47
SU 2	E 48/8.48	75.6	48.0	800/600	7.086687E	51.952653N	48
SU 3	E 48/8.48	75.6	48.0	800/600	7.140800E	51.952483N	53

Table 7: Correlation matrix for the daily energy output at the locations of *Coesfeld* and *Suedlohn* for the analyzed period.

	CO 1	CO 2	CO 3	CO 4	CO 5
SU 1	0.52	0.53	0.51	0.52	0.56
SU 2	0.54	0.53	0.52	0.53	0.57
SU 3	0.52	0.53	0.48	0.53	0.54

All correlations shown are significant on a 95% confidence level.

data (*Coesfeld*) and the target wind farm (*Suedlohn*) lead to a functional relation between, then, input and target variables that cannot be described well by a simple linear approach.

Thus, due to their ability to emulate any nonlinear functions, neural networks provide an ideal tool for such situations. The available energy output data were divided randomly in the training data, test data, and validation data (602 training data records, 65 test data records, and 65 validation data records).

Table 8: Summary of the results for the target site *Suedlohn* in the Westphalian Basin within the area of the North German Plain.

Target variable	Input variables	Network type	Network topology	RMSE (kWh)	r_S
$W_{SU\,1}$	$W_{(CO\,1\text{–}CO\,5)}$	MLP	5-6-1	1154	0.95
$W_{SU\,1}$	$W_{(CO\,1\text{–}CO\,5)}$	MLP	5-13-1	1092	0.95
$W_{SU\,1}$	$W_{(CO\,1\text{–}CO\,5)}$	RBF	5-33-1	1153	0.95
$W_{SU\,2}$	$W_{(CO\,1\text{–}CO\,5)}$	RBF	5-25-1	513	0.97
$W_{SU\,2}$	$W_{(CO\,1\text{–}CO\,5)}$	RBF	5-21-1	588	0.96
$W_{SU\,2}$	$W_{(CO\,1\text{–}CO\,5)}$	MLP	5-8-1	627	0.96
$W_{SU\,3}$	$W_{(CO\,1\text{–}CO\,5)}$	MLP	5-8-1	917	0.97
$W_{SU\,3}$	$W_{(CO\,1\text{–}CO\,5)}$	RBF	5-15-1	1179	0.95
$W_{SU\,3}$	$W_{(CO\,1\text{–}CO\,5)}$	RBF	5-17-1	1026	0.96

Table 9: Descriptive statistics of the measured MES time series and the time series modeled by the best artificial neural network ANN calculated over all available daily values.

Target variable	Average		Standard deviation		Maximum		Minimum	
	Measured	ANN	Measured	ANN	Measured	ANN	Measured	ANN
$W_{SU\,1\,(kWh)}$	3980	3995	3402	3296	17188	15327	363	264
$W_{SU\,2\,(kWh)}$	3582	3637	3229	3327	16103	15234	300	229
$W_{SU\,3\,(kWh)}$	3631	3619	3215	3062	16247	15560	287	592

The neural network modeling results for the energy yield values of the turbines SU 1 to SU 3 are given in Table 8 considering all results of the respective three best networks.

Only the power output values of all five turbines were used as input vectors for all neural networks.

The correlations of the network reproduced values with the corresponding measured values are all high and statistically significant ($r_S = 0.95$, $P \leq 0.05$) for the validation data sets of all target turbines (Table 8).

Since turbine SU 2 was affected most by the noise reduction restrictions, the absolute RMSE values are smallest for that turbine which has to be taken into account while interpreting the relatively low RMSE values. Nevertheless, the associated Spearman rank correlation coefficient shows that, despite altering operating conditions during nighttime, the neural networks produce satisfactory results (Tables 8 and 9). Nevertheless, the scattering of differences between modeled and observed values when considering multiple networks is an unwanted effect.

Since the accumulated RMSE is calculated from the squared residual values, the direction of deviation is not considered when constraining the length of the network learning process as shown in (3). Network ensembles provide a way to reduce these variations among a group of neural networks trained for the same purpose.

By weighted averaging of the output values of individual networks, the variations are smoothed and deliver improved results [34, 35]. The weight factors for the averaging were determined using the RMSE values of the incorporated networks.

It is noteworthy in the context of the present work that the results from the individual networks are satisfactory. For the application in practice, there is at least no urgent need for improvement. Nevertheless, since this study outlines the usefulness of neural networks for energy output modeling for wind turbines in order to close data gaps, the performance of a network ensemble was analyzed for turbine SU 1.

As shown in Table 10, the results could be improved using an ensemble composed of three neural networks. In comparison to the values gained through the best single RBF network (Table 8), r_s improved from 0.95 to 0.97, while the RMSE could be reduced by 82 kWh.

The distance between the two wind farms in the *Muensterland* lowland area is more than 20 km. Moreover, the more complex topography leads to a relationship between the power output values of the two wind farms that can hardly be captured by simple linear functions (Table 7). Nevertheless, the results show that the neural network approach delivers sound results using data that were collected over a period of 2 years only.

3.3. Mountainous Sites with Complex Topography. The region around the third pair of sites is located in the mountainous area within the Rhine Massif near the Southern rim of the Ebbe Mountains in central Germany (Figure 1) in the densely forested *Sauerland* region.

Accordingly, the topographic and in particular the orographic conditions will have a significant influence on regional scale and local scale wind fields [36]. Caused by the vast forested areas with different stand ages and scattered towns, the surface roughness is significantly higher compared to the previously presented sites causing higher shear stress on the airmass. These topographic conditions make the wind field modeling challenging and make a simple

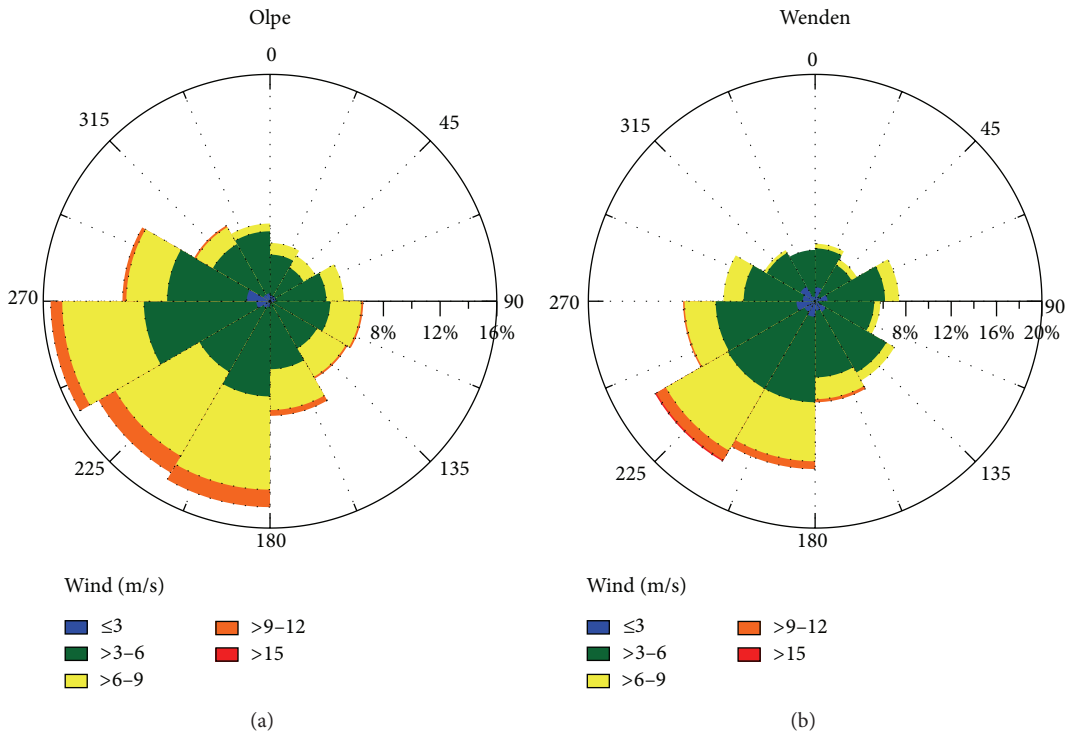

Figure 7: Average wind conditions at the sites *Olpe* and *Wenden* during the period from January 2004 through January 2008.

Table 10: Summary of the results for the calculation of the income of investment.

Target variable	Input variables	Network type	RMSE (kWh)	r_S
$W_{SU\,1}$	$W_{(CO\,1\text{–}CO\,5)}$	Ensemble (2 × MLP + 1 × RBF)	1010	0.97

SU 1 using an ensemble from the previously optimized individual networks.

spatial interpolation of wind velocities and wind energy yields, respectively, unfeasible. The technical specifications and geographical coordinates of the turbines used for the calculations for the mountainous site are given in Tables 11 and 12, respectively.

The area is mostly used for forestry and grazing. The wind farm *Wenden* is located 1.2 km South of the village *Wenden*. The wind farm *Olpe* delivering the input data for the neural network training and calculations is located 11 km North-Northeast of the wind farm *Wenden* and about 3.6 km Northeast of the town *Olpe*. The topography in the vicinity of both wind farms is undulating.

The data collection period for the turbines in *Olpe* and *Wenden* exceeds four years from January 1, 2004, until April 1, 2008. However, because the input vector is mathematically mapped on the output value by the neural networks, the number of variables in the input vector cannot be altered from the dimensionality used for the training and adjustment procedure when applying the trained network to new data. Hence, 1430 daily records that contain the synchronal operating data of all five wind turbines were available.

The wind conditions measured at 80 m above ground show significant differences between the two wind farm sites. At both locations the Southwestern wind directions exhibit the highest wind speeds. The West-Southwesterly and South-Southwesterly wind directions are more frequent in *Olpe* compared to the site in *Wenden* (Figure 7). The distributions of the wind direction indicate the effect of the orography in the complex terrain creating pronounced local scale wind fields with distinct distributions of wind velocity and direction at the two sites.

A training dataset of 1100 data records was used for the calculations on the site pair *Olpe*/*Wenden*. The test dataset and the validation dataset consisted of 165 data records each.

The Spearman rank correlation coefficients r_S are lower compared to the coastal sites and the sites in the Westphalian Basin area presumably caused by the more complex, nonlinear relationship between wind energy yields of the two wind farms. Nevertheless, correlations are still high reaching values from 0.90 to 0.93 (Table 13).

The comparison of the frequency distributions show that the energy output values in the most occupied class of less than 1000 kWh per day are better captured by the RBF networks in both cases (Figure 8). Also, for the other bins, the frequencies calculated by the best RBF agree best with the measured frequencies with the exception of a single value above 9000 kWh calculated by the RBF (topology 3-22-1) that

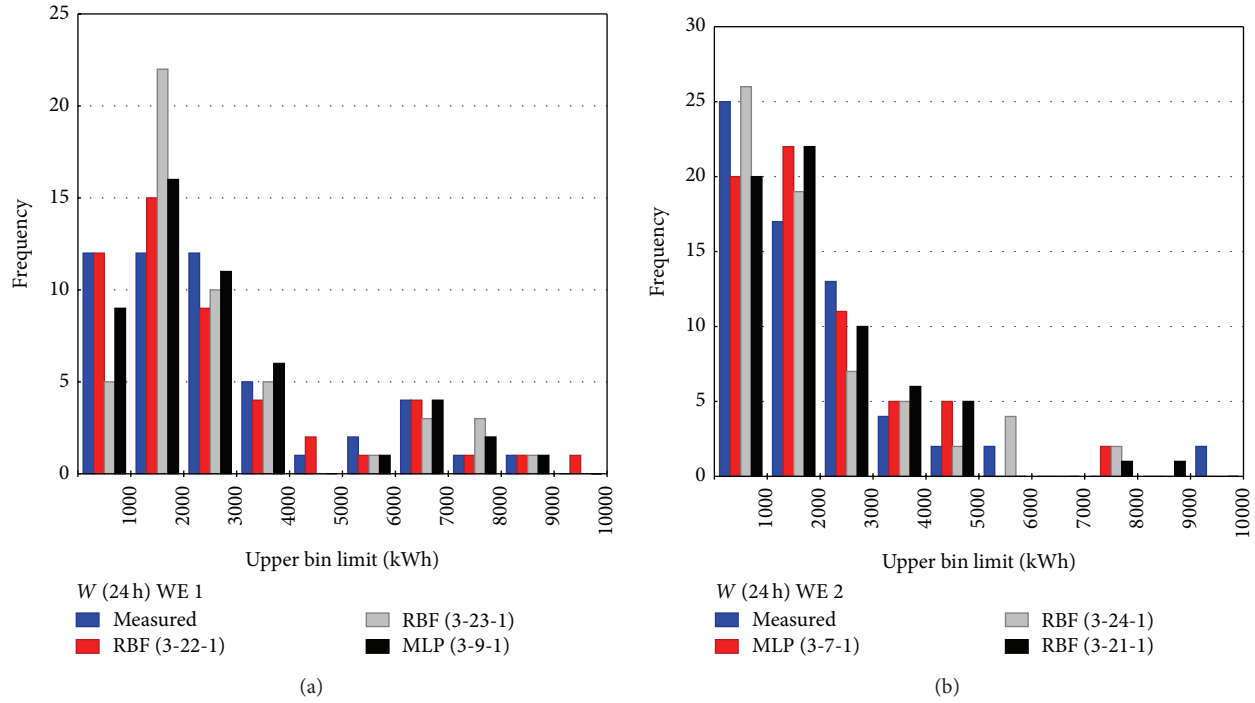

Figure 8: The frequency distributions for the measured data and the data calculated by the respective 3 best networks for the turbines WE 1 (a) and WE 2 (b).

Table 11: Coordinates and data of wind turbines at the site *Olpe*.

Turbine	Type	Hub height [m]	Rotor Ø [m]	Nominal power [kW]	Longitude (WGS 84)	Latitude (WGS 84)	Elevation (m a.s.l.)
OL 1	E 40/6.44	78.0	44.0	600	7.913383	51.054727	520
OL 2	E 66/18.70	86.0	70.0	1.800	7.917154	51.055103	528
OL 3	E 66/18.70	98.0	70.0	1.800	7.910023	51.053212	512

Table 12: Coordinates and data of wind turbines at the site *Wenden*.

Turbine	Type	Hub height [m]	Rotor Ø [m]	Nominal power [kW]	Longitude (WGS 84)	Latitude (WGS 84)	Elevation (m a.s.l.)
WE 1	E 40/6.44	78.0	44.0	600	7.871504E	50.960112N	423
WE 2	E 40/6.44	78.0	44.0	600	7.873792E	50.959684N	424

Table 13: Summary of results from the 3 best neural networks out of 20000 iteratively tested networks for the target turbines in the complex terrain of the *Sauerland* region.

Target variable	Input variables	Network type	Network topology	RMSE (kWh)	r_S
$W_{\text{WE 1}}$	$W_{(\text{OL 1–OL 3})}$	MLP	3-9-1	640	0.90
$W_{\text{WE 1}}$	$W_{(\text{OL 1–OL 3})}$	RBF	3-22-1	602	0.91
$W_{\text{WE 1}}$	$W_{(\text{OL 1–OL 3})}$	RBF	3-22-1	588	0.92
$W_{\text{WE 2}}$	$W_{(\text{OL 1–OL 3})}$	MLP	3-7-1	747	0.92
$W_{\text{WE 2}}$	$W_{(\text{OL 1–OL 3})}$	RBF	3-24-1	630	0.93
$W_{\text{WE 2}}$	$W_{(\text{OL 1–OL 3})}$	RBF	3-21-1	668	0.90

was not observed in the corresponding validation dataset (Figure 8(a)).

To account for the more complex topographic conditions, we tested the addition of more input variables that potentially carry the information needed to better map the input values onto the energy yields of the target turbines. While choosing additional input variables, one must avoid that redundant variables are chosen as an increased number of variables will affect the ability to find the global minimum of the error function during the network training process [13, 37]. This effect is also known as the "curse of dimensionality" and affects all multivariate optimization algorithms [38].

Sensitivity analyses are one method to ensure that only variables are used as input that are important to map the input vectors on the target values. Sensitivity analyses determine the influence of each variable on the minimum of RMSE by, first, determining the minimal error, while taking all input variables into account. In the next steps, the values of the input variables will be partially replaced by random values, while the other input variables remain unchanged. This is done consecutively through all variables. The minimum error E_r applying the random valuesis then compared to the original minimum error (E_0) achieved with the original values of the respective variable. This is done by simply calculating the ratio $R_E = E_r/E_0$. Thus, an error ratio R_E equal to 1 or less indicates that the variable does not add additional information. In that case, the variable can be considered disruptive or at least redundant and should not be used as input variable.

In Table 14, the error ratios for the network that was optimized for the reproduction of the energy output values of the turbine WE 1 are shown using the operating data recorded by the three turbines of the *Olpe* wind farm.

The ranks reflect the influence of each input variable on the result. The higher a specific variable is ranked, the more it contributes to the error minimization. Only variables that exhibit an error ratio $R_E > 1.1$ were used for the neural networks during further analyses [32]. The error ratios confirm that the energy output values have a large impact on the quality of the results. It is also shown that additional wind meteorological variables and operating data such as the instantaneous power contain important information to model the target variable. The fact that the wind direction exhibits an error ratio of 2.233 also underlines the importance of the topographic effects on the wind energy yields in complex terrain. Since the dimensionality of the input vector was increased, more neurons in the hidden layers were needed for the two RBF networks used [12].

Through application of the additional important variables the performance of the neural network could be increased (Table 15). The correlation between the energy output calculated with the neural network utilizing the additional input variables and the measured energy output could be increased from a maximum of 0.92 (Table 12) to 0.97 (Table 15). Furthermore, The RMSE accumulated over 165 validation data records was reduced by 20.4% from 599 on 477 kWh. Using an RBF network ensemble did not increase the correlation but slightly reduced the RMSE by another 1.9% (Table 15).

Table 14: Ranked error ratios for available input variables of the wind farm *Olpe*.

Rank	Variable	R_E
1	$W_{(\mathrm{OL}\,2)}$	2.790
2	$W_{(\mathrm{OL}\,1)}$	2.258
3	Average wind direction (OL 1)	2.233
4	Average instantaneous power (OL 1)	1.889
5	$W_{(\mathrm{OL}\,3)}$	1.808
6	Average instantaneous power (OL 2)	1.550
7	Minimum instantaneous power (OL 1)	1.507
8	Daily hours of operation (OL 2)	1.296
9	Minimum instantaneous power (OL 2)	1.278
10	Maximum instantaneous power (OL 2)	1.274
11	Maximum wind speed (OL 1)	1.216
12	Average of wind speed (OL 3)	1.208
13	Average of wind speed (OL 2)	1.204
14	Minimum wind speed (OL 3)	1.088
15	Minimum number of revolutions (OL 3)	1.082

Only the 15 variables with $R_E > 1$ are shown.

4. Conclusion

We presented a neural network approach to model the daily energy yields of wind turbines by training neural networks using the data of other wind farms. The method was deployed on three examples with different spatial setups and distances between the input sites and the target sites in exemplary regions covering a variety of topographic complexity. The results show that artificial neural networks provide a capable mathematical tool to deliver reliable results. The data modeled by the trained neural networks are highly correlated to the corresponding data measured by the operating and surveillance system of the turbines with coefficients of 0.9 and higher.

Differences between the predictions of the best networks are small for the coastal sites as well as for the sites in the mostly flat region of the Westphalian Basin. Both network types tested allow a sound and accurate filling of the data gaps. However, the RBF networks turned out to better capture extreme values compared to the respective best MLP networks.

The biggest advantage of the method of artificial neural networks to fill data gaps is the fact that no information about the relationships between the variables or the statistical distributions of the individual quantities must be assumed prior to the network training.

In combination with the fast learning procedure for RBF networks, this makes them a suitable approach for wind energy yield predictions in practice.

The operating data of wind turbines are recorded today by default with high temporal resolution and many important technical parameters as well as basic wind meteorological measurements.

Our results show that these additional measurements can significantly increase the performance of the neural networks

Table 15: Summary of results for the modeled energy output of turbine WE 1 using 13 input variables from the site *Olpe*.

Destination variable	Input variables	Network type and topology	RMSE (kWh)	Correlation (r_S)
$W_{\text{WE 1}}$	$W_{(\text{OL 1})}$ average instantaneous power (OL 1) Minimum instantaneous power (OL 1) Average wind direction (OL 1) Maximum wind speed (OL 1)	Ensemble 2 × RBF (15-92-1) (15-101-1)	477	0.97
	$W_{(\text{OL 2})}$ Daily hours of operation (OL 2) Minimum instantaneous power (OL 2) Average instantaneous power (OL 2) Maximum instantaneous power (OL 2) Average wind speed (OL 2)	RBF (15-95-1)	494	0.97
	$W_{(\text{OL 3})}$ Average wind speed (OL 3)	RBF (15-93-1)	486	0.97

especially in areas with complex topography. Therefore, the approach presented helps to reduce the uncertainty of future wind energy predictions conducted with wind flow models. The work presented delivers the methodical framework to be deployed anywhere on the globe where input and related output data is available to train the artificial neural networks.

Due to the increasing demand for energy the application of specifically trained networks for turbines in complex terrain is of great interest in practice. Distances between wind turbines used as input and the target locations can be significantly larger than the distances considered in our case studies. Modeled power output values could be used as input for a different ANN that is applied on turbines even further away from the turbines that deliver the primary input values. Exploring the capability of the ANN approach to work on larger spatial scales by using consecutive stages of input and output datasets is a goal for future research in order to utilize the inherent flexibility of artificial neural networks and increase the planning dependability for wind energy projects.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of the paper.

Acknowledgments

The operating data of the wind turbines was kindly provided by ENERCON GmbH, SOLVENT GmbH, and SL-wind energy GmbH. The authors thank these companies for their support. The authors also would like to thank Joshua Baur for editing the language of the manuscript.

References

[1] IPCC, "Climate Change 2013: The Physical Science Basis," in *Contribution of Working Group I To the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker, D. Qin, G. K. Plattner et al., Eds., p. 1535, Cambridge University Press, Cambridge, United Kingdom, 2013.

[2] P. Friedlingstein, R. A. Houghton, G. Marland et al., "Update on CO_2 emissions," *Nature Geoscience*, vol. 3, no. 12, pp. 811–812, 2010.

[3] G. P. Peters, G. Marland, C. le Quéré, T. Boden, J. G. Canadell, and M. R. Raupach, "Rapid growth in CO_2 emissions after the 2008-2009 global financial crisis," *Nature Climate Change*, vol. 2, no. 1, pp. 2–4, 2012.

[4] G. P. Peters, C. L. Weber, D. Guan, and K. Hubacek, "China's growing CO2 emissions - A race between increasing consumption and efficiency gains," *Environmental Science and Technology*, vol. 41, no. 17, pp. 5939–5944, 2007.

[5] X. Lu, M. B. McElroy, and J. Kiviluoma, "Global potential for wind-generated electricity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 27, pp. 10933–10938, 2009.

[6] EEG, "Renewable energy sources Act of 25 October 2008," last amended by the Act of 20 December 2012, Art. 5G I, 2730, 2012.

[7] I. Troen and E. L. Petersen, *European Wind Atlas*, Risø National Laboratory, Roskilde, Denmark, 1989.

[8] N. G. Mortensen, L. Landberg, I. Troen, and E. L. Petersen, "Wind Atlas Analysis and Application Program (WAsP), Vol. 1: Getting Started. Vol. 2: User's Guide," Risø National Laboratory, Roskilde, Denmark, 1993.

[9] R. Gasdch and J. Twele, *Wind Power Plants: Fundamentals, Design, Construction and Operation*, Springer, Berlin, Germany, 2nd edition, 2012.

[10] W. Winkler, M. Strack, and A. Westerhellenweg, "Scaling and evaluation of wind data and wind farm energy yields," *DEWI Magazine*, vol. 23, pp. 76–84, 2003.

[11] FGW, "Technical Guidelines for Wind Turbines. Part 6. Determination of wind potential and energy yields," Revision 8, Berlin, Germany, 2011.

[12] C. M. Bishop, *Neural Networks For Pattern Recognition*, Oxford University Press, Oxford, UK, 1996.

[13] D. W. Patterson, *Artificial Neural Networks. Theory and Applications*, Prentice Hall, Singapore, 1996.

[14] D. S. Broomhead and D. Lowe, "Multi variable functional interpolation and adaptive networks," *Complex Systems*, vol. 2, pp. 321–355, 1988.

[15] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge, UK, 1996.

[16] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 1999.

[17] S. Li, D. C. Wunsch, E. A. O'Hair, and M. G. Giesselmann, "Using neural networks to estimate wind turbine power generation," *IEEE Transactions on Energy Conversion*, vol. 16, no. 3, pp. 276–282, 2001.

[18] Z. Liu, W. Gao, Y. H. Wan, and E. Muljadi, "Wind power plant prediction by using neural networks," in *Proceedings of the IEEE Energy Conversion Conference and Exposition*, Raleigh, NC, USA, September 2012.

[19] W. A. Oost and E. M. Oost, "An alternative approach to the parameterization the momentum flux over the sea," *Boundary-Layer Meteorology*, vol. 113, no. 3, pp. 411–426, 2004.

[20] S. Ayoubi and K. L. Sahrawat, "Comparing multivariate regression and artificial neural network to predict barley production from soil characteristics in Northern Iran," *Archives of Agronomy and Soil Science*, vol. 57, no. 5, pp. 549–565, 2011.

[21] H. Z. Hamid Zare Abyaneh, "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters," *Journal of Environ Health Science and Engineering*, vol. 12, article 40, 2014.

[22] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, D. E. Rumelhart, J. L. McClelland, and PDP Research Group, Eds., vol. 1, pp. 318–362, Cambridge University Press, Cambridge, Mass, USA, 1986.

[23] E. Metcalfe, J. Teteh, and S. L. Howells, "Optimisation of radial basis and backpropagation neural networks for modelling auto-ignition temperature by quantitative-structure property relationships," in *Chemometrics and Intelligent Laboratory Systems*, vol. 32, pp. 177–191, 1996.

[24] M. Hestenes, *Conjugate Direction Methods in Optimization*, Springer, New York, NY, USA, 1980.

[25] G. J. Bowden, H. R. Maier, and G. C. Dandy, "Optimal division of data for neural network models in water resources applications," *Water Resources Research*, vol. 38, no. 2, pp. 2–11, 2002.

[26] R. Khosla, "Knowledge-Based Intelligent Information and Engineering Systems," in *Proceedings of the KES 9th International Conference*, Lecture Notes in Computer Science, p. 933, Melbourne, Australia, September 2005.

[27] A. Schmidt, C. Hanson, J. Kathilankal, and B. E. Law, "Classification and assessment of turbulent fluxes above ecosystems in North-America with self-organizing feature map networks," *Agricultural and Forest Meteorology*, vol. 151, no. 4, pp. 508–520, 2011.

[28] B. Smith and H. Link, "Anemometer measurements for use in turbine power performance tests," Preprint American Wind Energy Association (AWEA) Windpower Conference Portland, Oregon, USA, June 2002.

[29] A. Albers and H. Sacman, "Wind speed and turbulence evaluation for performance and gondola anemometer measurements," *DEWI Magazine*, vol. 10, pp. 51–62, 1997.

[30] J. A. Anderson, *An Introduction To Neural Networks*, MIT Press, Boston, Mass, USA, 1995.

[31] M. Qu, F. Y. Shih, J. Jing, and H. Wang, "Automatic solar flare detection using MLP, RBF, and SVM," *Solar Physics*, vol. 217, no. 1, pp. 157–172, 2003.

[32] A. Schmidt, T. Wrzesinsky, and O. Klemm, "Gap filling and quality assessment of CO_2 and water vapour fluxes above an urban area with radial basis function neural networks," *Boundary-Layer Meteorology*, vol. 126, no. 3, pp. 389–413, 2008.

[33] Y. N. Maharani, S. Lee, and Y. K. Lee, "Topographical effects on wind speeds over various terrains: a case study for Korean peninsula," in *Proceedings of the 7th Asia-Pacific Conference on Wind Engineering*, Taipei, Taiwan, November 2009.

[34] M. P. Perrone and L. N. Cooper, "When networks disagree: ensemble methods for hybrid neural networks," in *Artificial Neural Networks For Speech and Vision*, R. J. Mammone, Ed., pp. 126–147, Chapman & Hall/CRC, London, UK.

[35] B. Bakker and T. Heskes, "Clustering ensembles of neural network models," *Neural Networks*, vol. 16, no. 2, pp. 261–269, 2003.

[36] K. J. Eidsvik, "A system for wind power estimation in mountainous terrain. prediction of askervein hill data," *Wind Energy*, vol. 8, no. 2, pp. 237–249, 2005.

[37] J. M. Zurada, A. Malinowski, and I. Cloete, "Sensitivity analysis for minimization of input data dimension for feedforward neural networks," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 447–450, IEEE Press, London, UK, February 1994.

[38] R. E. Bellman, *Adaptive Control Processes*, Princeton University Press, Princeton, NJ, USA, 1961.