LONG TERM CORRELATION OF WIND MEASUREMENTS USING NEURAL NETWORKS

A NEW METHOD FOR POST-PROCESSING SHORT-TIME MEASUREMENT DATA

Dipl. Inf. Carsten Albrecht, Maike Klesitz
AL-PRO GmbH & Co. KG , Dorfstr. 100, D-26532 Großheide, Germany
Email: albrecht@al-pro.de, Tel.: +49 (0) 4936 6986-0, Fax.: +49 (0) 4936 6986-46
www.al-pro.de

Abstract

In the last years during the growth of wind energy especially in Europe, it became more and more clear that measurements taken out on-site for a quite short period, usually around one year, can differ significantly from the long term wind offer at the given site. Energy content of wind, for instance, easily can differ in the range of 30 % from year to year in the Middle-European climate zone. For this, even a high precise wind measurement is worth nothing without a reliable method for long term correlation. The usual approach for carrying out long term correlation of wind measurements are the so-called MCP (Measure Correlate Predict) methods, which typically base on linear regression methods.

This paper suggests an alternative method using the neural network technique to perform long term correlations. Taking advantage of the abilities of neural networks, long term correlation can be improved taking into account non linear relationship between the measurement data and the long term data source as well as the development of the wind conditions in the past few observations. Performing this new method, long term correlation and therefore micrositing and determination of wind conditions that can be expected on-site can be improved significantly.

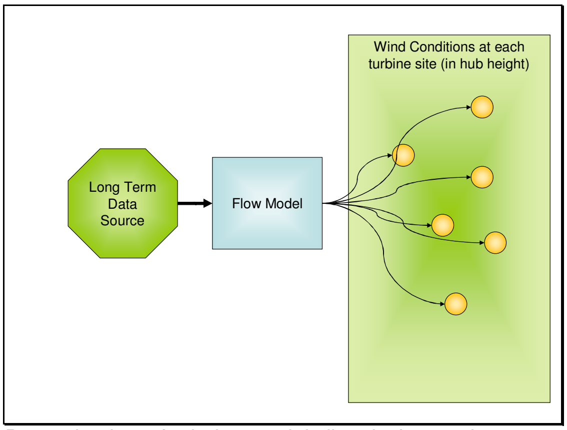
1 Introduction

For the determination of wind potentials – in particular in the planning phase – frequently one refers to site near wind measurements. Usually such measurements are only initiated in the course of the development of the project. Accordingly the available measuring period is generally short and typically moves within the range of one year.

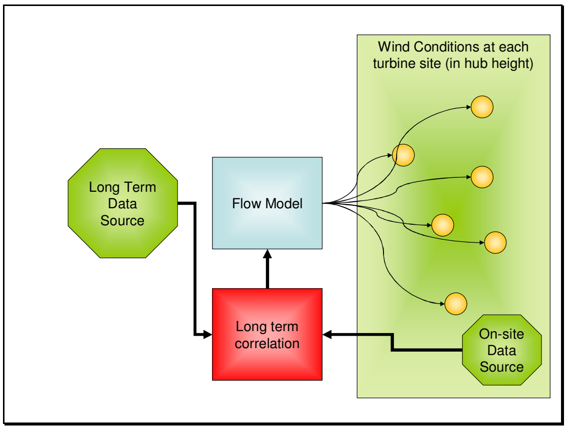
Then the received measurement results are converted with the help of a suitable flow model to the locations and hub heights of the planned wind energy plants. In order to get a statement –usually desired - about the wind energy arising within many years which can be expected, it is to be guaranteed before that the used data is long-term representative. Generally a long-term correlation is to be accomplished.

Alternatively also a wind data source raised of many years can be converted directly to the WEC locations with the help of a suitable flow model. The advantage of the first named method is that the less the used input data refers to the regarded location the more the accuracy of flow models usually decreases. However directly site-referred wind data raised of many years with sufficient accuracy and quality is available only rarely.

On the other hand the methods used for the long-term correlation are often afflicted with substantial inaccuracies. In order to come to a reliable estimate of the wind potential when using a short time location-referred measurement, a high-quality long-term correlation is necessary as well as a meaningful determination of the uncertainties of the used methods. This is essential in order not to lose the advantage of the accuracy of the flow model reached by the site-referred wind measurement.



Determination of wind potential directly from a long term data source.



Determination of wind potential with on site measurement.

2 Introduction of the methodology

2.1 General methodology requirements for long term correlation

Long term correlation of wind measurement data always – as a basic requirement – needs information about the long term wind potential in comparison with the period that has been measured on-site. Typically, this information is derived from measurement or modelled wind data available for the long term period of interest. In this case, each method for long term correlation consists of the following two steps:

1. Determination of a map between the long term data and the measured data on base of the concurrent period available for both data sets.
2. Application of the found map on the whole period of the long term data source.

For instance, a common method is based on a linear regression. In this case the map between long term data l and on-site data s is given as:

$$s = m \cdot l + b$$

Typically, the parameters m and b are determined for each wind direction separately. Both data sets must be available as time series.

This method assumes a pure linear relationship between both datasets. Indeed this is very often the case for measurements located not too far away from each other and in a comparable terrain situation.

For more complicated situations the relation typically contains non-linear components and can not be described in such a simple way.

More sophisticated methods are needed to realize a reliable long term correlation in those cases.

2.2 Basic Idea

Basic need of determining the map between the data sets can be reduced in recognizing certain meteorological situations and their meaning for the on-site wind conditions. In the described case this is done simply by the wind data of a long term station close to the site.

Unfortunately, very often such station does not exist. In this case the relation between long term and on-site data might be much more complicated. The idea of the introduced method is to use neural networks for determination of the map and performing long term correlation. The hope is to take advantage especially from the following features of neural networks:

1. Ability to recognize non-linear relationships.

Neural networks are well known to be able to recognize non linear relations between data sets. As especially those relations are expected for the present task it also can be expected that this ability is helpful for improving long term correlation.

2. Ability to take past data into account.

Certain neural network architectures have the opportunity to take past observations into account. This might be helpful for two reasons:

First, if long term and on site measurement are located in a larger distance to each other, events that happen at the one site, depending on the wind direction, need some time to proceed to the other location. For this, observations taken at the same time would not show a correlation which indeed is present.

Second is that to recognize a certain meteorological situation it might not be sufficient to take just one observation into account, the development over some past observations at the long term could tell a lot more about what to expect on site.

3. Ability to use more than one long term data source.

Even if this is not a special feature of neural networks, it can be expected that neural networks, as relations become even more complicated with the increase of information, are especially able to handle more data sources. The developed method has been created for network architectures that are able to handle up to four long term data sources.

Certainly, with the increase of information the chance to find a good, reliable and stable map between the long term data sources and the on-site measurement – if the method is able to divide significant information from those not needed in the particular case, which again might depend on the given weather situation. So, for certain conditions it might be necessary to take all available data sources into account while in other cases only one source will do. Neural networks should be able to take a decision here.

Also a combination of all these abilities is expected to be helpful for the present task, so for a decision which station to chose if more than one long term source is available could be easier taking past information of each long term station into account, and so on.

2.3 Analyzing available long term data sources

2.3.1 Synoptic measurements with regional relation to the site

Mainly it can be said that this type of data is good for long term correlation. As for all data sources the better the correlation between on-site and long term data the better is the achieved long term correlation. This is typically the case for measurements located close to the site and/or in a similar terrain situation.

The problem, especially for remote areas is that such a data source very often is not available, measurement periods are to short or the database is inconsistent due to changes in measurement instruments, recording or the environment of the station.

2.3.2 Model data

This means data, especially time series, derivingof any meteorological model, typically given for each grid point of such a model.

Depending on the resolution of the model, the big advantage of such data is that it is available quite close to the site.

It is a matter of course that the quality of the data depends directly on the ability of the used model to compute the real meteorological conditions – and also on the input data the model has been driven with.

2.3.3 Reanalysis data

This data is a special case of model data taken from world wide reanalysis models. Two common datasets are available, the European ERA 40 and the North American NCAR/NCEP data.

The data is available back to 1948 in case of the NCAR/NCEP dataset. The grid resolution is quite coarse with 2.5 degrees. For the geographical height of Germany this means a grid point about each 200 km.

The big advantage of these datasets is their world-wide availability. So, in principle, long term correlation can be carried out at each site world wide using it. Although in each particular case consistency and data quality has to be checked carefully.

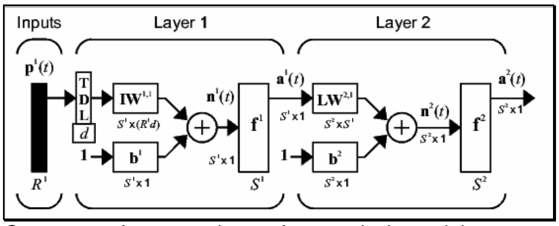
The present method has been designed with special regard to this data set, pre-processed as World Wind Atlas [1]. With the ability to include up to four long term datasets, for instance the four surrounding grid points of the site could be choosen.

2.4 Neural networks for long term correlation

It is a broad literature base available on theory and practical applications of neural networks, therefore only a very brief introduction into this theme now.

Neural networks try to copy the biologic functionality of real neuron systems in an electronic way. The basic idea is to connect several neurons that perform a very simple function like the identity or simple linear terms with weighted connections. Then, the network is trained by adjusting these weights. Often also an additional bias parameter for each neuron is introduced and trained.

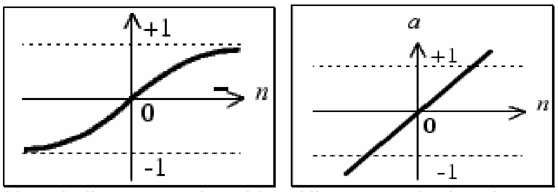
The network architecture used in the present application is a so-called focussed time delay architecture.



Structure of a two layer focussed time delay neural network.

The main characteristic of this kind of neural network is the so-called tapped delay (TDL) before the input layer. This tapped delay enables the network to include past data sets into the computation of the actual output. Other architectures which also include recursive structures, typically taking the network output as a part of the input vector have not been tested yet.

For the transfer function of the neurons, depending on the task, a broad number of functions are thinkable and suggested in literature. For the present study two typical cases, the hyperbolic tangent sigmoid and a simple linear transfer function have been tested. A common approach is to have linear neurons in the last layer while the transfer function in the other layers may differ. This approach has been followed in this study.



Hyperbolic tangent sigmoid and linear transfer function

A main characteristic of neural networks is that once they have been set up they have to be trained before they are at all able to produce useful output. After setup, the weights and biases are initialized by default values or by random. Then a training data set consisting of inputs and corresponding outputs is needed for training.

Training then takes place with various training algorithms adjusting the parameters of the net. This is usually an iterative procedure, typically several 100 training cycles are needed to adjust the network's parameters to the given application.

One of the main problems in training neural networks is to avoid so-called over fitting with means in simple words that the net learns to reproduce exactly the input data – but nothing else. It is easy to understand that any neural net, similar to other mathematic functions, can easily reproduce any output if only there are enough free parameters available. It is a main task to avoid this over fitting. For the present application two different approaches for training algorithms and techniques to avoid over-fitting have been tested so far. It is planned to continue testing on more training algorithms in the future.

The first algorithm that has been tested is a back propagation algorithm that combines adaptive learning rate with momentum training and is called traingdx. For this algorithm the early stopping technique has been chosen to avoid over fitting. By this method the training data is split in two parts. While training takes place on the one subset only, the second – usually smaller - one is taken to evaluate the performance of the net.

The second tested algorithm is called Bayesian training, with regularisation technique that bases on modifying the performance function by adding a term that consists of the mean of the sum of squares of the network weights and biases.

For a detailed description of the training algorithms and a bibliography on neural network theory refer to the Matlab help [5].

2.5 Implementation

The developed method has been implemented using the software package Matlab [5]. The developed software can train neural networks with up to 3 internal layers and an additional output layer. For each internal layer, the type of transfer function and the number of neurons can be chosen. The training function for the designed net can be switched between the two described training methods; in the case of the traingdx additionally the fraction of data that is used for training can be specified. Also it is possible to give higher wind speeds a higher weight. Also it is possible to let the application calculate the standard deviation of the residuals occurring when the trained net reproduces the observation with the long term data. Therefore, an important requirement on the actual technical guidelines is fulfilled [3]. Furthermore, it is possible to re-apply this “noise rate” to the whole computed long term time series. This might improve reproducing the real energy content of the measurement data, a similar technique is used in the MCP Tool of the software WindPRO [6].

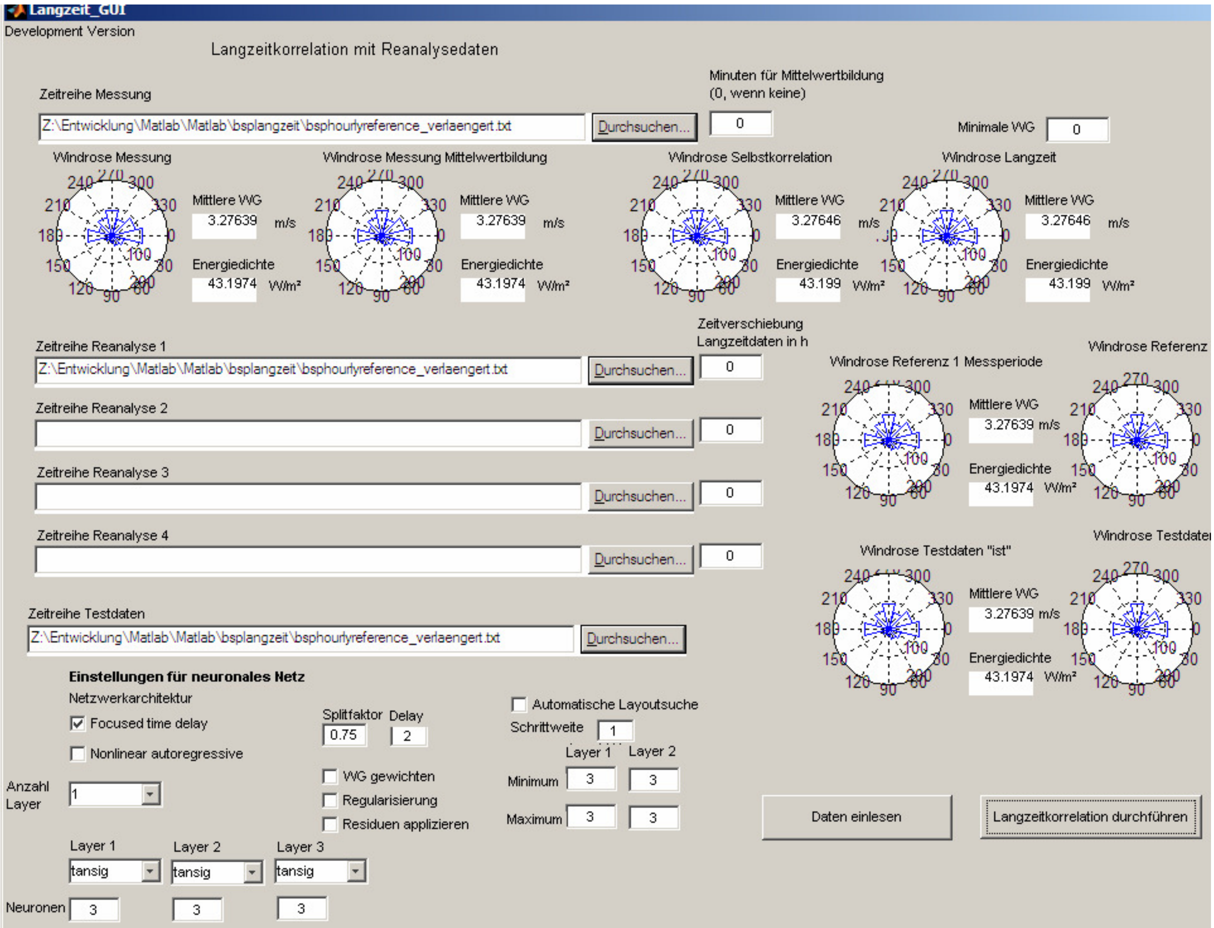
Furthermore, the delay steps can be set, in other meaning the number of past data sets of the long term data source that should be taken into account.

The software is able to long term correlate a time series with up to 4 long term data sources. Averaging of the data is possible as well as having different time zones in the measurement and the long term data. Additionally it is possible to include a test data set deriving from the on site measurement but not part of the training data to check the quality of the correlation on a data subset that has not been used for training the neural network.

The only format accepted for the time series actually is a semicolon separated ASCII file as generated by the WindPRO meteo objects [6] when writing time series into a file. The tool is also able to interpret the filtered data column produced by WindPRO when filtering the time series.

Two versions of the application have been developed. The interactive version allows setting the network architecture parameters manually. The batch version is controlled by an Excel file giving up to 100 network architectures and settings. Each of the network architecture is tested, the main results of the correlation are written into the same file.

The second version of the software was used for the testing described in the next section.



Screenshot of the application, interactive version.

3 Testing and validation

3.1 Test concept

During the testing, it has been focussed especially on the following aspects:

1. Comparison of the two training algorithms used.
2. Improvement achieved by using delays (past data sets).
3. Improvement achieved by applying residual distribution.
4. Comparison with standard MCP algorithms.

A set of 18 network architectures has been tested in each case on two sites so far. Each site has been long term correlated using several long term data sources as described below.

3.2 Test cases

Two sites have been tested up to now, each with several long term data sources. The proceeding to prepare the data was the following in each case:

1. Each time series (long term and site data) has been checked, errors and implausible data has been filtered out.
2. Each long term data source has been fit to the available on-site data. So when later a long term correlation was performed using this data the result quality can be checked against the real observation.
3. A subset of the available on-site data, usually the last part, was taken as measurement data and training database for the neural networks.
4. Another subset was taken as test data to determine the result quality. Therefore, result quality has been tested on a data set which was not a part of the training data for the neural networks

3.2.1 Sicily

The site ST6 is located in the central part of the island. Measurement data is available for the period between October 2003 and December 2004. The period May 2004 to December 2004 has been used as measurement data, the period from October 2003 to April 2004 as test data.

The following long term data sources have been tested:

3.2.1.1 ST3:

A measurement in the close neighbourhood of ST6. This test case is regarded as an easy task that should also simple regression methods can handle.

3.2.1.2 World Wind Atlas:

The two nearest points located at 12,5°°East, 37,5° North and 15°°East, 37,5° North have been used as data source, each of them separate and also together. As the grid for the NCAR/NCEP reanalysis model is too coarse to even resolve the Italian peninsula one would expect that this long term data source would not correlate well with the measurement data at all. However, especially the time series from the western data point shows good correlation with the measurement. Therefore, difficulty of this test case is estimated as medium.

3.2.2 France

The site is located in the Vendée in the western part of France. Measurement data is available for the period between September 2003 and June 2005. The period September 2004 to June 2005 has been used as measurement data, the period from September 2003 to August 2004 as test data.

The following long term data sources have been tested:

3.2.2.1 Méteo France:

A station provided by the French weather service, about 30 km away from the site. Difficulty is estimated as medium to high.

3.2.2.2 **World wind atlas:**

The point closest to the site, located at 0° West, 45° North has been tested as well as the 3 surrounding grid points located at 0° West, 45° North, 2,5° West, 45° North and 0° West, 47,5° North together. Difficulty is estimated as high.

3.3 Network architectures

The test performed series initially consisted of 20 network architectures. During the testing it was found out that the last 2 architectures have been too complicated to train on the used PC's in many cases, so they have been skipped. 18 architectures remained for further evaluation.

The neurons of the internal layers always have been set up as tansig neurons. The tansig function is expected to be able to handle difficult non-linear problems better than a simple purelin function. On the other hand, the purelin might be the better choice if the relation between the data is close to a simple linear regression.

	# Neurons		
# Layers	Layer 1	Layer 2	Layer 3
1	3	0	0
2	3	3	0
3	3	3	3
1	6	0	0
1	9	0	0
1	12	0	0
1	21	0	0
1	30	0	0
2	3	6	0
2	3	9	0
3	3	6	6
3	3	9	9
3	6	6	6
3	12	12	12
2	12	12	0
2	12	21	0
2	12	30	0
2	30	30	0

Table 1, Tested network architectures

3.4 Results

To determine the result quality, for chapter 3.4.1 to 3.4.3 the absolute deviation in energy content on the test data set has been taken as reference parameter. For chapter

3.4.5 the production of an E-40 turbine has been taken as reference.

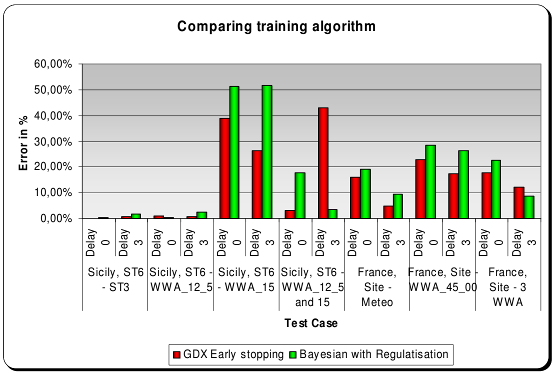
3.4.1 Comparing training algorithms

A first result of all tests was that for each application sometimes totally different network architectures lead to good results. There was only a main trend to observe that with complexity of the problem also the complexity of the neural network architectures that delivered good solutions increased. No rule in finding out about what architecture is good for a certain problem type could be recognized so far.

Comparing results of the two training algorithms a first observation was that using the GDX early stopping gave a broader bandwidth of result quality from totally off to very good. In nearly all cases the best result that could be achieved beats the Bayesian regularisation method. On the other hand, the Bayesian regularisation delivered more regular results.

One observation during the testing was that obviously the parameter for the number of iterations to continue after the stop criteria is reached in the early stopping was set to a too high value. This has been changed but not tested yet.

The diagram below compares the results of the best architectures for each case



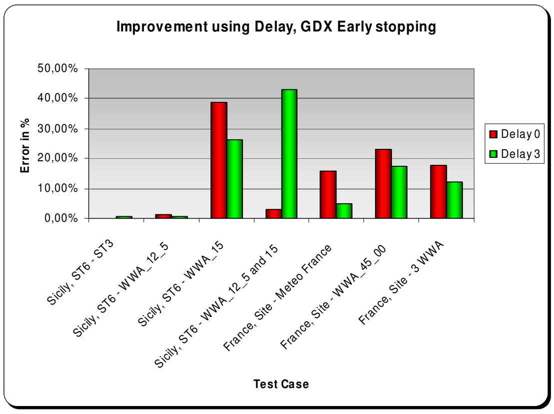
It can be seen clearly that in the Sicilian case obviously using ST6 as long term data leads to extremely good results – as well as the first World Wind Atlas grid point. The second one, located 15° east leads to bad results and obviously is not a good choice for long term correlation at all.

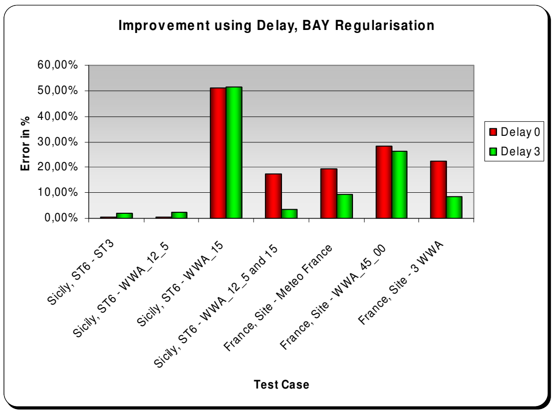
3.4.2 Testing delay improvement

For this case again the results of the best networks of each run are compared. A test run without any delay is compared with a run including the 3 past data sets so far. Typically, best results have been achieved by different network architectures. This was expected as including more information like past data should lead to more complex neural network architectures. Again, a common rule which architecture might do best could not be recognized before.

It was a clear picture that including delays leads to better results. Only one case that might be caused by an error during the test run showed different results.

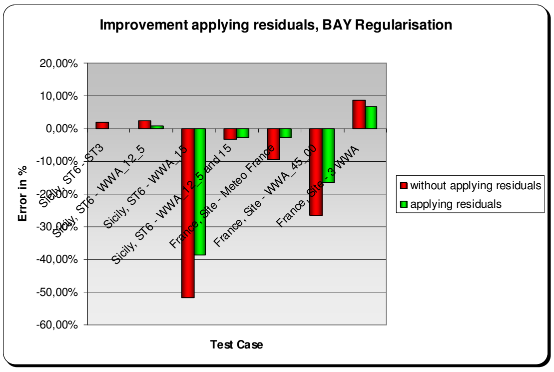
The diagram below compares the result quality for each test case and both training algorithms.





3.4.3 Applying residuals

This has only been tested with the Bayesian regularisation training so far. Again the best results have been compared. In this case the deviation (not the absolute deviation) has been evaluated. The results showed an improvement when applying residuals. This differs from later observation that real turbine production is over-estimated in this case.



3.4.4 Developing a method to determine useful network architectures

With the observation that network architectures "doing a good job" are very different from case to case it becomes an important question how to recognize a network well usable for long term correlation when only the concurrent data and the results that the software can deliver are available.

It has been found that minimizing the mean error on the test data and the self correlated concurrent test data delivered good results while, as expected, not the absolutely best network could be identified by this method in any case.

A disadvantage of this method is the need of an independent test data set. This means in practical use that the available data has to be split in two, or in case of the GDX early stopping even in three subsets. This means that the number of data sets available for training the neural network decreases.

Developing a better method of determination is in progress actually.

3.4.5 Comparison with common MCP methods

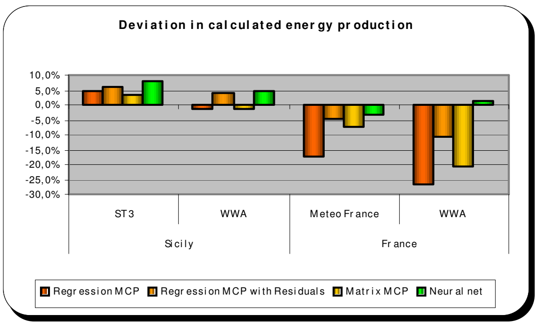
This test is carried out following a highly practical approach. The production of an Enercon E-40 6.44 turbine has been compared, using the real data, the data of two common MCP approaches, regression MCP and Matrix MCP as implemented in the WindPRO [6] MCP tool.

The computation of the turbine production was done by using the WindPRO meteo module using a Weibull fit.

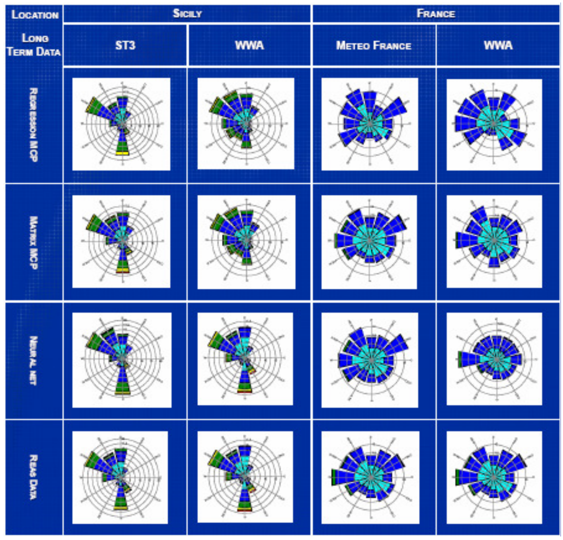
To categorize the test cases it was once assumed that long term data from a measurement close to the site was available. The second assumption was that only the World Wind Atlas data could be used.

For the neural networks, the one found with the rule developed in 3.4.4 was used. This ensures that in practice is is not only the case that there is a network existing that can do the job but one is also able to identify it.

The results are shown in the diagram below.



It can be seen, that fort he Sicilian case no significant change in the result quality occurs – all methods deliver good results. The French case shows clear improvements by using the neural network, especially when using the World Wind Atlas data.



Comparison of wind roses computed

4 Summary, conclusion

A new application to long term correlate wind measurements by using neural networks has been developed and tested on two sites with different long term data sources so far.

The results achieved are promising. Especially in complex cases where the relation between long term data and measurement is difficult neural nets gave – partly significantly - better results than common MCP methods. It also could been shown clearly, that applying past observations into the calculation improves the results – as well as including more than one long term data source.

On the other hand the testing also showed a lot of open questions to answer. Development of improved criteria to find out good network architectures for the given set of measurement data and long term data is one of the main requirements. Some improvements in the implementation already took place; their effect has to be tested now as well.

Finally a broader test data base with more cases to check is desired to ensure these conclusions.

References

[1] World Wind Atlas (WWA): digital wind atlas, Sander+Partner GmbH/Switzerland; www.sander-partner.ch

[2] National Center for Atmospheric Research (NCAR), Boulder, Colorado, USA; www.ncep.ucar.edu/ncar. National Centers for Environmental Prediction (NCEP), Camp Springs, Maryland, USA; www.ncep.noaa.gov

[3] Standard for the preparation of wind energy assessments, Revision 2, Advisory board for wind energy assessment in the German Wind Energy Association BWE.

[4] Technical Guidelines for Wind Energy Converters, Part 6, Determination of wind potential and energy yield, Revision 5, Fördergesellschaft Windenergie, www.wind-fgw.de

[5] MATLAB 7.2.0.232 (R2006a), The MathWorks, Inc. www.mathworks.com

[6] WindPRO 2.5 SP3, EMD International A/S, www.emd.dk