



## Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection

Schlechtingen, Meik; Santos, Ilmar

*Published in:*  
Mechanical Systems and Signal Processing

*Link to article, DOI:*  
[10.1016/j.ymssp.2010.12.007](https://doi.org/10.1016/j.ymssp.2010.12.007)

*Publication date:*  
2011

*Document Version*  
Early version, also known as pre-print

[Link back to DTU Orbit](#)

*Citation (APA):*  
Schlechtingen, M., & Santos, I. (2011). Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mechanical Systems and Signal Processing*, 25(5), 1849-1875. <https://doi.org/10.1016/j.ymssp.2010.12.007>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection

Meik Schlechtingen and Ilmar Ferreira Santos

Department of Mechanical Engineering, Section of Solid Mechanics  
Technical University of Denmark  
Phone: +45 45256269  
Fax: +45 45931475  
Email: ifs@mek.dtu.dk

18<sup>th</sup> November 2010

---

## Abstract

This paper presents the research results of a comparison of three different model based approaches for wind turbine fault detection in online SCADA-data, by applying developed models to five real measured faults and anomalies. The regression based model as the simplest approach to build a normal behavior model is compared to two artificial neural network based approaches. They are a full signal reconstruction and an autoregressive normal behavior model. Based on a real time series containing two generator bearing damages the capabilities of identifying the incipient fault prior to the actual failure are investigated. The period after the first bearing damage is used to develop the three normal behavior models. The developed or trained models are used to investigate how the second damage manifests in the prediction error. Furthermore the full signal reconstruction and the autoregressive approach are applied to further real time series containing gearbox bearing damages and stator temperature anomalies.

The comparison revealed all three models being capable of detecting incipient faults. However, they differ in the effort required for model development and the remaining operational time after first indication of damage. The general nonlinear neural network approaches outperform the regression model. The remaining seasonality in the regression model prediction error makes it difficult to detect abnormality and leads to increased alarm levels and thus a shorter remaining operational period. For the bearing damages and the stator anomalies under investigation the full signal reconstruction neural network gave the best fault visibility and thus lead to the highest confidence level.

*Keywords: Condition monitoring, Neural networks, SCADA-Data, Fault detection, Health condition, Wind turbine, Online monitoring, Normal behavior models, Generator bearing damage, Gearbox bearing damage*

---

## 1. Introduction

Condition monitoring of wind turbine components is of increasing importance. The size of wind turbines used nowadays has reached a level where the availability of the turbine is very crucial. Downtimes are very costly. It is therefore worth increasing the effort spent to monitor the turbine condition in order to reduce unscheduled downtime and thus costs.

Condition monitoring (CM) systems can be used to aid plant owners in achieving these goals. They aim to provide operators with information regarding the health of their machines, which in turn, can help them improve operational efficiency by allowing more informed decisions regarding maintenance [1].

The available CM systems mostly require high level knowledge about the problem domain. However, this knowledge is difficult to access and does often not exist. Physical models can thus seldom be built.

On the other hand there is a large amount of historical operational data available, which can be used to give an indication about the turbine condition. By application of advanced signal analysis methods,

focused on trends of representative signals or combination of signals, significant changes in turbine behavior can be detected at an early stage [2].

Another possibility of identifying changes in signal behavior are model based approaches. Thereby the historical operational data is used to develop models capable of predicting a certain output signal, when given one or more input signals. For wind turbine signals these approaches are well suited, since many signals can be found to be correlated to other signals simultaneously measured, e.g. the wind speed or the power output.

One advantage of using normal behavior models to monitor wind turbine signals lies in the reduction of prior knowledge about the signal behavior. Another important property is that with normal behavior models the possibility of monitoring the signal is widely decoupled from the operational mode. In practice simpler monitoring approaches such as by defining thresholds, are difficult to establish due to the various operational modes, which cause signals to widely fluctuate. If thresholds are to be defined they

must be specified for several operational modes individually.

The normal behavior models are developed at a stage, where the turbine components can be considered healthy. Afterwards, the model is used to estimate a specific signal. The estimation error can give an indication of signal behavior changes and thus incipient faults.

There is a number of research activities carried out in this field. They mainly deal with artificial intelligence approaches. Among the highest developed is a system called SIMAP [3] and a multi agent system (MAS)[1], [4] currently under development. Both systems make use of artificial intelligence techniques to set up the normal behavior models of online SCADA signals.

In this article the initial findings of an ongoing three year Ph.D. research project are presented. The project aim is to develop a condition monitoring system that primarily uses ordinary SCADA data usually available to wind turbine operators. In the first project stage algorithms are to be developed with which all wind turbine SCADA data can be monitored and anomalies detected. This will be done with normal behavior models in a similar manner as it is presented here.

At the present project state normal behavior models to monitor the following signals are developed:

- Power output
- Generator bearing temperature
- Generator stator temperature
- Generator slip ring temperature
- Shaft speed
- Gearbox oil sump temperature
- Gearbox bearing temperature
- Nacelle temperature

Normal behavior models to monitor further SCADA data are currently under development.

In the second project stage anomaly patterns will be related to specific faults occurred, laying the basis for statements about the component and finally the turbine condition. This is likely to be done with neural networks (NNs) for pattern recognition.

The research is based on SCADA data from ten different operating offshore turbines of the same type in the two MW class at present.

The intention of this article is to compare two artificial intelligence approaches to a regression based approach in combination with data smoothing techniques. The analysis is based on real measured faults.

In the literature a comparative analysis of regression and artificial neural network models is described in [5] for power curve estimation. The comparison showed a high accuracy and a good performance of the NN approach. The estimation error achieved was in the range of 2% and the conclusion

drawn was that NNs can be used to estimate wind power generation efficiently as a diagnostic tool[5].

The main original contribution of this paper is the direct comparison of the fault visibility of real measured wind turbine faults in the prediction error of normal behavior models. The CM systems described in literature mainly use autoregressive approaches to predict the desired signal. This approach is most suitable to slow changing signals such as temperatures. Some wind turbine signals have a low autocorrelation, e.g. the turbine power output and thus are difficult to model with this kind of model. It is thus important to see the limits of each of the model type approaches in order to identify the one appropriate for a given application or a needed combination.

In section 2 the time series used for research are introduced and the damages or anomalies highlighted. Also the sensor positions are illustrated. In section 3 there is a brief description of how the regression based model is set up and how smoothing is applied. In the following section 4 it is briefly described how the NN is set up and how it is trained to achieve a good performance. In section 5 and 6 the trained networks are applied to time series described in section 2 and a comparative fault investigation for damages and anomalies is shown.

## 2. Time series containing anomalies

Before introducing the time series relevant for this paper, the relevant sensors are shown in a schematic of the wind turbine generators (WTG) under investigation.

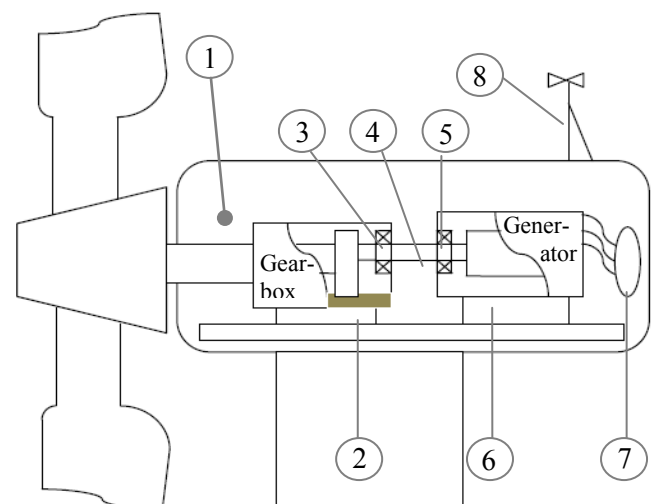
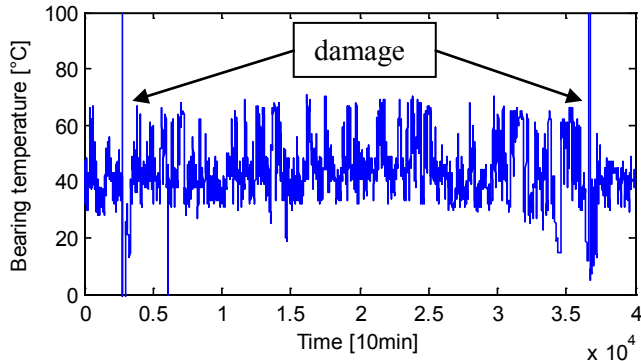


Figure 1: Schematic of sensors in a wind turbine [6]

- 1: Nacelle temperature
- 2: Gearbox oil sump temperature
- 3: Gearbox bearing temperature
- 4: High speed shaft revolution speed
- 5: Generator bearing temperature
- 6: Stator temperature (3 phases)
- 7: Power output
- 8: Ambient temperature

## 2.1. Generator bearing anomaly

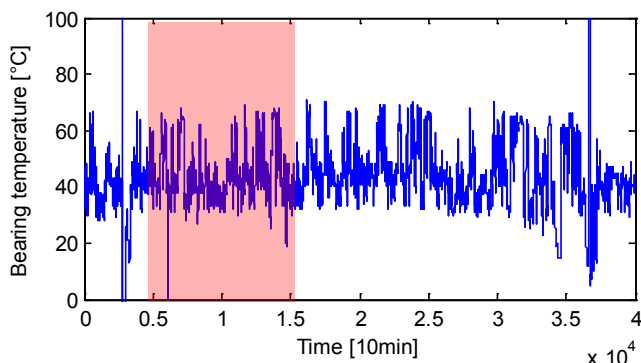
Next to the overall performance of NNs and regression based approaches, it is equally important to investigate the fault visibility in the estimation error. Basis of this analysis is a time series containing two catastrophic generator bearing damages of a 2 MW offshore wind turbine.



**Figure 2: Time series containing two generator bearing damages measured on WTG A**

The time series shows the bearing temperature evolution over the operational time. Both catastrophic bearing damages required a bearing replacement. For this reason, the time series is particularly well suited for the performed analysis, since new bearings have a high likelihood to show normal behavior. The specific reason for the two bearing damages to occur so close to each other is, however, unknown. For the period under investigation no other fault is reported for this generator.

For training of the NNs and the regression model development the period after the first bearing damage is used.



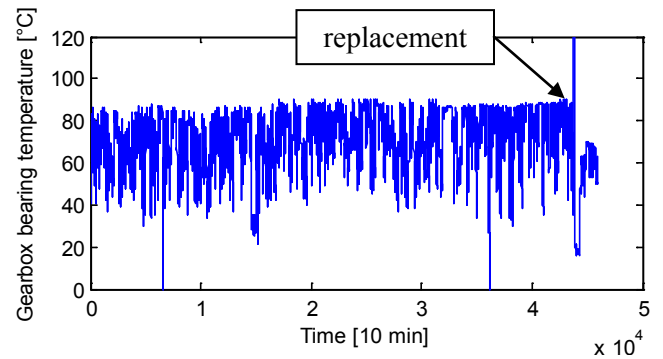
**Figure 3: Period used for training/ regression model development WTG A**

The trained networks and the regression model are then applied to the full time series. Hence the fault manifestation of the second bearing damage in the estimation error can be investigated and the three approaches compared. In this paper all basic analysis is performed on this time series.

The following time series are used to achieve a more general view on the limitation of the full signal reconstruction (FSRC) and autoregressive approach.

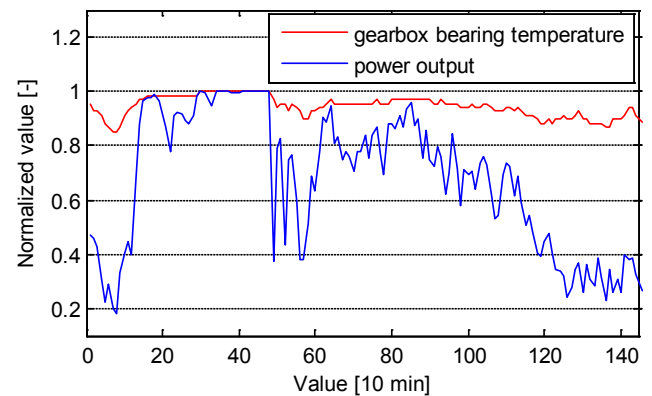
## 2.2. Gearbox bearing temperature anomaly I

The time series shown in Figure 4 illustrates the gearbox high speed shaft bearing temperature measured at another turbine.



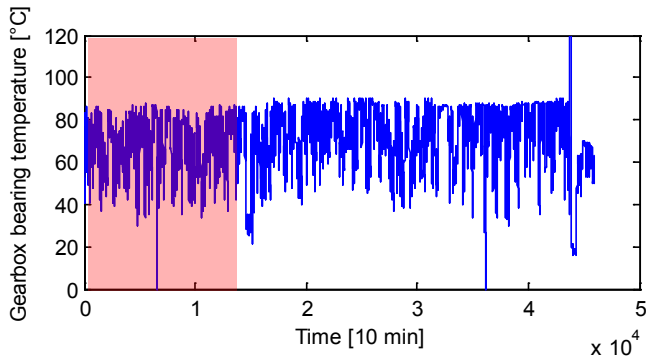
**Figure 4: Time series containing a gearbox bearing damage measured on WTG B**

Although the time series shows a gradual temperature increase over time, identification of temperature increase based on thresholds is only possible under full load condition and is thus dependent on the operational mode. The large power output fluctuations cause problems when defining thresholds for each individual operational mode (e.g. 50-100 kW or 200-250 kW power output). The difference in autocorrelation between for instance the bearing temperature and the power output leads to the same temperature level being present in several power output bins. Figure 5 shows an example of the power fluctuations over time and the related gearbox bearing temperature.



**Figure 5: Exemplary power and gearbox bearing temperature fluctuations over time**

For the gearbox under investigation, there is no information about an earlier bearing replacement. For this reason the first three months of the time series available are used for model development as shown in Figure 6.

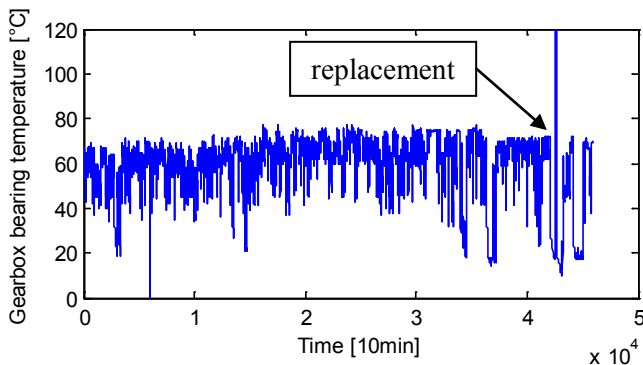


**Figure 6: Period used for training WTG B**

The damage was recognized by the operator 2.5 months before the bearing was finally replaced. When the incipient fault was recognized the turbine power output was restricted to 60% of the nominal power to allow a scheduled bearing replacement. There is no further information about the type of damage or on the exact cause available. This holds for all damages and anomalies presented in this paper.

### 2.3. Gearbox bearing temperature anomaly II

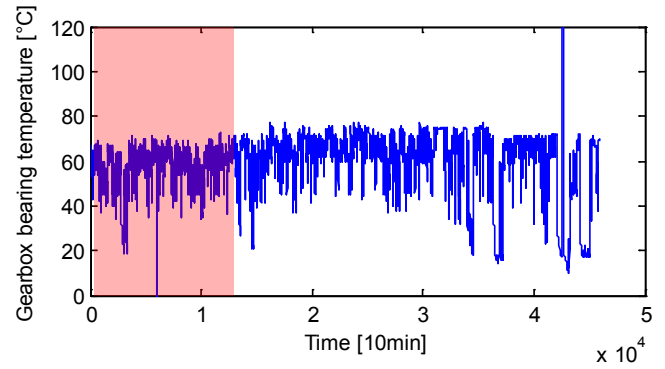
In the data supplied by the operator also a further gearbox high speed bearing damage is measured on a different turbine. The time series is shown in Figure 7.



**Figure 7: Time series containing a gearbox bearing damage measured on WTG C**

The time series shows a trend in maximum temperature measured right at the beginning of the time series. After 20000 values the maximum temperature measured stabilizes. 7000 values prior the bearing damage the maximum power output is restricted to 50% of the nominal power which lead to a drop in maximum temperature level.

Although the time series shows a trend in the first five months of data available, these data are used to train the networks. The present trend will have an influence on the network performance, but training the network after the first 20000 values would not leave enough data for fault analysis. Therefore the influence on the performance is accepted, due to lag of earlier data being accessible. The training period is highlighted in Figure 8.

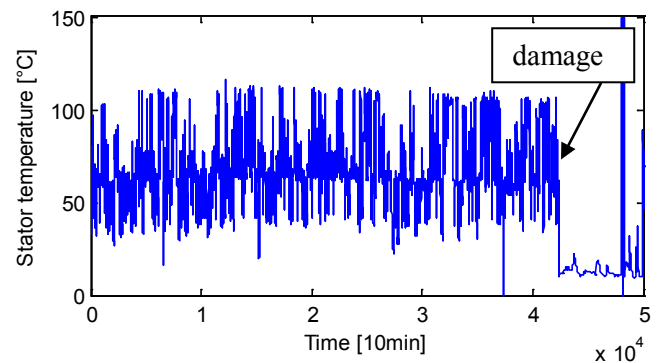


**Figure 8: Period used for training WTG C**

All data currently available for this research cover a period of 14 months recorded after two to three years of turbine operation. It is therefore possible that training is performed with data indicating signs of wear. In this case only relative changes from this status are possible to identify.

### 2.4. Generator stator temperature anomaly I

In order to further generalize the comparative analysis also other types of faults are investigated in this research. Figure 9 shows the stator temperature evolution up to the point where a generator exchange was required due to generator reconnection problems. This is when the turbine is not able to automatically connect to the grid, if the turbine & the grid is o.k. and the wind speed is above the cut in wind speed.



**Figure 9: Time series of the stator temperature measured on WTG D**

The time series contains no obvious anomaly. Only a small seasonality effect is present as a result of the annual change in ambient temperature. The first three months of data available are used for training.

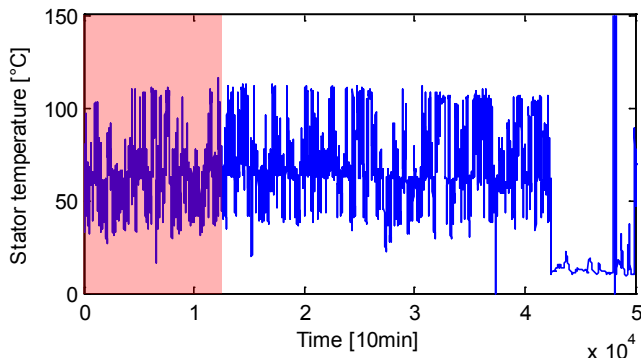


Figure 10: Period used for training WTG D

Apart from the information about the wind turbine status code that highlighted the issue no information about the cause is available.

### 2.5. Generator stator temperature anomaly II

The second stator temperature anomaly is a consequence of service performed on WTG E. The time series showing the anomaly is illustrated in Figure 11.

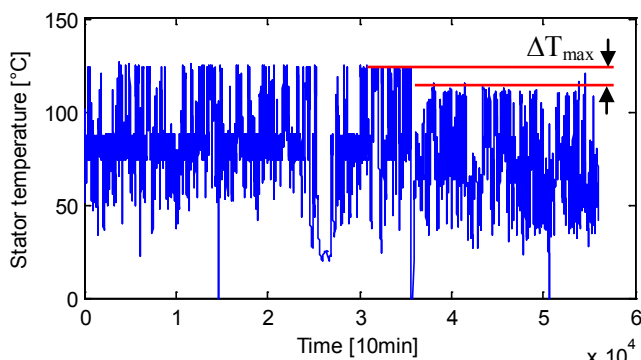


Figure 11: Time series of the stator temperature measured on WTG E

The service action had an influence on the stator temperature. In the consequence the stator temperature dropped by 10 °C. Possible causes can be cleaning of the water cooling system during service or insufficient volume flux due to low system pressure in the expansion tank. If the pressure is increased in the expansion tank during service the cooling system has a larger volume flux and in the consequence is more efficient.

The time series shows how the service action improved the situation by lowering the temperature level. If it is possible to detect this anomaly with a normal behavior model the opposite - degradation of the cooling system efficiency over time - is expected to also be detectable. However, this can finally not be proved by the data supplied.

The period for used for training is shown in Figure 12.

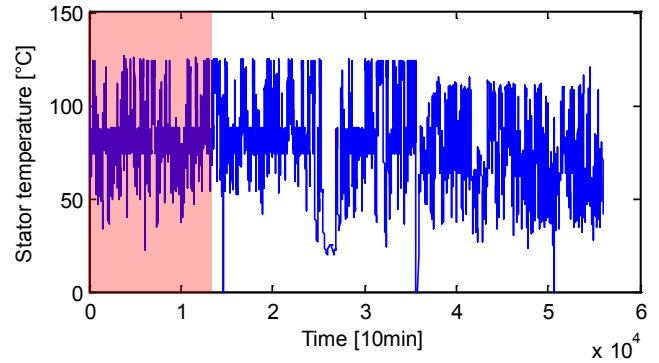


Figure 12: Period used for training WTG E

### 3. Regression model development

Many wind turbine signals can be found to be closely correlated to other simultaneously measured signals. By using the linear cross-correlation the related signals and their lag to the desired signal can be found. The actual lag removal is an important step to reduce the scatter and improve the prediction.

The identified correlated signals can then be used to build the regression model.

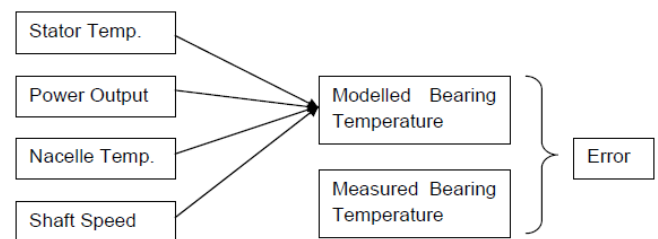


Figure 13: Regression model schematic [6]

Especially the stator temperature was found to be highly correlated to the generator bearing temperature as it is visible in Figure 14.

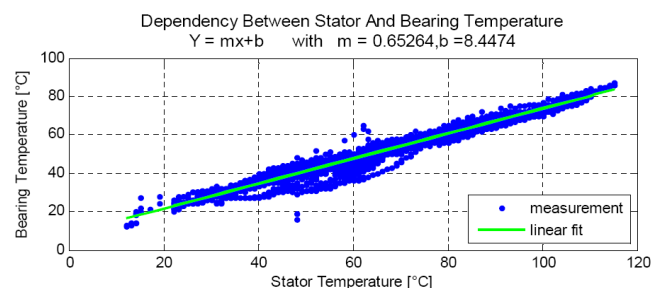
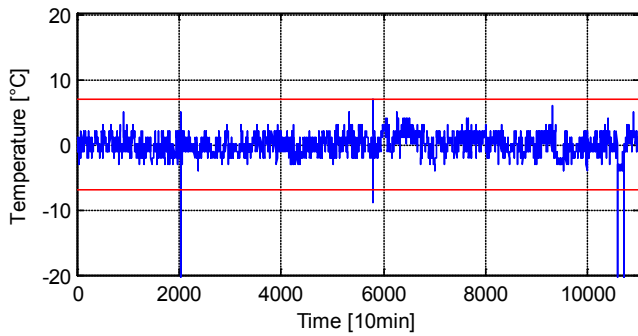


Figure 14: Correlation between the stator and the bearing temperature [6]

The regression of both signals for the specific generator type is well presented by a first order polynomial. In the general case, it can however be also a higher order polynomial function. Due to the high correlation between the stator and the bearing temperature the linear model is already accurate. The three remaining signals, i.e. the power output, nacelle temperature and the shaft speed are used to reduce the number of prediction outliers.

The prediction error for the full regression model is illustrated in Figure 15:



**Figure 15: Estimation error for the full regression model**

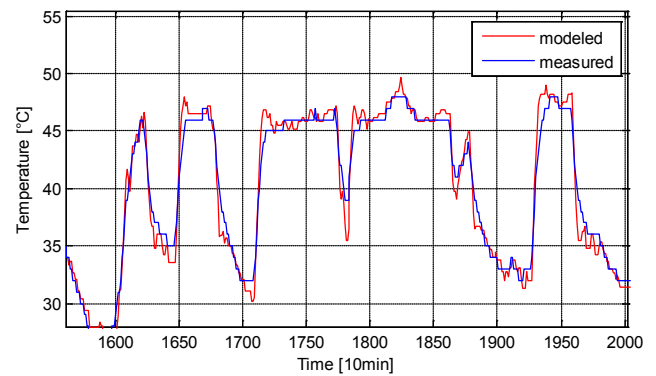
Apart from the few outliers, the regression model is capable of predicting the bearing temperature in the range of  $\pm 5$  °C accurate. The threshold value is five times the standard deviation. The remaining outliers are caused by extreme transient situations such as turbine run up or shut down. The reason is an insufficient description of the cooling or heat transfer process and the difference in heat capacity. The stator temperature in these situations is no longer a good representation of the bearing temperature. In case of cooling, the model should be decoupled and natural cooling should be simulated for instance by Newton's exponential cooling law. This law states that the temperature of a cooling body falls exponentially towards the temperature of its surroundings with a rate which is proportional to the area of contact between the body and the environment [7].

The high ratio of stator heat capacity over power loss smoothes out any high frequency fluctuations. In this sense normal operational temperature profiles are modeled well.

The problem can be either overcome by implementation of separate models for transient situations or by simple filtering. Filtering of non-operational periods is possible, since determination of the component condition based on online SCADA data is most reasonable if the turbine is operating; e.g. determining a bearing temperature if the shaft is not rotating is not expected to give information about the bearing condition. The same is with the stator temperature or the power output.

When anomalies, such as offsets are detected during operation the non-operational data can, however, be used to identify purely sensor related issues.

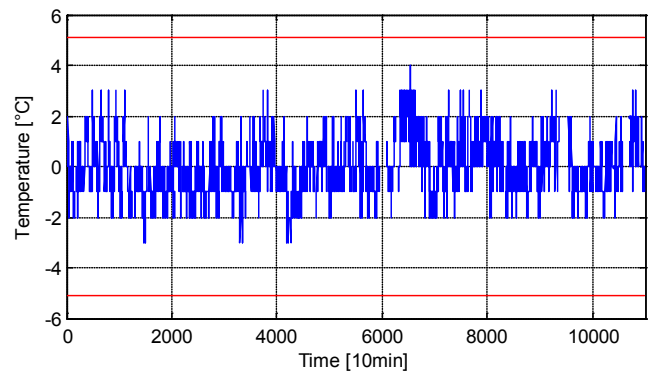
The difference in heat capacity and mass of the matter surrounding the temperature sensors also causes another phenomenon to occur. Since the stator temperature sensor is placed right at the stator coils (heat source) the signals shows larger fluctuations as the bearing temperature signal. Hence the predicted bearing temperature signal also shows higher order fluctuations. This effect is emphasized in Figure 16.



**Figure 16: Estimated and measured bearing temperature [6]**

By use of a smoothing filter the higher order variations can be filtered and the prediction error reduced.

When transient situations are filtered and the smoothing filter is applied the prediction error in the considered time period becomes:



**Figure 17: Bearing temperature estimation error**

The prediction error is finally down to  $\pm 4$  °C. The final model is now kept fixed and can be applied to the full time series in order to see how the bearing damage manifests in the prediction error.

#### 4. Neural network model setup and training

The regression based model is rather simple to handle and it is a trivial task to estimate how the model will behave in case of a signal change, since the underlying function is known. The characteristic of the model deviation can be interpreted in a way that the signal deviating from the normal behavior can be identified.

The multilayer perceptron (MLP), on the other hand, is a complex model with a nested form of nonlinearity designed to preserve the information content of the training data. However, it uses a black box approach to globally fit a single function, into the data, thereby losing insight into the problem [8].

There a number of different factors important to successfully apply NNs. A large variety of literature is available in the field of NNs. The ones referred to in the context of this research are: [9],[10],[11] and[12].

Generally one of the major concerns is to identify the suitable network set up for the given application. Important issues are:

- Identification of input signals
- Network type and transfer/activation function
- Input data pre-processing and validity check
- Number of training patterns
- Network structure (number of layers, number of neurons, number of neurons in each layer, network connectivity)
- Training method
- Weight initialization

The issues will be addressed separately in brief and the relevant aspects for this research are mentioned.

#### 4.1.1. Input signals

The input signals used are generally the same as for the regression model. There are basically two ways of predicting the generator bearing temperature signal. The first is to use the regression between the bearing temperature and other correlated signals in a way to independently reconstruct the signal. The model input signals are (see Figure 1):

- El. Power output (t-lag) {pos.7}
- Nacelle temperature (t-lag) {pos.1}
- Generator stator temperature (t-lag) {pos.6}
- Generator speed (t-lag) {pos.4}

The second is to identify changes in the auto-regression of the bearing temperature. The input signals are the same, but are extended with the previous bearing temperature.

- Generator bearing temperature (t-1) {pos.5}
- El. Power output (t-lag) {pos.7}
- Nacelle temperature (t-lag) {pos.1}
- Generator stator temperature (t-lag) {pos.6}
- Generator speed (t-lag) {pos.4}

The autoregressive approach is used by [3] and [1] to identify gearbox failures.

Both kinds of models are compared to the regression model.

#### 4.1.2. Network type and transfer/ activation function

In this research multi-layer perceptrons in combination with Sigmoid transfer function are applied. Since the weights are found iteratively the curve shape of the transfer function has a strong influence on the number of iterations and the stability of the process. Functions with high gradients may lead to oscillations if the weight modification is not sufficiently small enough.

#### 4.1.3. Input data pre-processing and validity check

Input data pre-processing is an essential part to successfully apply NNs. Although NNs are able to handle fuzzy or incomplete data they are sensitive to invalid data. Therefore some pre-processing must be carried out. This is particular important when training a network. The network might not give an optimal generalization otherwise. The principal pre-processes applied are:

- (a) Validity check
- (b) Data scaling
- (c) Missing data processing
- (d) Lag removal

**(a) Validity check** - For a smoother network generalization it may sometimes be necessary to remove some of the data outliers [10]. A simple method of a validity check is a data range check. For example, if oven temperature data has been collected in degrees centigrade values in the range 50 to 400 would be expected. A value of -10 or 900, is clearly wrong[12]. However, defining a validity range is not a trivial task. Bearing temperatures for instance can reach a very high level exceeding 200 °C in the event of serious failures. The normal operational range is up to 90 °C usually. The range would therefore have to be chosen extremely carefully to not classify real measured values as invalid.

Ideally attempts should be made to check that the following conditions are met when the data is assembled into the input vectors:

- all elements of the vectors are within the expected ranges;
- all elements of the vectors are mutually consistent;
- for a supervised learning application, the target vectors are consistent with the input vectors [12].

Consistency checks are difficult to put into effect as the overall target is to identify abnormal signal behavior. Consistency checks therefore have a very limited use in this research. It is useful in identifying faulty sensors. It is therefore important to monitor the amount of data classified as invalid as indication of possible problems with the data import, processing or the sensor itself. The data used for model development in this paper are checked for their ranges and for their consistency by filtering extreme outliers and data with unexpected high gradients. Furthermore constant rows are removed.

**(b) Data scaling** - In this paper data scaling is performed by applying the following equation:



$$S = \frac{V - V_{\min}}{V_{\max} - V_{\min}} \quad \text{Equation 1}$$

S... Normalized variable

V... Variable

**(c) Missing data processing** - Unknown or missing values are particularly harmful during training. If they are not classified as NaNs but 999 for instance the network might try to fit to these values. This will increase the generalization error of the network. Thus it is first of all important to clarify how missing values are labeled.

Since there is a large amount of data available, no approximation was performed. Instead missing input and target values are treated by neglecting the data set.

**(e) Lag removal** - Wind turbine signals usually do not respond immediately to changes of operational conditions. The delay is individual to each signal and must be removed when setting up a static normal behavior model. The lag can be identified by use of the cross-correlation of the signals. It was found that neglecting the signal lags led to an increase in prediction error (broader scatter). The regression based model shows that there is a strong linear relation between the input and the output variables. Therefore the cross-correlation function is used to identify the lag in the linear dependency. The input variables are then shifted to build a consistent input-output dataset.

#### 4.1.4. Number of training patterns

There are no acceptable generalized rules to determine the size of the training data for suitable training. The pattern chosen for training must cover upper and lower boundaries and a sufficient number of samples representing particular features over the entire training domain [10]. The number of required training patterns thus depends on the complexity of the problem. In [1], three months of operational data are used for the gearbox bearing temperature. For the examples presented in this paper this amount of data is sufficient, too. Generally the period of data required should be as short as possible, since the components are subjected to wear and degrade from the first hour of operation.

#### 4.1.5. Network structure

The network structure, i.e. the number of layers, the connectivity and the number of neurons can be generally chosen arbitrarily. Determining the architecture is an iterative process [1]. The process aim is to find a network structure that gives the best generalization.

For the purpose of this research a two layer network is used for the given application.

**(b) Number of Neurons** - In [12] it is recommended that the optimum number of neurons should be found by performing at least 10 runs where only the number of neurons is changed. The network architecture which gives the best generalization should be chosen. A similar method is also recommended by [10] for MLP networks. This method reduces the risk of finding solutions that do not generalize well. This method is therefore used in this research.

**(c) Network connectivity** - In [1] and [3] a feed forward network type is chosen. Due to its simplicity this network type is good to start with, unless the specific application requires a more sophisticated type. For the given application this type proved to be sufficient.

The inputs are connected to each neuron of the hidden layer, which in turn is connected to each neuron of the output layer.

The final network architecture for the autoregressive generator bearing temperature NN model is:

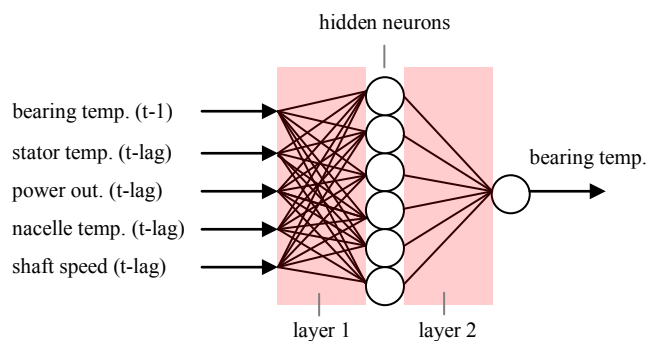


Figure 18: Network architecture - feed forward network [6]

The network architecture for the FSRC generator bearing temperature model is similar, but has only five neurons in the hidden layer.

#### 4.1.6. Training method

Generally there are two different modes of training NNs: batch mode and pattern mode (sequential learning) [12]. In [10] it is recommended using batch mode to begin with. If the behavior of the network to unseen data is not satisfactory pattern mode should be tried.

There are also two other training methods available, which are a modification of the simple batch mode and may reduce the risk of local minima. They are batch gradient decent and gradient decent with momentum. In the latter mode not only the local gradient of the error function is calculated, but also the general trend. Thus local minima may be survived and generalization improved. For this reason this training method is used for the research carried out in this paper.

### 4.1.7. Weight initialization

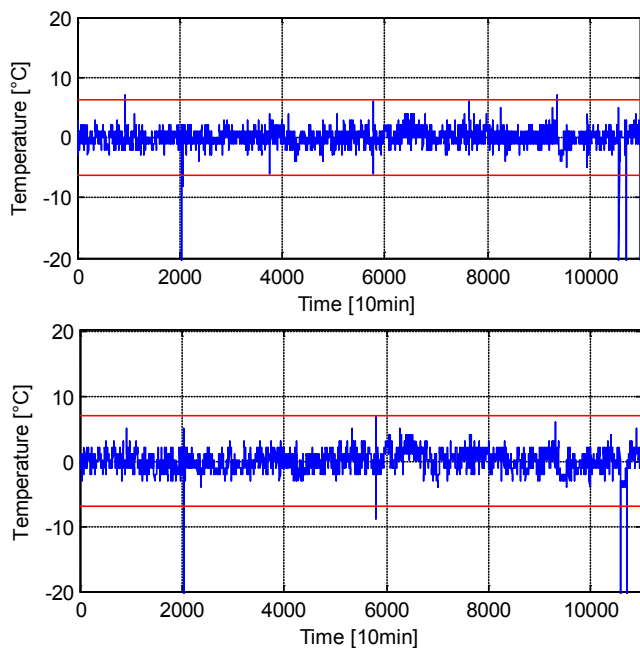
When training MLPs with the error back-propagation algorithm the first step for any training run is to initialize the weights in both layers by setting each one to a small random number, typically in the range  $-0.01$  to  $+0.01$  [12].

The initialization with small weights allows the identification of unused input signals. For those signals the weights will stay small close to the initial value. The inputs can be removed to increase the computational efficiency.

### 4.2. FSRC neural network model

The set up network is capable of predicting the generator bearing temperature, when given the stator and nacelle temperature as well as the shaft speed and the power output. It was found that in principle the stator temperature and the nacelle temperature are sufficient quantities to set up the model. However, using the power output and the shaft speed additionally reduces some of the outliers.

The network training is performed with the same input data as used for the regression model to produce comparable results. In Figure 19 the prediction error for the FSRC NN and the simple regression model are shown.



**Figure 19: Prediction error for the training data set; FSRC NN model (top); Regression model (bottom)**

Although the standard deviation is somewhat reduced, the two models give comparable results. Both models show large prediction errors for transient situations. Here the model complexity has a strong influence. When more input signals are given to the network and more hidden units (neurons) are used, transient situation can be found to be better

represented. Possible further inputs are for instance the statistics of the 10 min periods, i.e. the standard deviations. Also the generator slip ring temperature or the cooling fan activity may be used to increase the model accuracy. For the current research the fan activity was not supplied and could therefore not be accounted for. The standard deviations and the generator slip ring temperature were applied, but finally left out in order to simplify the models.

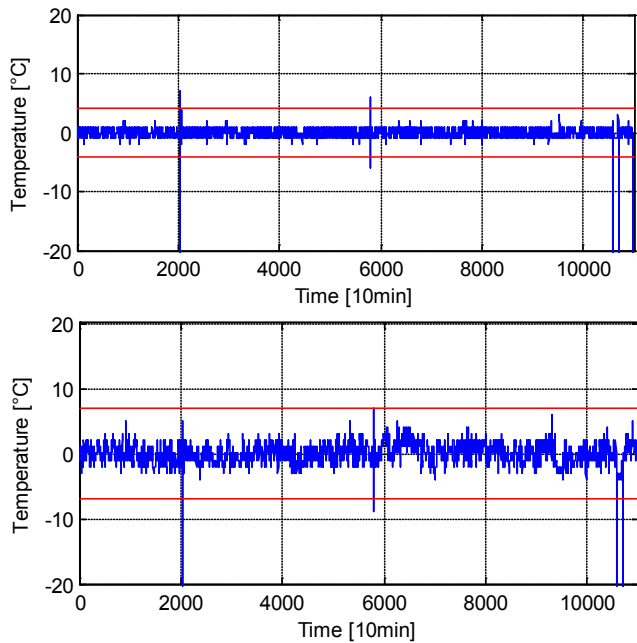
The more input signals are used the more computational time is needed. The overall target for the current research is not to develop models that are most accurate, but models that are first of all robust in order to reduce the number of false alarms and ease the anomaly interpretation process.

Fault identification does, however, become more difficult when using more input signals. With NNs, the identification of the normal behaving signal must be determined by fuzzy logic, rather than a direct interpretation of the deviation, as it is possible for the simple regression model. In SIMAP [3] a fuzzy expert system is implemented to identify the abnormalities. The more input signals are used, the more complex the interpreting logic must be. Thus the aim of the model implementation procedure is to find model inputs that give the best representation of the desired signal, but keep the model complexity low.

To overcome the problem of single predictions triggering an alarm a further threshold value is introduced by Caselitz and Giehardt [13]. If more than five subsequent values exceed the limit, an alarm is triggered [13]. For the developed models in this paper, it is sufficient using a threshold of three subsequent exceedings to prevent false alarms.

### 4.3. Autoregressive neural network model

Training of the autoregressive model is done in the same way as the FSRC model. Again several training runs are performed to find an optimal network architecture. The figure below shows the prediction error for the training data set and a comparison to the simple regression model.



**Figure 20: Prediction error for the training data set; Autoregressive NN model (top); Regression model (bottom)**

The autoregressive model gives very accurate results. The reason is the high autocorrelation of the bearing temperature signal caused by the large amount of mass surrounding the bearing.

The main heat source in the generator is the stator loss. The heat emitted is transmitted to the bearing through radiation, conduction and convection. Bearings on the other hand are very efficient and the amount of heat released by friction is order of magnitudes smaller. Thus the bearing temperature can be well expressed by its previous temperature plus a trend coming from the stator temperature sensor, illustrated in Figure 1 pos 6.

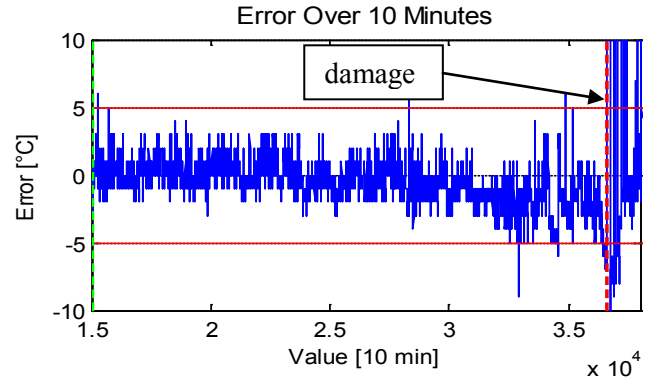
The prediction error is in the range of  $\pm 1$  °C most of the time. Transient situations are modeled more accurately, but the extreme transients are still not well represented. The optimization procedure used does optimize the mean squared error on the test data set. This data set should contain all situations to be modeled. Dominating situations like normal operation lead to a best fit in this operation mode. Transient situation inaccuracies could thus be reduced when using a data set that contains more of these operation modes. This will, however, influence the overall network generalization.

## 5. Fault identification and comparison

Once the regression model is set up and the NNs are trained the models can be used to predict the bearing temperature and identify the second fault present in the time series shown in Figure 2.

### 5.1. Regression model

The developed relationship between the signals in terms of their polynomial functions and coefficients is kept fixed as found during the model development phase. Figure 21 shows the evolution of the prediction error from the time series exceeding the one used for model development until the second bearing damage.

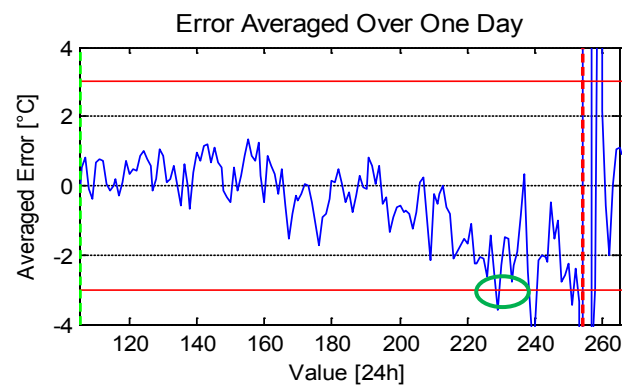


**Figure 21: Prediction error of the regression model**

The error indicates a growing shift in the error mean beginning 83 days prior to the total bearing damage. However, identifying this trend with high accuracy is not trivial. Here a threshold based approach proved practical. The choice of the threshold value is a tradeoff between the model sensitivity against anomalies and the number of false alarms.

Since the subsequent number of predictions exceeding the alarm limit is set to three, an alarm was first triggered 33 h prior the damage.

Figure 22 shows the one day average values of the prediction error for the period under consideration.



**Figure 22: Averaged prediction error of the regression model**

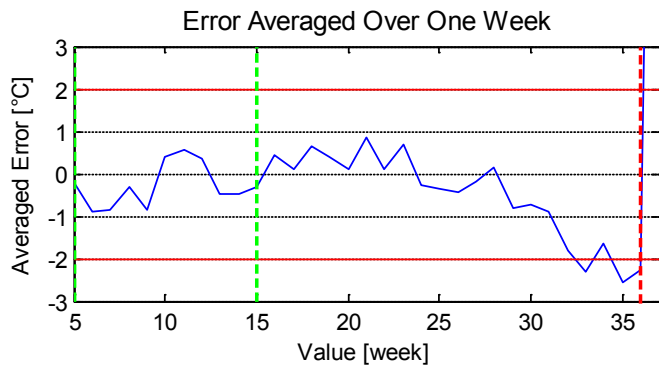
The averaged signal eases the trend identification. On the other hand it is visible, that the trend is first increasing before it finally decreases.

The signal amplitude constantly increases up to the point of failure. Also the development of a suitable analysis tool for growing amplitudes is aimed to be developed at a later stage of the research project.

The amplitude increase together with the shift in mean. This leads to a first alarm limit violation 25 days prior to the damage. By this time the confidence level of a damage being present is high, since looking at the

historical values from this point clearly shows the trend in the signal.

Seasonality is a major concern of the developed model and the method in general as it complicates the anomaly identification. The regression between the bearing, nacelle and ambient temperature is weak. Also the amount of data used for building the polynomial expressions covers only a fraction of the annual temperature circle. In order to implement the regression with higher accuracy a full cycle is required, or more advanced tools need to be applied. Figure 23 shows the one week average prediction error including the model development period.



**Figure 23: Averaged prediction error including the model development period**

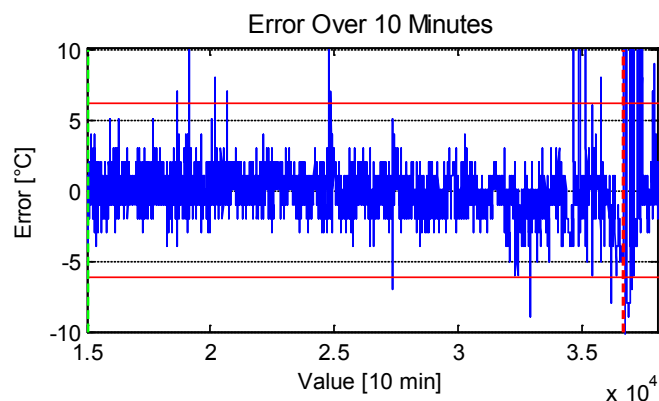
It is evident that up to six weeks prior to the damage, a clear trend cannot be identified due to the present seasonal effect.

## 5.2. FSRC neural network model

Once training is successfully completed the model predicts the bearing temperature when given previously unseen input signals.

The prediction strongly depends on the weight initialization. The random initialization leads to quite different model outputs during the optimization procedure. It was found, that not all of them are capable of predicting the damage. For the final model no alarm is triggered on the basis of 10 min averages.

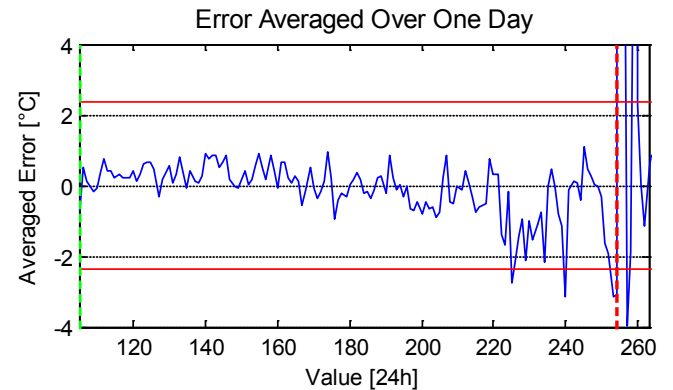
Figure 24 shows the prediction error development prior to the damage.



**Figure 24: Prediction error of the full signal reconstruction neural network model**

Visually an anomaly can be detected, beginning 4000 values ( $\approx 27$  days) prior to the damage.

Figure 25 shows the one day averages of the prediction error.



**Figure 25: Averaged prediction error of the full signal reconstruction neural network model**

In comparison to the regression model, the averaged error has reduced amplitude. The model is more accurate, leading to reduced alarm limits. An alarm is triggered 30 days before the bearing breaks. Also for this model an increase in amplitude is visible when the final break comes closer. The general nonlinear neural network is capable of better representing the ambient temperature effect. This causes the underlying seasonality to be less pronounced. Hence seasonality is not a major concern of this type of approach. Figure 26 shows the long term prediction error development averaged over one week.



**Figure 26: Averaged prediction error of the FSRC neural network model including the training period**

The seasonal effect contained in the signal is  $\pm 0.3$  °C in comparison to  $\pm 0.5$  °C in the regression model. The one week averages prove useful to detect long term signal trends. The current fault situation, however, does not contain a long term trend. An anomaly in form of a trend can be identified from six weeks prior the break. The evidence level of an anomaly is low up to the point an alarm is triggered by alarm limit violation (ALV).

### 5.3. Autoregressive neural network model

From network training it is known that the autoregressive model has a very high accuracy due to the large heat capacity. This model is thus capable of detecting very small changes in the autoregression of the temperature signal. The signal autoregression changes, when an internal heat sink or source is added. This could be for instance a bearing damage, where the bearing internal heat increases.

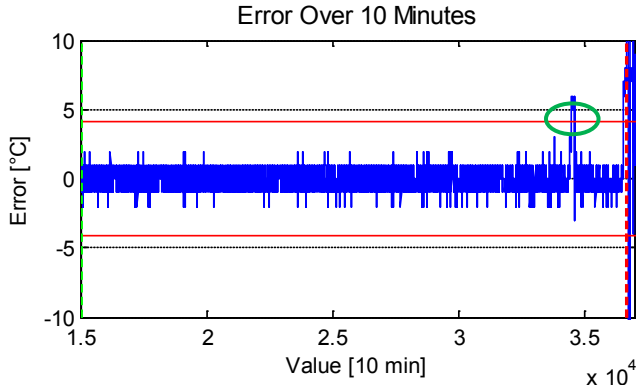


Figure 27: Prediction error of the autoregressive neural network model

On 10 min bases, the prediction accuracy is in the range of  $\pm 2$  °C for most of the time series. 14 days prior to the damage the prediction error exceeds 5 °C. The alarm limit is violated and an alarm is triggered. After this event, the prediction error falls back to its normal operational range of  $\pm 2$  °C.

Figure 28 shows the averaged prediction error over one day.

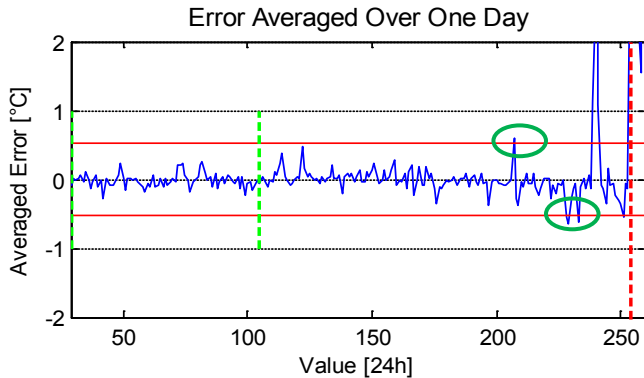


Figure 28: Averaged prediction error of the autoregressive neural network model

On an averaged basis there are more peaks exceeding the alarm limit. The earliest alarm is generated 50 days in advance. This is followed by further alarms 25 and 20 days prior the damage. By this time the evidence level of damage being present is high, when considering the fact that the prediction error amplitude is increasing. Whether the model deviates positively or negatively depends on the turbine operational mode and the final weights and biases.

The remaining seasonality in the prediction is shown in Figure 29.

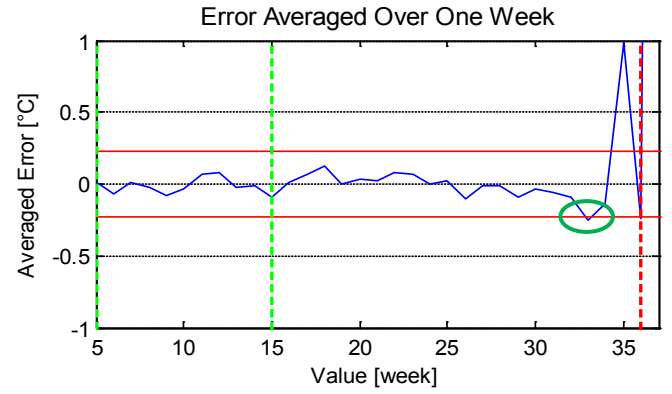


Figure 29: Averaged prediction error of the autoregressive neural network model including the training period

The seasonality is close to zero. In general the ambient temperature comes into effect, as the cooling system is more efficient for low ambient temperatures. This effect is less pronounced in this model, since the bearing temperature is determined by the previous value plus a trend given by the other input signals. The information about the cooling system efficiency and the ambient conditions are contained in the level of the previous bearing temperatures recorded.

When a bearing damage is present, the bearing internal heat source more strongly determines the bearing temperature. This heat source depends on the load and the rotational speed. For speed variable machines, the bearing temperature is thus expected to show higher fluctuations. The autoregression changes and the prediction error increases. This is the main difference to the other two models set up, which aim to map the relation between the ambient conditions and the target signal.

### 5.4. Comparison

All three models are able to identify the second bearing damage before the bearing finally fails. There are, however, differences in the remaining operation time before total failure. The table below summarizes the results for the different models based on first ALV and the point where the anomaly can be identified with higher confidence. This is when the alarm limit is violated the second time or a clear trend can be identified.

	First ALV	Second ALV/ Trend
Regression	25 days	18 days
Full signal reconst. NN	30 days	25 days
Autoregressive NN	50 days	25 days

Table 1: Comparison of the fault visibility between the three different normal behavior model type approaches

The results of the comparison are in principal expected to be dependent on the type of damage and operational modes prior to the damage.

For the considered bearing damage, the autoregressive NN model gives the earliest anomaly indication. The fault visibility increases for all models, when averaging the 10 min values over one day.

The regression model has the highest alarm limits and a large seasonality in the prediction. The fault can, however, still be identified. When setting up the NNs with the optimization procedure developed, seasonality effects are reduced and the alarm limits are lower, increasing the fault sensitivity.

During regression model development, it was found that each turbine must be treated as individual. The cause can be found to be different component suppliers used by turbine manufacturers. The turbines are built with components, e.g. generators from different manufacturers, even if the turbines are to be erected in the same wind farm. It was found to be advantageous to first implement the regression with the highest correlation to the target signal. Otherwise an optimal model cannot be found. This leads to an increased effort required when implementing an automatic model development procedure to handle a larger fleet of turbines. Here the NN training process has the inbuilt capability of giving more weight to the signals that represent the patterns in the output signal best. Hence knowledge about the individual contribution of each signal is not required and the training procedure attempts iterating to solutions that give the best performance.

Another problem with the regression model is that the input signals are correlated between each other. Thus implementing the regression of one input signal influences the correlation of the prediction error to the other input signals. The main challenge is to find a model complexity that has implemented the main signal correlations and has a low correlation of the prediction error to any of the input signals.

A full nonlinear NN has this optimization process built in the training procedure. Here the main challenge is to define the right network complexity and architecture prior to the training process.

Comparing the two NN models set up, it becomes evident, that the alarm limit of the autoregressive model is very close to the normal operational fluctuations of the prediction error. This makes it difficult to identify normal from abnormal behavior. The FSRC model on the other hand has a larger distance to the alarm limits when no signs of damage are present. Thus it has a reduced risk of producing false alarms.

## 6. Further analysis results

The results of the two NN based approaches are very close for the generator bearing damage under investigation in this paper. In order to further identify the differences between these two approaches two gearbox bearing damages and two stator temperature anomalies are investigated by use of the FSRC and the

autoregressive NN model only. The corresponding time series are shown in Figure 4 to 12.

### 6.1. Gearbox bearing damage I

#### 6.1.1. Model development

The relevant signals to model the gearbox bearing temperature are found by using the cross-correlation function and by considering the underlying physical process. For the FSRC model the model input signals are (compare Figure 1):

- Gearbox oil sump temperature (t-lag) {pos.2}
- High speed shaft revolution speed (t-lag) {pos.4}
- Power output (t-lag) {pos.7}
- Nacelle temperature (t-lag) {pos.1}
- Ambient temperature (t-lag) {pos.8}

For the autoregressive model the most relevant input signal representing the operational mode of the gearbox is the power output. The model inputs are:

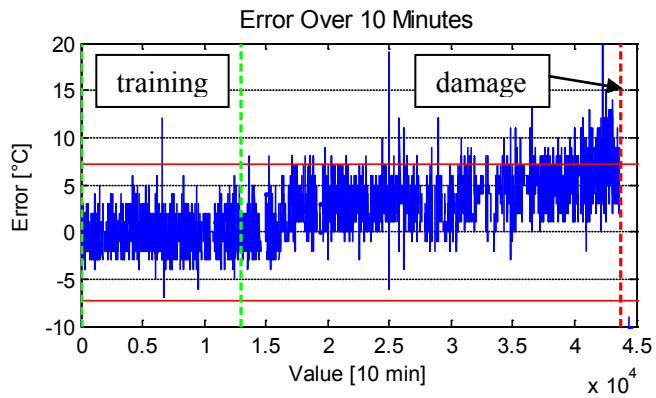
- Gearbox bearing temperature (t-2) {pos.3}
- Gearbox bearing temperature (t-1) {pos.3}
- Power output (t-lag) {pos.7}
- Nacelle temperature (t-lag) {pos.1}
- Ambient temperature (t-lag) {pos.8}

The architecture of the feed forward MLP network is similar to the one used for modeling the generator bearing temperature, i.e. a two layer network. In contrast, Zaher (et al.) [4] use a three layer network when analyzing a gearbox damage. Using more than two layers was, however, not found to be advantageous for the signal constellation used in this paper.

The number of neurons used in the hidden layer is four for the autoregressive NN and five for the FSRC NN respectively. The same data pre-processing is performed as in the previous example.

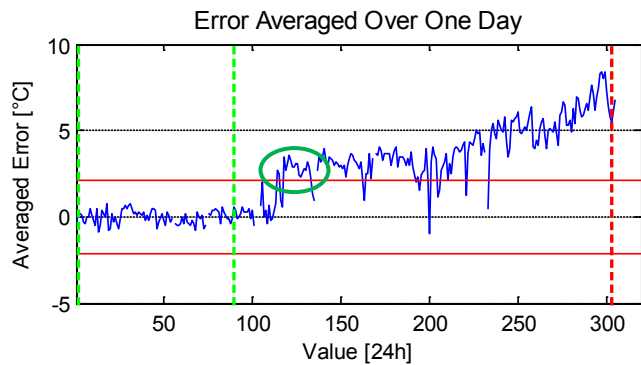
#### 6.1.2. FSRC neural network model

The model is trained by using the first three months of 10 min. average values of the time series. In Figure 30 the model prediction error is shown for the entire time series including the training period.



**Figure 30: Prediction error for the FSRC NN model**

The alarm limits are with  $7.2\text{ }^{\circ}\text{C}$  in the same order of magnitude as for the generator bearing temperature. 3000 values after the training period, the prediction error starts increasing continuously. The trend becomes more obvious when the prediction error is averaged over one day, illustrated in Figure 31.



**Figure 31: Averaged prediction error for the FSRC NN model**

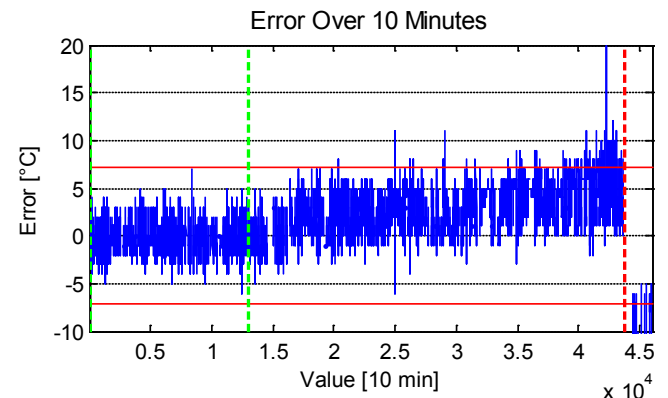
It is visible that a first ALV is present 190 days prior to the catastrophic bearing damage. The confidence level is, however, low, since the violation may be considered as an outlier. After this event the prediction error falls back to its normal operational range. A second ALV takes place about 184 days before the catastrophic bearing damage. Here the confidence level increases. The prediction error stays on an abnormal level for several days. The mean error is  $2\text{-}3\text{ }^{\circ}\text{C}$  and stays almost constant up to 100 days prior the damage, where the model error continuously increases. The final model deviation is  $7\text{ }^{\circ}\text{C}$ .

### 6.1.3. Autoregressive neural network model

In contrast to the generator bearing model, the autoregressive NN model for the gearbox bearing required the past two values rather than just one as model input to set up an accurate model. Sanz-Bobi (et al.) [3] and Zaher (et al.) [4] also used two previous values as model inputs to model gearbox related signals. Although it was not necessary for prediction of the generator bearing temperature, it is expected to give more accurate results generally. This is in particular true, when the signal to be modeled has a high autocorrelation. For thermal signals related to the

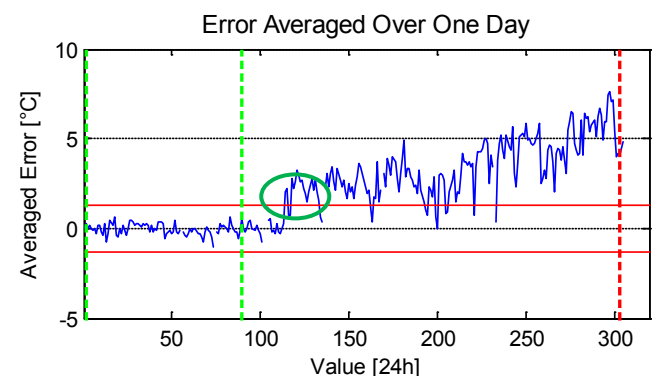
gearbox this is because of the large system mass and heat capacity.

Figure 32 shows the prediction error for the autoregressive NN model including the training period.



**Figure 32: Prediction error for the autoregressive NN model**

The alarm limit is  $\pm 7.1\text{ }^{\circ}\text{C}$ . An ALV is first visible 12000 values after the training process has been completed, i.e. 83 days prior to the damage. By this time a shift in mean is already visible. Again the one day average can be used to better identify the underlying trend and to increase the sensitivity by means of reduced alarm limits.



**Figure 33: Averaged prediction error for the autoregressive NN model**

Due to the alarm limits being reduced to  $1.3\text{ }^{\circ}\text{C}$ , an ALV is raised 185 days prior the damage. Considering the low variations prior to this event the probability of an anomaly being present is high. The prediction error stays on a high level for several days, before it shortly falls back. With the damage progressing, the variations do drastically increase. This is a clear indicator that the model deviation is a function of the operational mode. Further analysis will be performed by checking the cross-correlation between the prediction error and the rotational speed or the power as an alternative indicator of anomaly at a later stage of the research. For the actual damage it is, however, not necessary since pure threshold monitoring in terms of ALV highlights the anomaly.

### 6.1.4. Comparison

Both models are capable of identifying the anomaly in advance. The overall differences are small. The autoregressive NN model does show smaller fluctuations during training and thus has a higher accuracy. While the accuracy on 10 min average base is on the same level as for the FSRC NN model, the autoregressive model is more accurate on one day average base. In fact this leads to evidence of anomaly when the first ALV is observed.

Another difference is the final model error when the bearing breaks. The autoregressive model has an average deviation of 5 °C and the FSRC model of 7 °C. Considering the different alarm limits both models do perform equally well, as the final model error can be found as 350% of the alarm limit on one day average base. The table below summarizes the results of the comparison concerning anomaly detection.

	First ALV	Second ALV/ Trend
Full signal reconst. NN	190 days	184 days
Autoregressive NN	185 days	186 days

**Table 2: Comparison of the fault visibility between the FSRC and the autoregressive NN model for the gearbox bearing temperature anomaly I**

The first ALV might give a first indication of anomaly, but its usefulness is questionable due to the possibility of being considered an outlier. Ideally the model deviation should lead to a high confidence level right from the beginning, as a false alarm might be present otherwise. In this sense the autoregressive NN model performs somewhat better, although the second ALV takes place at almost the same time instant. For the generator bearing damage this behavior was found to be the other way around.

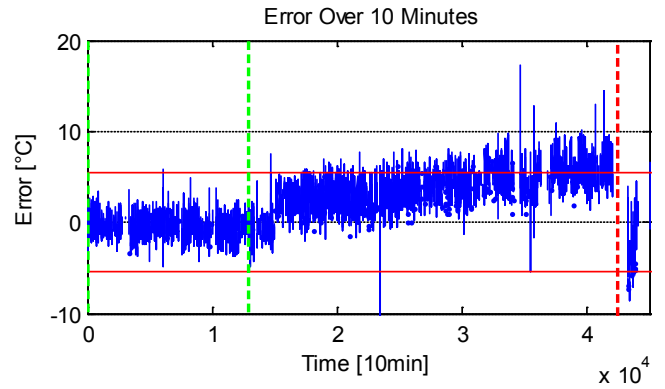
## 6.2. Gearbox bearing damage II

### 6.2.1. Model development

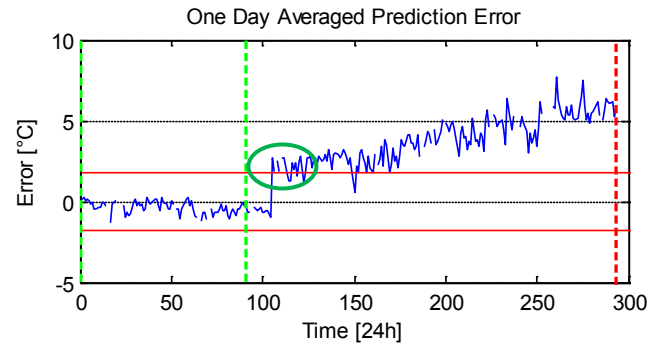
The models are set up analog to section 6.1.1.

### 6.2.2. FSRC neural network model

The corresponding time series showing the temperature evolution is shown in Figure 7. Below the prediction error is illustrated on 10 min average and on one day average base.



**Figure 34: Prediction error for the FSRC NN model**



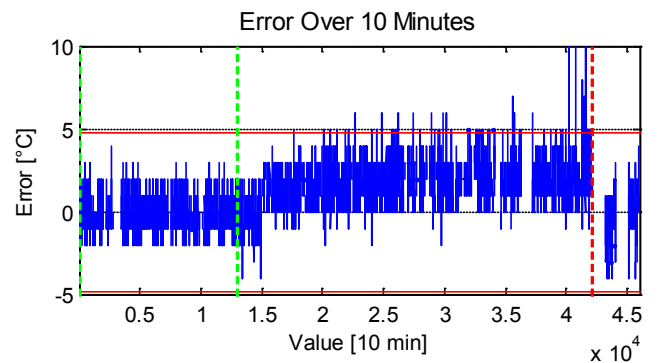
**Figure 35: Averaged prediction error for the FSRC NN model**

Like for the earlier shown gearbox bearing damage also this damage manifests in the prediction error by a trend. The anomaly starts with a steep increase in prediction error of 2.5 °C on average. Consequentially an ALV occurs 187 days prior the bearing replacement. The prediction error stays on a high level. The final model deviation is 6 °C.

The fault pattern is very similar to the one shown in section 6.1.2. Although the similar start of anomaly (steep increase in prediction error) may lead to the conclusion of an exceptional event such as high loads experienced by the two turbines, this cannot be proved as the 10 min average wind speeds show no unusual event in this period.

### 6.2.3. Autoregressive neural network model

The prediction error for the autoregressive NN model is shown in Figure 36 and Figure 37 below.



**Figure 36: Prediction error for the autoregressive NN model**



## 6.3. Generator stator temperature anomaly I

### 6.3.1. Model development

The input signals for the generator stator temperature models are found by using the cross-correlation function and considering the underlying physical process. For the FSRC model the model input signals are: (compare Figure 1)

- Power output (t-lag) {pos.7}
- High speed shaft revolution speed (t-lag) {pos.4}
- Nacelle temperature (t-lag) {pos.1}
- Ambient temperature (t-lag) {pos.8}

The autoregressive model is set up by using two previous values of the stator temperature and the power output as indicator for the operational mode of the generator. The model inputs are:

- Stator temperature (t-2) {pos.6}
- Stator temperature (t-1) {pos.6}
- Power output (t-lag) {pos.7}
- Nacelle temperature (t-lag) {pos.1}
- Ambient temperature (t-lag) {pos.8}

### 6.3.2. FSRC neural network model

The corresponding time series showing the generator stator temperature is shown in Figure 9. Besides the turbine downtime caused by the generator replacement no anomaly is obvious in this time series. The prediction error of the normal behavior model developed is shown in Figure 38 and Figure 39.

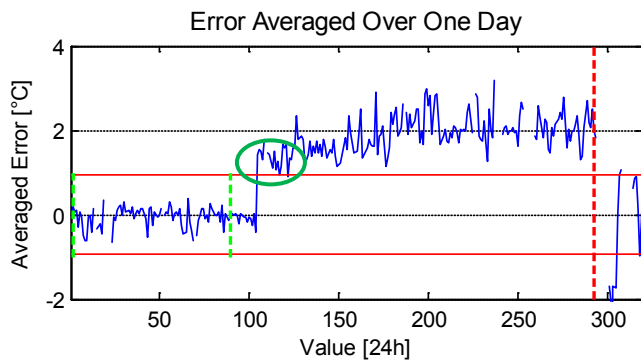


Figure 37: Averaged prediction error for the autoregressive NN model

Also for this type of model the anomaly starts with a steep increase in prediction error 187 days prior the bearing replacement. The averaged prediction error stays above the alarm limit until the bearing is replaced. A second ALV is raised one day later. The final model deviation is 2 °C.

### 6.2.4. Comparison

Due to the type of damage being potentially the same as for the presented gearbox bearing damage I, the manifestation of the anomaly in the prediction error is similar. Comparing the two types of approaches the autoregressive NN approach predicts the bearing temperature more accurately. With respect to anomaly detection of this type of fault this property does not lead to earlier or more pronounced model deviations in case of fault occurrence.

The table below summarizes the results gained concerning anomaly detection for the two approaches.

	First ALV	Second ALV/ Trend
Full signal reconst. NN	187 days	186 days
Autoregressive NN	187 days	186 days

Table 3: Comparison of the fault visibility between the FSRC and the autoregressive NN model for the gearbox bearing temperature anomaly II

Both approaches highlight the anomaly at the same time. However, there are differences in the way the fault manifests in the prediction error. While the FSRC approach leads to a continuously increasing model deviation, the autoregressive approach increases only 100 days and then oscillates around a mean model deviation of 2 °C. If the condition of the bearing is to be evaluated on defined thresholds of the model deviation, the autoregressive approach leads to a stagnating condition level. Furthermore the final model deviation is 200 % of the initial alarm limit, for the autoregressive approach, while it is 300 % of the alarm limit for the FSRC model.

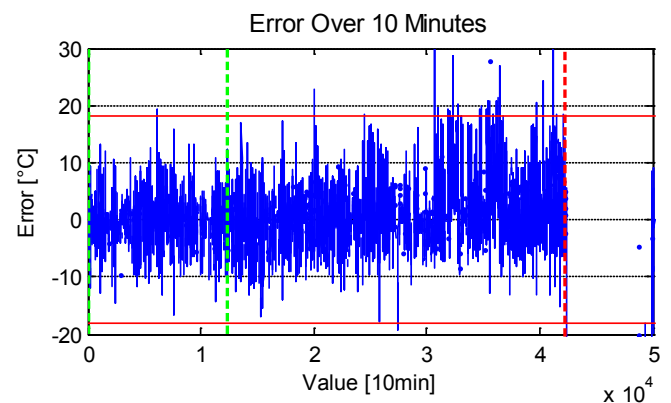


Figure 38: Prediction error for the FSRC NN model

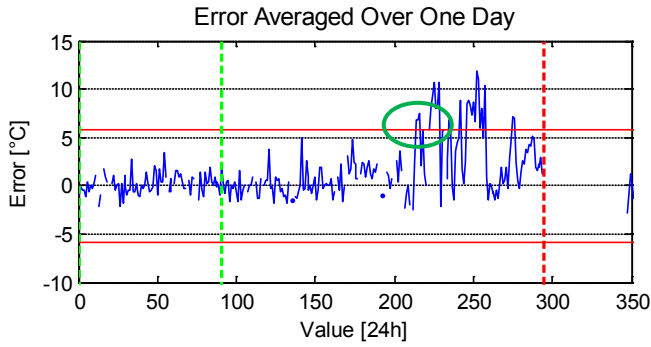


Figure 39: Averaged prediction error for the FSRC NN model

The prediction error shows a starting increase in amplitude 30 days after training is completed. For the current fault this is potentially caused by a degradation of the generator stator, i.e. the stator has increased losses.

The first ALV occurs 81 days before the generator cannot be connected to the grid anymore. The prediction error stays on a high level for four days, leading to a second ALV be raised one day later. After further four days the prediction error falls back to the normal operational range.

Although by this time the ALV might be considered as an outlier, five days later the prediction error again increases and the error exceeds 10 °C. This is 70 days before the generator needs to be replaced. The maximum model deviation is 12.1 °C.

### 6.3.3. Autoregressive neural network model

Below the corresponding prediction error for the autoregressive NN model is shown.

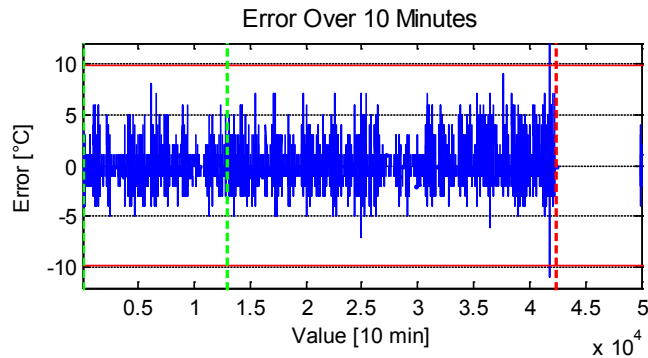


Figure 40: Prediction error for the autoregressive NN model

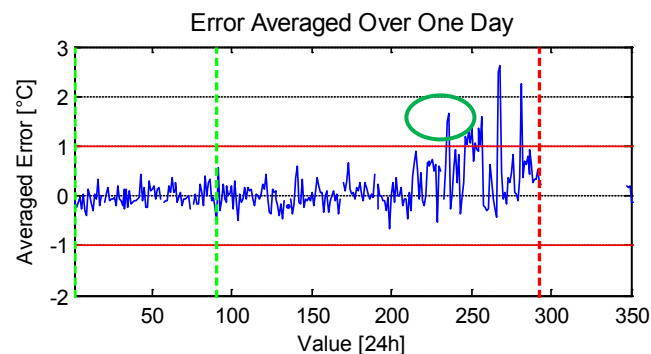


Figure 41: Averaged prediction error for the autoregressive NN model

Also the autoregressive NN model shows an increase in amplitude starting 30 days after training is completed. On one day average base the alarm limits are down to the measurement accuracy of  $\pm 1$  °C. The first ALV is present 59 days before the generator needed to be replaced. After this violation the prediction error falls back below 1 °C and the alarm limit is violated a second time 48 days prior the generator damage. The maximum model deviation is 2.6 °C.

### 6.3.4. Comparison

Both approaches show clear signs of anomalies in the prediction error. Differences exist, however, in the operational time left after the anomaly is detected and the amplitude of model deviation.

The remaining operational time after the anomaly is highlighted is summarized in Table 4.

	First ALV	Second ALV/ Trend
Full signal reconst. NN	81 days	70 days
Autoregressive NN	59 days	48 days

Table 4: Comparison of the fault visibility between the FSRC and the autoregressive NN model for the generator stator temperature anomaly I

The FSRC model highlights the anomaly 22 days earlier than the autoregressive model.

The fault pattern in the prediction error for both model types is similar so that there is little difference in automatically detecting the anomaly. Anomaly detection for both model types can be based on thresholds. In the current research project the amplitude of model deviation is used to classify the component condition. For the autoregressive NN model the maximum model deviation is 260 % of the initial alarm limit, while it is 200 % for the FSRC NN model.

## 6.4. Generator stator temperature anomaly II

### 6.4.1. Model development

The models are set up analog to section 6.3.1.

### 6.4.2. FSRC neural network model

The time series illustrating the anomaly in the raw temperature data is shown in Figure 11. The prediction of the FSRC NN model is shown in the two figures below.

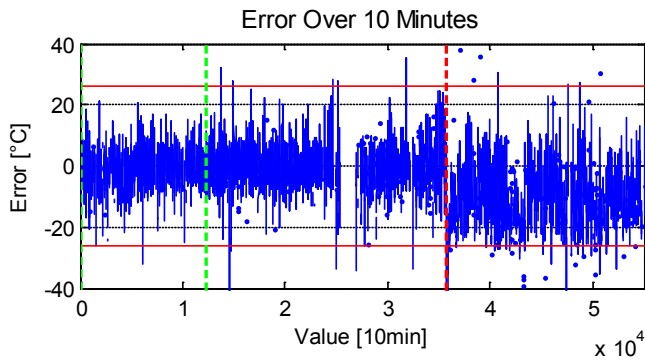


Figure 42: Prediction error for the FSRC NN model

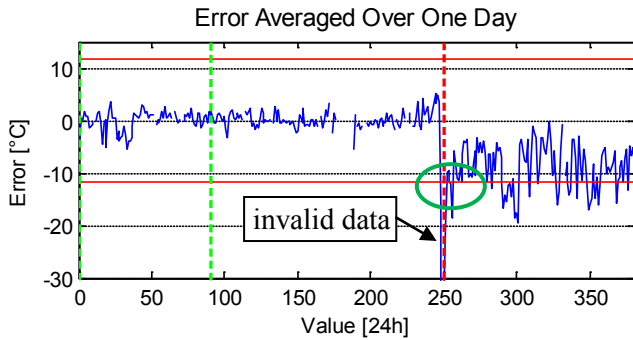


Figure 43: Averaged prediction error for the FSRC NN model

The prediction error shows no sign of anomaly before the service action begins. During service several invalid data points are generated, emphasizing the need for excluding the data generated during service from the analysis. Here many turbine manufacturers supply a digital indicator or the time in the 10 min average period where the turbine was under service. For the current data set such an indicator was, however, unavailable.

The anomaly in the data becomes obvious right after service is finished. From then on, the alarm limit is frequently violated due to a shift in mean of 10 °C. The amplified amplitude after service indicates a change in the dependency between power output and the stator temperature. If this type of fault (pressure decrease in the water cooling system) happens slowly over time, the prediction error will incorporate a trend plus an amplitude amplification. The fault pattern is likely to look similar to the earlier discussed stator temperature anomaly I.

The uncertainty in diagnosing the exact fault cause is not considered a problem at the current state of research, as in any case the fault requires service personal to access the turbine and further investigate the root cause, when the next scheduled visit takes place. As more faults and fault patterns are going to be investigated during the research project it is expected to be possible to distinguish different root causes by considering the prediction error fault patterns.

The maximum peak value of the prediction error is 19.8 °C.

### 6.4.3. Autoregressive neural network model

Figure 44 and Figure 45 show the prediction error of the autoregressive NN model.

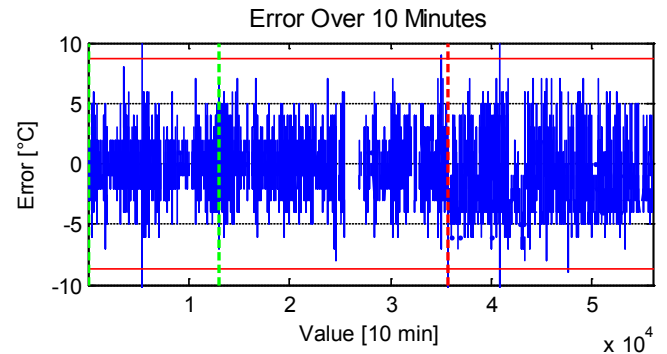


Figure 44: Prediction error for the autoregressive NN model

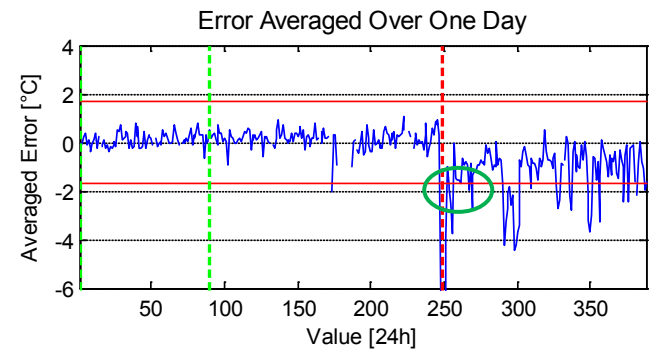


Figure 45: Averaged prediction error for the autoregressive NN model

Also here the invalid data generated during service are visible and stress the need to filter the data where service personnel is at the turbine in order to reduce the number of false alarms. After service the prediction error mean shifts by 1.5 °C and the amplitudes of the oscillations are increased. The alarm limits are frequently violated and the prediction error shows a dependency on the operational mode. The maximum model deviation during the observed period is 4.5 °C.

### 6.4.4. Comparison

The main difference of the two approaches for this type of anomaly lies in the amplitude height and the shift in mean. While the FSRC model shows the shift in mean 1:1 with the visual shift in the raw temperature data, the autoregressive model has a smaller shift. Like for the previously discussed stator temperature anomaly the autoregressive model has its advantage in having larger amplitudes in case of anomaly in comparison to the alarm limits.

Both model types highlight the present anomaly right after service is finished. The autoregressive model is more accurate in predicting in the temperature under normal conditions, leading to very tight alarm limits.

## 7. Conclusion

The comparison for the two types of bearing damages and the stator temperature anomalies reveal some differences between the three approaches.

The linear regression model is based on correlation, which in general terms is linear analysis. Wind turbine signals on the other hand may contain nonlinear relationships. In this sense, a regression model may not be applicable to all wind turbine signals. Since this paper only gives an introduction to possible applications, the procedure in favour must be widely applicable to monitor all SCADA-data. Although this was not tested during the current research, it is expected that NNs provide this feature.

All methods require proper data pre-processing and validation in order to allow optimal generalization.

Due to its simplicity, the simple regression approach may be used to monitor for instance the individual coil temperatures of the generator stator or rotor system by use of sister signals. For the generator bearing temperature, it did, however, prove unpractical due to the required effort to set up a working model. Since each turbine must be treated as individual, a high degree of automatization development is required to reduce the effort.

The NN based models used in this research, perform better and give an earlier confidence about a damage being present. However, they are more difficult to interpret. In SIMAP [3] fuzzy logic is employed to identify the abnormal behaving signal.

Apart from a simple alarm limit violation to identify abnormal behavior, further analysis tools are required to identify anomalies. This research and development was done by Sanz-Bobi (et al.) in 2006 and is employed in their anomalies detection and the health condition assessment module as part of SIMAP.

The three bearing damages and the two stator temperature anomalies indicate that the form of deviation can be quite different. For the anomalies under research the FSRC and the autoregressive NN demonstrate a high ability to detect anomalies at comparable time instants. An exception was found by the generator damage discussed in section 6.3. Here the FSRC model highlighted the anomaly considerably earlier.

In case of the bearing damages investigated the FSRC models allow easier abnormal behavior identification due to larger shifts in mean, once the damage or anomaly progresses.

An autoregressive model therefore does not always seem advantageous although Zaher (et al.) and Sanz-Bobi (et al.) used this approach to identify gearbox damages. They were able to detect incipient problems 6 month (Zaher et al.) and 2 days (Sanz-Bobi et al.) before the actual failure [1].

For reliable fault identification, the confidence level is of major importance. False alarms must be prevented. The system currently under research will be tested in a field test on the data coming from an

onshore wind farm consisting of 18 turbines in the next project stage in order to test the reliability in fault detection. For this wind farm also detailed information about carried out services and occurred damages will be available.

Reanalyzing data in order to recognize patterns after a certain fault has occurred is a rather trivial task. The major challenge, however, is to confidently identify incipient faults in online data. Here advantages of each model type depend on the type of damage.

Autoregressive models should be used when a FSRC model cannot be set up due to lag of correlated signals or if the autoregressive behavior is of particular interest. Furthermore it is more applicable to high inertia systems, i.e. slow changing signals.

The FSRC models set up in this research worked well due to the high correlation to other closely correlated signals available. FSRC models also have another important property that makes them more advantageous to use than autoregressive approaches: Since the signal is fully reconstructed, it is independent on previous measurements. For this reason, not only the relative signal changes can be monitored, but also the absolute changes. This statement holds for all types of anomalies. The property is expected to allow identification of purely sensor related issues and other types of faults.

The drawback is that FSRC models are expected to have a higher risk of producing false alarms. This is because they may require more input signals.

The comparison thus showed that FSRC models can be favorable to use over autoregressive approaches. The regression based models are simple to interpret and may be applied to simple problems, but are outperformed by the NN approaches in case of the anomalies investigated.

Future research should focus on applying the findings to further SCADA signals. Thereby the most appropriate NN approach must be identified for each signal to be modeled individually. Also more research is required to investigate the potential of false alarms and the general fault visibility. This will be part of the ongoing Ph.D. research project by applying the procedure to a larger turbine fleet and long time monitoring of the signals behavior.

## References

- [1] **Zaher, A., McArthur, S.D.J. and Infield, D.G.** (2009). "Online Wind Turbine Fault Detection through Automated SCADA Data Analysis." *Wind Energy*. 12 pp.574-593.
- [2] **Verbruggen, T.W.** (2003). "Wind Turbine Operation & Maintenance based on Condition monitoring." ECN-C-03-047.
- [3] **Sanz-Bobi, M. A., del Pico, J. and Garcia, M. C.** (2006). "SIMAP: Intelligent System for Predictive Maintenance Application to the health condition monitoring of a windturbine gearbox." *Computers in Industry* 57 pp.552-568.

- [4] **Zaher, A.S. and McArthur, S.D.J.** (2007). "A Multi-Agent Fault Detection System for Wind Turbine Defect Recognition and Diagnosis. "*Proc. IEEE Lausanne POWERTECH*. pp. 22-27.
- [5] **Shuhui, L., et al.** (2001). "Comparative Analysis of Regression and Artificial Neural Network Models for Wind Turbine Power Curve Estimation. "*Journal of Solar Energy Engineering*. 123 pp.327-332.
- [6] **Süttmann (Schlechtingen), M.** (2010). "Master Thesis: Condition Monitoring in Wind Turbines - A Drive Train Monitoring System. "
- [7] **Blundell, S. and K. M., Blundell.** *Concepts in thermal physics*. 2006. 978-0198567691.
- [8] **Haykin, S.** (1998). "Neural Networks: A Comprehensive Foundation. "ISBN: 0132733501. Prentice-Hall.
- [9] **Rafiq, M.Y., Bugmann, G. and Easterbrook, D.J.** (2001). "Neural network design for engineering applications. "*Computers and Structures*. 79 pp.1541-1552.
- [10] **Swingler, K.** (1996). "Applying Neural Networks - A Practical Guide. "ISBN: 0340705892. Elsevier.
- [11] **Tarassenko, L.** (1998). "Guide to Neural Computing Applications. "ISBN:0340705892. Elsevier.
- [12] **Caselitz, P. und Giebhardt, J.** (2001). "Advanced Maintenance and Repair for Offshore Wind Farms using Fault Prediction Techniques. "Institut für Solare Energieversorgungstechnik e.V. - Department of Control Engineering.
- [13] **Kung, S.Y. and Hwang, J.N.** (1988). "An Algebraic Projection Analysis for Optimal Hidden Units Size and Learning Rates in Back-Propagation Learning. "Princeton University.