

A global condition monitoring system for wind turbines

Schlechtingen, Meik; Santos, Ilmar

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Schlechtingen, M., & Santos, I. (2013). A global condition monitoring system for wind turbines. DTU Mechanical Engineering. (DCAMM Special Report; No. S150).

DTU Library

Technical Information Center of Denmark

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.

A Global Condition Monitoring System for Wind Turbines

by

Meik Schlechtingen

DEPT. OF MECHANICAL ENGINEERING
Solid Mechanics



TECHNICAL UNIVERSITY OF DENMARK

In cooperation with



EnBW Erneuerbare Energien GmbH

Title of the thesis:

A global condition monitoring system for wind turbines

Ph.D. student:

Meik Schlechtingen

E-mail: m.schlechtingen@enbw.com

Supervisor:

Prof. Ilmar F. Santos

E-mail: ifs.mek.dtu.dk

Address:

EnBW Erneuerbare Energien GmbH

Admiralitätstr.4

20459 Hamburg, Germany

Copyright © 2013 Meik Schlechtingen

DCAMM Special Report no.: xx

ISBN: xx

Preface

This thesis is submitted in partial fulfillment of the requirement for obtaining the degree of Ph.D. in Mechanical Engineering at the Technical University of Denmark (DTU).

The idea for this Ph.D. project arose already prior my master degree while working in industry at a wind turbine manufacturer. Part of my job was to analyze wind turbine measurement data proactively in order to minimize unnecessary downtime and keep the turbine availability high. This work was very labor intensive and due to the large amount of data it should be considered luck, that I discovered some gearbox bearing degradation by means of bearing temperature analysis (step increase of 5°C in bearing temperature). This defect was not picked up by any automatic algorithm. Data are collected to a large extent, but apparently there was a lack of efficient automatic analysis methods, highlighting such anomalies. On the one hand I was happy about the success, on the other hand the question arose how many of those anomalies remain undiscovered.

During my master course I dedicated several projects to this topic, all confirming, that there are efficient algorithms available which could potentially be used for anomaly detection. None of them were fully validated or applied in larger scale in wind industry. Motivated by curiosity and excitement I wanted to push this research further with a view to setting new standards and opening up for large-scale industry application. I decided to do this in terms of a Ph.D. project since a systematic research approach seemed best suited to achieve this. However, such a project calls for industrial cooperation, which made me start looking for a strong industry partner, operating a large number of wind turbines and as such having access to data. With EnBW Erneuerbare Energien GmbH such a partner was quickly found. Not only did EnBW (a German utility company) take care of financing the project entirely, they also employed me as a Ph.D. student throughout the three year project.

In particular, I would like to thank Prof. Dr. Ilmar F. Santos from the department of Mechanical Engineering at the Technical University of Denmark, who supervised this thesis for the great support, the many fruitful talks, the motivation in difficult times and the constructive criticism during the project.

Additionally I would like to thank Dr. Werner Götz, formerly managing technical director at EnBW Erneuerbare Energien GmbH not only for replying to my informal e-mail at first contact during my search for an industrial partner, but also for the trust he had in me and the outcome of the project right from the beginning. I owe him and his quick decision making the possibility of focusing on the research project without any restrictions.

A special thank you also belongs to my manager Michael Boll, Offshore Operations Manager at EnBW Erneuerbare Energien GmbH for the trust and support throughout the project: for keeping my back free from day-to-day business, supporting investments, accepting me as an equally valued employee and for the unconditional trust with the direction of my project. His support was crucial to achieving the outcome of my work.

A further thank you belongs to Dr. Eric Gross at the Institute of Reliability at the Technical University of Hamburg (Germany) for giving me the possibility for an external research stay. The many fruitful talks and discussions during the course of this stay widened my understanding of signal processing and its involved algorithms. His knowledgeable and useful hints and comments supported me in quickly getting used to this field.

Finally yet importantly, I would like to give a special thank you to my family for giving me the strength and encouragement for starting this project and accepting all consequences and compromises involved in this decision. In particular it is my wife who suffered from my self-criticism and my doubts in difficult times of the project. Nevertheless, she and my little son (born in the middle of the project) had to cope with me being at DTU reams of times for several days, putting additional load on her shoulders. That is definitely not something that I take for granted.

Hamburg, 14th of February 2013

Meik Schleutgen

Resumé (in Danish)

Energiprisen fra vindmølleparker (offshore-parker specielt) udfordrer samfund i forbindelse med krav om lavere og mere miljøvenlig genereret elektrisk energi. Målet om en prisreduktion indfries ved hjælp af anvendelse af tilstandsmonitoreringssystemer, som kan hjælpe vindmøllejeeren med at reducere reparationstiden samt generelle omkostninger. På baggrund af dette forskningsprojekt er et globalt tilstandsmonitoreringssystem foreslået og resultatet for dette er en platform som forsyner operatører med forskellige tilgængelige informationskilder.

En af de mest almindelige informationskilder med hensyn til komponenters tilstand er Supervisory Control And Data Acquisition (SCADA) data, fx temperature, strømstyrke og spændingsmålinger fra forskellige komponenter.

Ved at bruge nyligt udviklede Adaptive Neuro-Fuzzy Interference System (ANFIS) modeller, udtages information fra disse data på baggrund af modeller om normal adfærd; data som ellers består af signaler med stor varians. Nyheden her er brugen af ANFIS modeller i denne sammenhæng, som traditionelt set er dækket af neurale netværk eller regressionsbaserede fremgangsmåder. Fordelene ved kortere træningstid og resultatets sporbarhed i systemet, motiverer for ibrugtagning af disse lovende modeltyper.

Metoder beskrevet i litteraturen tager udgangspunkt i autoregressive metoder, dvs. at en forsinket version af outputsignalerne bruges som input ved siden af andre højt korellerede signaler. I denne forskning foreslås en komplet signalrekonstruktionsmodel, hvori outputsignalet er rekonstrueret fuldstændigt ved hjælp af andre korellerede signaler. Dette har medført at fordele som fx synlighed af maskinfejl og gennemløbstiden indtil en maskinfejl er estimeret, er blevet observeret.

I SCADA dataene er vindmøllens outputeffekt et meget vigtigt signal at monitorere. Outputeffekten har en direkte indflydelse på operatørernes indtægt og derfor også på energiprisen. Modeller om normal adfærd beskrevet i litteraturen ignorerer indflydelsen af vindretningen og den ambiente temperatur, som teoretisk set influerer den forventede outputeffekt med op til 20 %. Et nyt sæt af modelinput er foreslået og valideret i denne forskning. Der er opnået højere prognosenøjagtighed og derfor også tidligere påvisning af uregelmæssigheder med hensyn til afvigelse i møllens ydelse.

En anden informationskilde som operatører kan have adgang til er vibrationsmålinger fra vindmøllens transmission. I forbindelse med gearbokslejerens tilstand er der for nyligt foreslået en model til at separere diskrete signaldele (fx fra tandhjulene) fra tilfældige signaldele (fx fra lejerne). Denne model er blevet benyttet og valideret i denne forskning. Denne avancerede metode kaldet "signal pre-whitening" øger synligheden af fejlmønstre i indgrænsede spektrum

på en meget effektiv måde. Valideringen er baseret på faktiske målte lejefejl i høj hastighed og i middel hastighed.

Derudover er der i litteraturen foreslået en halvautomatisk metode til diagnosticering af lejer, og denne er opdateret med de mest lovende teknikker indenfor forbehandling af signaler for at øge synligheden af fejl. Ydermere er den udviklet til at være en fuldautomatisk diagnosticering. Til dette formål er en frekvenskomponentindikator udviklet til at udtrække frekvensindholdet i et indgrænsede spektrum hvilket anlægger fundamentet til den automatiserede diagnosticering.

En modificeret parameter, kaldet the Kurtosis of the Amplitude Envelope Spectrum (KEAS), er foreslået og dens ydeevne er sammenlignet med den originale parameter, the Spectral Kurtosis, som bruges til at indikere alvorligheden af fejl.

Ydermere, er avancerede metoder til monitorering af vindmølletandhjul i korte træk blevet sammenlignet, og de mest lovende parametre er blevet identificeret ved hjælp af et tandhjulpår med false brinelling.

Slutteligt er et global fuzzy expert system blevet udviklet og det giver muligheden for at sammenkoble al tilgængelig information med hensyn til fuzzy logic regler.

Det er vigtigt at fremhæve at denne forskning er baseret på faktiske målte data, som kommer fra to vindmøllerparker med vindmøller af forskellige type og mærke. Alt i alt har data fra 39 operationelle vindmøller været til rådighed i dette forskningsprojekt.

Abstract

The cost of energy generated from wind power plants (particular if located offshore) is challenging societies in terms of desiring cheaper and more environmentally friendly generated electrical energy. The high cost reduction targets can be aided by broad application of condition monitoring systems, which bear the potential to support plant owners reducing turbine downtime and lowering costs. In this research a global condition monitoring system is proposed, which provides a platform to take advantage of the different information sources available to operators.

One of the most common sources for information about the component condition is Supervisory Control And Data Acquisition (SCADA) data, e.g. temperature, current or voltage measurements from different components.

Using newly developed Adaptive Neuro-Fuzzy Interference System (ANFIS) models, a normal behavior model based approach is taken to extract information from these data, which otherwise are covered by the high signal variance. A novelty here is the application of ANFIS models in this context, which is traditionally covered by neural networks or regression based approaches. Advantages in training speed and back traceability of results are the motivators to apply these promising model types.

Methods proposed in literature usually take an autoregressive approach, i.e. a delayed version of the output signal is used as input alongside other highly correlated signals. In this research, a full signal reconstruction method is proposed in which the output signal is entirely reconstructed by using other correlated signals. Benefits in fault visibility and lead-time to failure estimates are observed.

A very important signal to monitor contained in the SCADA data is the wind turbine power output. The power output has a direct influence on the revenue of the operators and thus the cost of energy. Normal behavior models from literature neglect the influence of the wind direction and the ambient temperature, which can, theoretically speaking, influence the expected power output by up to 20%. A new set of model inputs is proposed and validated in this research. Higher prediction accuracy and thus earlier anomaly detection in case of turbine performance deviations is achieved.

Another information source potentially available to operators is vibration measurements from the wind turbine drive train. With respect to the condition of gearbox bearings, a recently proposed method to separate discrete (e.g. originating from gears) from random (e.g. originating from bearings) signal components is applied and validated in this research. This state of the art method named “signal pre-whitening” enhances the fault pattern visibility in the envelope spectra in a very efficient manner. The validation is based on real measured high-speed and intermediate speed bearing faults.

Additionally, a semi-automated method for bearing diagnosis suggested in literature is updated with the most promising signal pre-processing techniques to enhance fault visibility and further developed leading to fully automated fault diagnosis. For this purpose a frequency content identifier is developed extracting the frequency content from the envelope spectrum building the basis for automated diagnosis.

A modified parameter, namely the Kurtosis of the Amplitude Envelope Spectrum (KEAS), is proposed and its performance is compared to the original parameter, the Spectral Kurtosis (SK), which is used to indicate the fault severity.

Moreover, state of the art methods to monitor wind turbine gears have been briefly compared and the most promising parameters are identified using a gear pair with false brinelling.

Finally, a global fuzzy expert system is developed giving the possibility of linking all available information in terms of fuzzy logic rules.

It is important to highlight that this research is based on real measured data coming from two wind power plants with turbines of a different type and brand. In total, operational data from 39 turbines were available for this research project.

Table of contents

Preface	iii
Resumé (in Danish).....	v
Abstract	vii
Table of contents	ix
List of abbreviations.....	xi
List of symbols	xiii
1 Introduction	1
1.1. Problem description / Motivation.....	1
1.2. State of the art in wind turbine condition monitoring	2
1.3. Goals, originality and thesis structure	7
1.3.1. Goals.....	8
1.3.2. Originality.....	8
1.3.3. Thesis structure.....	9
1.4. Available data.....	10
2 General system setup.....	15
2.1. Description	15
3 SCADA Data analysis	17
3.1. Publication [P1] (published).....	17
3.2. Publication [P2] (published).....	45
3.3. Publication [P3] (published) & [P4] (submitted)	56
3.3.1. Extended research and development - Additional information	87
4 Vibration data analysis	91
4.1. Publication [P5] (submitted)	91

4.1.1. Extended research and development - Additional information	112
4.1.1.1. HSS bearing defects	112
4.1.1.2. IMS bearing defects	116
4.2. Spur gears	119
4.2.1. Cepstrum analysis	120
4.2.2. Cyclostationary analysis	122
4.2.3. Time synchronous average.....	122
4.2.3.1. Ideal filtering.....	124
4.2.3.2. Linear prediction.....	128
4.2.3.3. Amplitude and phase demodulation	129
4.2.3.4. TSA conclusions	131
4.2.4. Wavelet transform.....	132
4.2.5. Transmission error	136
4.2.6. Sideband energy ratio.....	136
4.2.7. Discussion	139
5 Global fuzzy expert system	140
5.1. System overview and description	141
5.2. Application examples	144
5.2.1. Example 1: HSS bearing defect	146
5.2.2. Example 2: IMS bearing defect	149
5.2.3. Example 3: Generator NDE bearing defect	152
6 Conclusions and future aspects	155
6.1. Conclusion	155
6.2. Future aspects	157
7 References.....	158
7.1. References cited in thesis.....	158
7.2. References cited in papers [P1-P5]	162
7.3. Standards.....	164

List of abbreviations

AC	Alternating Current
AIC	Akaikes entropy based Information Criterion
ANC	Adaptive Noise Cancellation
AR	AutoRegressive
BPFI	Ball Pass Frequency Inner race
BPFO	Ball Pass Frequency Outer race
BSF	Ball Spin Frequency
CEP	Cepstrum Editing Procedure
CMS	Condition Monitoring System(s)
COE	Cost Of Energy
Cross.	Cross prediction model
DC	Direct Current
DE	Drive End
Diff.	Difference
DOG	Difference Of Gaussians
DRS	Discrete/ Random Separation
Ext.	External
FCI	Frequency Content Identifier
FDS	Fault Detection System
FFT	Fast Fourier Transformation
FL	Fuzzy Logic
FRF	Frequency Response Function
FSRC	Full Signal ReConstruction
FTF	Fundamental Train Frequency
GAPLS	Genetic Algorithm combined with Partial Least Squares
Gen.	Generator
HSS	High Speed Shaft
IMS	Intermediate Speed Shaft
Int.	Internal
KEAS	Kurtosis of the Envelope Amplitude Spectrum
KFM	Kohonen Feature Maps
LP	Linear Prediction

MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MAS	Multi Agent System
Meas.	Measurement
MED	Minimum Entropy deconvolution
MF	Membership Function
Mod.	Modulation
NBM	Normal Behavior Model
NDE	Non Drive End
NN	Neural Network
Num.	Number
OT	Order Tracking
Pres.	Pressure
PSD	Power Spectral Density
PW	Pre-Whitening
RMS	Root Mean Squared
Rot.	Rotor
RPM	Revolutions Per Minute
SANC	Self Adaptive Noise Cancellation
SCADA	Supervisory Control And Data Acquisition
SD	Standard Deviation
SER	Sideband Energy Ratio
SIMAP	Intelligent System for Predictive Maintenance
SK	Spectral Kurtosis
SRM	Stochastic Resonance Model
STFT	Short Time Fourier Transform
Temp.	Temperature
Tr.	Tracked
TS	Time Series
TSA	Time Synchronous Average
WTG	Wind Turbine Generator
WT	Wavelet Transform

List of symbols

a	Scale
$a(k)$	Weighting coefficient
B_i	Frequency component i
b	time
E	Prediction error
f_0	Gaussian shape parameter
f_s	Shaft rotation frequency
K_e	Kurtosis
M	Number of spectral data points
M	Sideband index
m	Number of sideband pairs
N	Number of teeth
n	Number of measurements
p	Number of previous values (model order)
q	Quefrency amplitude values used for reference
r_{max}	Maximum value occurring during normal operation
SER	Sideband energy ratio
$S_{max\ i}$	Maximum linear amplitude i of the sideband
t	Time instant
$thres$	Threshold
$\hat{x}(n)$	Predicted current value
$x_{m,i}$	Measurement i
x_m	Measurement
$x_{p,i}$	Prediction i
x_p	Prediction
$x(t)$	Time variant signal
y	Time series
σ	Standard deviation
$\psi(t)$	Mother wavelet

1 Introduction

Rising prices for fossil fuel and tragedies like the nuclear accident in Fukushima (Japan) motivate many countries to support and invest in renewable energy sources and their exploitation. In consequence, it can be observed that the number of wind turbines has massively increased in the past decade. Especially in the densely populated countries in Europe, there is a tendency to build offshore wind power plants with the size of several hundred MW each. Not only is the motivation for this the usually higher energy yield due to the better wind conditions offshore, but also the visual impact is thought to be lower leading to a higher acceptance in societies.

Despite the higher energy yield, the Cost Of Energy (COE) from offshore wind power plants is still high in comparison to the COE from onshore wind turbines or energy derived from fossil fuels. This is mostly driven by the high investments required for power plant planning and installation. In addition, service and maintenance costs during operation are high because of the limited turbine accessibility especially in the winter period. A turbine fault requiring a physical inspection can cause long downtime due to the harsh weather conditions at sea preventing access. Furthermore, turbine visits are expensive because of the high transport cost by either boat or helicopter. For operators it is thus of major importance to know about the condition of their turbines, to be able to schedule service actions and trigger repairs in periods with lower weather risk before the component finally fails.

Here Condition Monitoring Systems (CMSs) can aid the power plant operators in taking informed decisions. Most available CMSs focus on vibration data analysis of the main drive train components (main bearing, gearbox and generator) only. Specifically offshore, but also if turbines are placed in remote locations onshore, failures of auxiliary equipment and secondary components like fans, pumps, converters, etc. consequentially gain importance by a higher contribution to the total turbine downtime. Because of this, wind turbine CMSs shall be able to monitor main and auxiliary components, as well as the total turbine performance.

1.1. Problem description / Motivation

From the previous discussion, it seems obvious that operators should consider all available information from their turbines to gain knowledge about the turbine and component condition. Possible information sources are:

- Supervisory Control And Data Acquisition (SCADA) data, e.g. 10 min average values of temperatures, pressures, currents, voltages, etc.,
- vibration measurements of the main components,

- oil samples,
- and others.

The SCADA data are currently mostly used for backward analysis after fault occurrence to clarify the fault root cause, but not for condition monitoring at large scale. For smaller and many older wind turbines, SCADA data is the only available information source. Wind turbines are born with a SCADA data sensor setup, while vibration condition monitoring systems need to be purchased separately. In 2011, Yang and Jiang [1] pointed out that this data is the cheapest source for developing CMSs. Thus, it seems natural to utilize this source of information in parallel with vibration measurements to achieve the desired reduction in COE. In the wind industry this is not yet done, because of the lack of suitable algorithms and missing automatization in fault analysis, making an excessive data analysis labor intensive.

If operators already perform labor-intensive SCADA data analysis, the different information sources are not yet combined. This is because specialized third party experts often perform vibration data analysis, while the operators themselves analyze the SCADA data. In future, this separation is likely to diminish with the increasing number of units owned by operators. Hence there is a need to find expert systems capable of aiding operators in monitoring their assets and implement their knowledge gained, combining all available information.

However, technical details about the turbine or components design are seldom available to operators, formulating the need for self-learning algorithms with thresholds defined automatically.

1.2. State of the art in wind turbine condition monitoring

In 1989 Petersen [2] presented a knowledge based approach to carry out a technical and economic assessment of smaller wind turbines, taking into account the investment, operational costs, maintenance costs and costs for lost production due to interruptions. This project emphasized the importance of improved maintenance by implementation of a maintenance planning system [2]. The knowledge based system allowed scheduling maintenance based on historical and on-site monitored data as well as maintenance logs [2].

In 1990, Clayton et.al [3] added another possible source of information by developing a method to monitor the condition of composite rotor blades via an infrared technique. This was among the first monitoring methods especially suited for wind energy application.

Around this time, there were already a number of monitoring methods available for other applications. Among these were, for instance methods to monitor turbine generators [3], machines with rolling element bearings [4], induction motors [5] or electric machines in general [6]. For the latter, Penman and Yin [7] published the results of their feasibility study of using unsupervised learning artificial Neural Networks (NNs) for condition monitoring of electrical machines in 1994. Although a NN approach was used for induction motors before, they found that Kohonen Feature Maps (KFM) are more useful for this application due to their ability to learn without any a priori information of the distribution of the data. Moreover, they noticed a trend in condition monitoring methods towards common processing of the data from a wide range of different sensors, which leads to the concept of performance mapping of complete processes. They found that using multiple sensors also reduces the dependency of

acquired information from a system on any particular techniques and that therefore more confidence can be placed in the result of the diagnosis [7].

However, except for the rotor blade CMS none of these systems was applied to Wind Turbine Generators (WTGs). Caselitz, Giebhardt and Mevenkamp [8] were the first applying and combining several available condition monitoring methods to the special application in WTG in 1994. They described the importance of a Fault Detection System (FDS) to not only improve the operational safety, but also allowing for predictive condition based maintenance schemes.

FDSs evaluate measured process data (e.g. using spectral analysis for vibration monitoring) to isolate incipient faults at a very early stage, that is before they become optically or acoustically manifest [8]. Their development project in cooperation with Carl Schenck AG, Darmstadt (Germany) focused on the following main components:

- Rotor
- Gearbox
- Generator
- Bearings, shafts
- Yaw system
- Tower, WTG structure

In this context, they reviewed and investigated several different approaches to identify and isolate faults used in other applications.

System identification approach

The process model parameters are identified from input-output measurements and compared with a nominal parameter set. The difference or, alternatively, the identification error signal forms the fault indicator, the so-called residual. Since in general the identified parameters resemble physical process quantities, this residual often leads to a very straightforward fault diagnosis.

Observer based approach

In this approach, the residual is the observation error or innovation in the case of the Kalman filter calculated from the process measurements and the output of a reference model. The observer can be designed with reduced sensitivity against model uncertainties and structured external disturbances, but still very accurate models for the essential process dynamics are required.

Signal analysis approach

Fault indicators are derived from process measurements via limit and trend checking of process signals and by means of various spectral analysis methods, e.g. Fast Fourier Transform (FFT), Cepstrum or envelope curve analysis.

Expert system and artificial intelligence approaches

This approach is used when a process is too complex to be modeled analytically. Qualitative (expert) process knowledge can be used to evaluate relations between measured signals and current operating conditions. Frank [9] investigated fuzzy techniques in this context. In some applications, NNs have been used to substitute analytical process models. With a NN trained

on a reference process (e.g. new machine) residuals can be generated similar to the model based approaches [10].

The review of existing fault identification approaches by Caselitz, Giebhardt and Mevenkamp [8] lead to the conclusion that model based approaches can at best be applied to single components of a WTG. On the other hand, spectral analysis methods do not require input measurements and they are known to be well suited for the separation of relevant periodical signals from stochastic disturbances [8]. Thus, their development activities were focused on the spectral analysis methods in the signal processing section [8].

In their paper, they carried out a frequency analysis of the power output of a wind turbine generator. From the spectrum, it became evident that the electrical power output can be used for a detailed analysis of structural oscillations and even drive train vibrations [8].

Caselitz et.al [11] investigated not only the electrical power output of the turbine, but also applied spectral analysis to the tower bending and tower torsion moments, to increase the fault visibility in the wind turbine structure in 1996. In the scope of their development were failures like imbalance or fatigue of rotor blades, wear or impending cracks of shafts, bearings or tooth wheels of the gearbox, insulation failures of the generator or offsets in the pitch or yaw system.

They used computer simulations of the dynamic system to help with the interpretation and validation of the experimental data. Additionally a commercially available vibration monitoring system for application in other industries was tested on gearboxes and bearings of a test turbine. The test revealed that the commercial system, which was designed for power station turbines was not well suited for a WTG drive train because the costs were relatively high and the system was not designed to handle rapid power changes.

In 1997 Caselitz et.al [12] started a field test with installations in several turbines of different manufacturers. Moreover a fault classification system to combine information from different measurements based on fuzzy techniques was introduced. As a result, the reliability of fault detection was significantly improved [12]. The CMS was designed as a separate system which communicates with the plant control system via standard ports to be generally applicable in all horizontal axis turbine types. However, because of the special measurement and signal processing hardware required, an integrated realisation on the basis of conventional WTG control systems was deemed to be unfeasible. This necessitated the development of a flexible interface [12].

Figure 1.1 shows the signals used for the system developed and the integration of the CMS in a WTG.

The system set up in this manner allowed for real time analysis and monitoring. In addition to their prior work they now also included algorithms to monitor the overall turbine performance via a power curve analysis (power output vs wind speed). Gearbox and bearing faults were monitored by an envelope curve analysis of the frequency spectra.

Furthermore the idea of a CMS network in a wind power plant is described. The advantages were expected to be in an optimized fault classification by using comparisons between WTGs operating under identical conditions and thus in new monitoring functions significantly improving the efficiency of fault detection and the performance of each single WTG in a wind power plant [12].

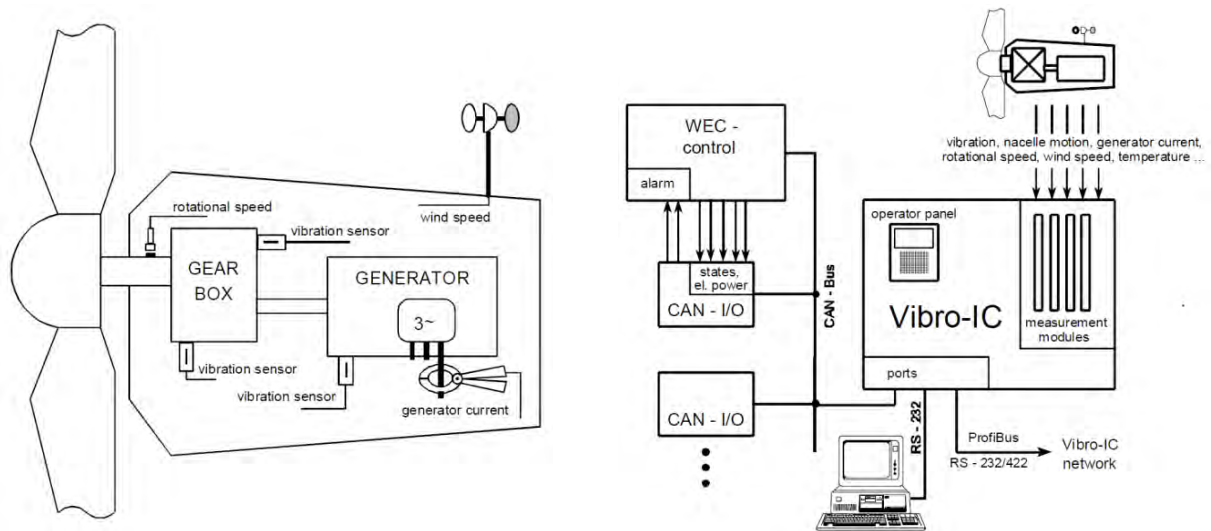


Figure 1.1: Measurements for condition monitoring [12] (left); Integration of the CMS in a WTG [12] (right)

Two years later Caselitz and Giebhardt [13] presented the results of the field test and the development of a processing unit. The system was able to monitor the main components and performance of the turbine by using a so called object model. This model consists of several layers; the time signal acquisition layer, signal processing layer generating characteristic values and the evaluation and diagnosis layer capable of sending out alarms.

A novation was that the objects in the evaluation and diagnosis layer are to perform a learning mode to determine warning and alarm levels autonomously. The levels are calculated from the mean and standard deviation values with weight factors as parameters [13].

However the field test revealed a limitation of the spectral analysis methods used for gearbox and bearing monitoring. It was found to be very difficult to obtain the detailed knowledge about the construction parameters of the WTG gearbox and bearings [13]. Moreover it was found that the analysis of the electrical power output described in earlier papers [8] [11] by the use of spectral analysis, is less useful than focusing on the nacelle oscillations induced by the rotor to detect and distinguish rotor fault conditions [13]. In overall, it was concluded from the experience gained during the field test, that with the developed system there is a CMS fit for the market to obtain rising effects on the economics of all types of WTGs.

However, the detailed knowledge of the component design parameters required to apply the developed CMS causes major problems in practical application of this system, since this information is seldom available to operators.

Already in 1991 there was a method developed by Chow, Mangum and Yee [5] to overcome this shortage. Their NN approach for real time condition monitoring of induction motors did not require prior knowledge about the component. In a learning phase a training algorithm determines the dependencies between various inputs and the desired output signals. These dependencies can then be used in the application or monitoring phase to detect deviations from the normal (learned) behaviour. Their NN based incipient fault detector provided an accuracy greater than 95% [5]. The method described refers to the expert system and artificial intelligent approach investigated for the use in wind turbine applications by Caselitz, Giebhardt and Mevenkamp [8].

The particular advantage of independency of prior knowledge about the component or system lead to an increasing interest in those techniques for wind turbine applications once the limitation became evident.

In 2001 a comparative analysis of regression and artificial neural network models for wind turbine power curve estimation was published by Li, O'Hair and Giesselmann [14]. Wind turbine power curve monitoring can give a good indication of various faults like yaw and blade pitch errors or blade defect. The comparison showed that the regression model relies on a well defined function to give accurate results, while the NN model relies on good training. Since the definition of an accurate regression function is difficult, the NN model performed better [14].

Two years later Verbruggen [15] published a final report of a project, reviewing different condition monitoring methods for wind turbine operation and maintenance. It covered a large amount of wind turbine components and also included an experimental study. For each of the components a possible monitoring method was described. Moreover there is also a section included describing a global monitoring system. Verbruggen [15] states that there is already a lot of information available in the signals measured in wind turbine, e.g. information about temperatures, pressures, currents, etc. Normally this information is only used at the level of safeguarding. He recommends using this information to detect significant changes in wind turbine behaviour. This idea mainly focuses on the analysis of 10-minute averages. A major advantage is that no additional investments in hardware is necessary. Only development costs are involved [15].

In this context also model based condition monitoring by use of NN is mentioned for instance for the pitch mechanism. However, there is no specific information given in the project report. McArthur, Booth and McFadyen [16] presented an agent-based anomaly detection architecture for condition monitoring in 2005. This system allows combining different sources of information to arrive at an overall statement about the component condition. This is similar to the object model developed by Caselitz and Giebhardt [13], but without the need of detailed knowledge about how the fault manifests in the data.

The research activities by McArthur, Booth and McFadyen [16] focus on an agent working in the absence of this knowledge. In this sense algorithms are based on artificial intelligence approaches i.e. NNs, and thus incorporate learning fingerprinting the plant item. Any deviations from this "learned" behavioural envelope will result in the highlighting of an anomaly [16].

As a result, in this system individual agents will perform different categories of functions and will intercommunicate to implement the overall anomaly detection functionality [16].

The prototype anomaly detection system was tested on a 2.5 MW diesel engine driven alternator system. A total of 50 parameters were measured from the engine/alternator system. These consisted of engine load, generator output, individual cylinder temperatures and pressures, exhaust temperatures and pressures, thermal measurements from various engines and alternator locations, etc. [16]. The prototype system demonstrated the efficiency of applying agent based technologies to this problem.

In 2006 Sanz-Bobi, Garcia and del Pico [17] presented an Intelligent System for Predictive Maintenance (SIMAP) applied to the health condition of a wind turbine gearbox. Their system makes use of backpropagating artificial NNs to set up normal behaviour models and to detect abnormal signal behaviour. Like the MAS [16] it uses a number of different modules

(agents) to evaluate the overall component condition. The system makes use of already available signals as suggested by Verbruggen [15], e.g. component temperatures or digital operational signals. They show the ability of their system to not only detect faults like bearing defect in advance, but also to predict the remaining lifetime of the component in the diagnosis expert module. This module scans the data for certain patterns recorded during previous faults.

One year later (2007) Zaher and McArthur [18] applied the idea of a MAS [16] for condition monitoring purposes to wind turbine signals. The development focused on processing and analysing signals of the three main components gearbox, generator and rotor blades.

The development consists of five agents, detecting temperature, power performance and vibration anomalies and a downtime classifier. The downtime classifier agent serves two main purposes. Its first use is essentially the output of its main data analysis operation. It is activated upon receiving alarms from the temperature anomaly detection agent. Once an alarm is received it identifies downtime periods by monitoring the active power output trends of the turbine. If a period of downtime is identified, an attempt to provide some form of justification for each period found can be established through the information provided in the alarms and also by analyzing the associated temperature trends of the components. The second use for the output is to build a probability distribution of the various downtime periods and utilise this information for developing a model of the turbine's operational behaviour [18].

The power curve agent employs a comparison of the actual measured power curve to an averaged power curve of a number of healthy turbines.

The temperature anomaly detection agent makes use of normal behaviour models. This is achieved through the use of back propagation NNs and the module is thus similar to the anomaly detection module developed by Sanz-Bobi, Garcia and del Pico [17]. Additionally to the SIMAP system, it is also scheduled to make use of vibration signals and an additional processing unit installed on the turbine. Details about which methods will be used for the vibration signals are not given.

The MAS under development was not tested on real data by the time the article was published.

1.3. Goals, originality and thesis structure

The literature review revealed many research activities in the recent years, pointing in the direction of a global CMS to evaluate wind turbine performance and its components conditions. In particular, for single wind turbine components a large variety of monitoring methods is already available. These methods are either exclusively developed for wind turbine application or are taken from other applications and adopted to the special wind turbine characteristics such as the dynamic load situation or variable speed operation.

A number of research activities were undertaken to combine different measurements to gain an overall assessment of the system condition. They mainly differ in the input data used, required knowledge of the component design parameters as well as knowledge of how a fault manifests itself in the data. While the system developed by the Caselitz et.al [11] requires a high level of prior knowledge, SIMAP [17] is based on training algorithms and thus reduces the amount of knowledge required.

SIMAP is among the most advanced systems towards predictive and condition based maintenance schemes. The system is still under development and current research activities focus on the improvement of the link between health condition of the wind turbine and its required maintenance. Several methods for characterization of a "virtual stress" of the wind turbine for different failure modes are currently under investigation. In addition, efficiency of the executed maintenance actions is a main objective of the investigation.

The MAS presented by Zaher and McArthur [16] is intended to also consider other data sources such as vibration measurements. Their activities focus on the development of a MAS able to operate in the absence of detailed component knowledge.

1.3.1. Goals

In this Ph.D. project the focus lies on performing further research to combine the advances made by Sanz-Bobi, Garcia and del Pico [17] and Zaher and McArthur [16] with regards to the number of components monitored via a normal a behavior model approach. Furthermore, the output is combined with other information sources such as vibration measurements of the main components to arrive at a global CMS for wind turbines. Hence, the following goals can be summarized:

- Research of algorithms reducing the amount of prior knowledge required, especially for small components (e.g. pitch or azimuth motors)
- Development of modules analyzing all relevant wind turbine signals such as temperatures, accelerations, currents, voltages, speeds, power output, etc. operating in the absence of design parameters and prior knowledge
- Identification and validation of state of the art methods for vibration data analysis suitable for wind turbine applications
- Research to combine the information from the different sources forming a global condition monitoring system for wind turbines
- Research of a knowledge database to automate analysis
- Validation of the CMS using real world data and defects

Information combination is not trivial, as the domains, underlying processes, sample frequencies and measurement availabilities are quite different. For large-scale application in industry, it is of utmost importance that the algorithms and the CMS itself are highly efficient in terms of required manual analysis effort. The desired algorithms should therefore incorporate a high degree of automation to be cost effective.

1.3.2. Originality

The main original contribution of the research lies in the evaluation of different algorithms for wind turbine condition monitoring using real measured data from modern pitch regulated turbines. Efficient algorithms are identified and methods proposed to monitor the broadly available SCADA data, building the basis for monitoring not only new, but also older existing turbines with a high degree of automation.

Moreover recently developed promising methods to monitor machine vibrations (in particular bearings) are surveyed, tested and validated using real measured wind turbine gearbox bearing defects. Especially the validation of these algorithms at the industrial level is of interest, due to the novelty of these algorithms. Using recommendations found in literature for semi-automated bearing diagnosis, research is performed to fully automate bearing diagnosis.

Based on fuzzy logic an expert system is build, giving the expert/ operator a large degree of freedom when setting up rules to automate the condition evaluation. The novelty here is that the fuzzy logic based expert system can handle different types of information in a way that is general enough to allow rules to be applied to other turbines of the same type, but treating each turbine with its individual characteristics. This is required, since in practice wind turbine manufactures are allowed to supply the wind turbines equipped with components from different sub suppliers, giving each turbine its individual characteristic. Hence, although the turbines may be of the same type, their behavior during fault may differ.

Moreover, the CMS is validated by application to two wind power plants with each turbine rating two MW or more, i.e. state of technology type turbines.

1.3.3. Thesis structure

The thesis is based on international scientific journal papers. Three of them have already been published and two are currently under review. However, in this thesis chapters and sections are added to set the papers into context and supply additional information important for the description and the understanding of the global CMS.

The thesis is structured according to information source considered, i.e. SCADA data and machine vibration data.

Chapter 1 gives a general introduction, together with a literature review to describe the state of the art in wind turbine condition monitoring at the time the project was initiated. Moreover, the available data sets and information sources are briefly described in section 1.4.

Chapter 2 describes the general system setup to emphasize the available information sources and their interactions in terms of the global CMS for wind turbines.

Chapter 3 summarizes the research results to analyze SCADA data. Here four papers are presented, which are published in this context.

Chapter 4 describes the research performed to analyze the available vibration data. This chapter is further divided according to the components monitored, since each component type requires different algorithms. Namely the component types are:

- Bearings,
- Gears, and
- Rotor

The section about bearing analysis algorithms is mostly covered by a paper, but additional validation of the algorithms is supplied in an additional section, which was not published in

the original paper. The algorithms identified to monitor the component type gears are not covered in a paper due to lag of validation cases. This section gives an overview about the different state of the art algorithms implemented to monitor this type of component. It is included for completeness only as the parameters identified are further treated by the global fuzzy expert system described in chapter 5. The performance of the identified algorithms is briefly compared and discussed using a case, where false brinelling is present in the spur gear stage of a WTG gearbox.

The algorithm applied to monitor the component rotor is of simple nature and its validation was not possible. Hence, no separate section is dedicated to this component, instead a few comments about the used algorithm are made.

Chapter 5 gives an inside to the algorithms and methods used to build the global fuzzy expert system, linking all available information in terms of fuzzy logic rules. Part of this system is already described in a paper presented in chapter 3 and the paper of chapter 4, but the description in the papers is isolated concerning the information source treated. Hence, the global fuzzy expert system, being a combination of both described rule based expert systems is described in more detail. Three application examples are given to emphasize the CMS working principle.

Chapter 6 contains the main conclusion drawn from research and a discussion of future aspects.

1.4. Available data

For research, a large amount of data is available from two power plants at different locations. Table 1.1 summarizes the basic power plant information.

Table 1.1: Basic power plant information

	Power plant 1	Power plant 2
Location	Onshore	Offshore
Number of turbines	18	21
Rated power / turbine	2 MW	2.3 MW
Commissioning date	April 2009	April 2011
Pitch regulation	Hydraulic	Hydraulic
Generator type	Doubly fed	Synchronous
SCADA measurements	Yes (191370 10 min avg. values per signal)	Yes (88405 10 min avg. values per signal)
Vibration measurements	No	Yes

For power plant 1 about 191370 10-minute average values per signal are available for research. Due to the later commission, this figure is 88405 10-minute average values per signal for power plant 2. More detailed information about the sensor setup as well as a list of sensors used for condition monitoring of power plant 1 can be found in the paper: "Wind turbine condition monitoring based on SCADA data - Part I System description" at the paper section 3.1 (thesis page 45). In a similar fashion, the available SCADA data of power plant 2 is summarized in Table 1.2 and a turbine drawing to visualize the sensor and component positions is depicted in Figure 1.2.

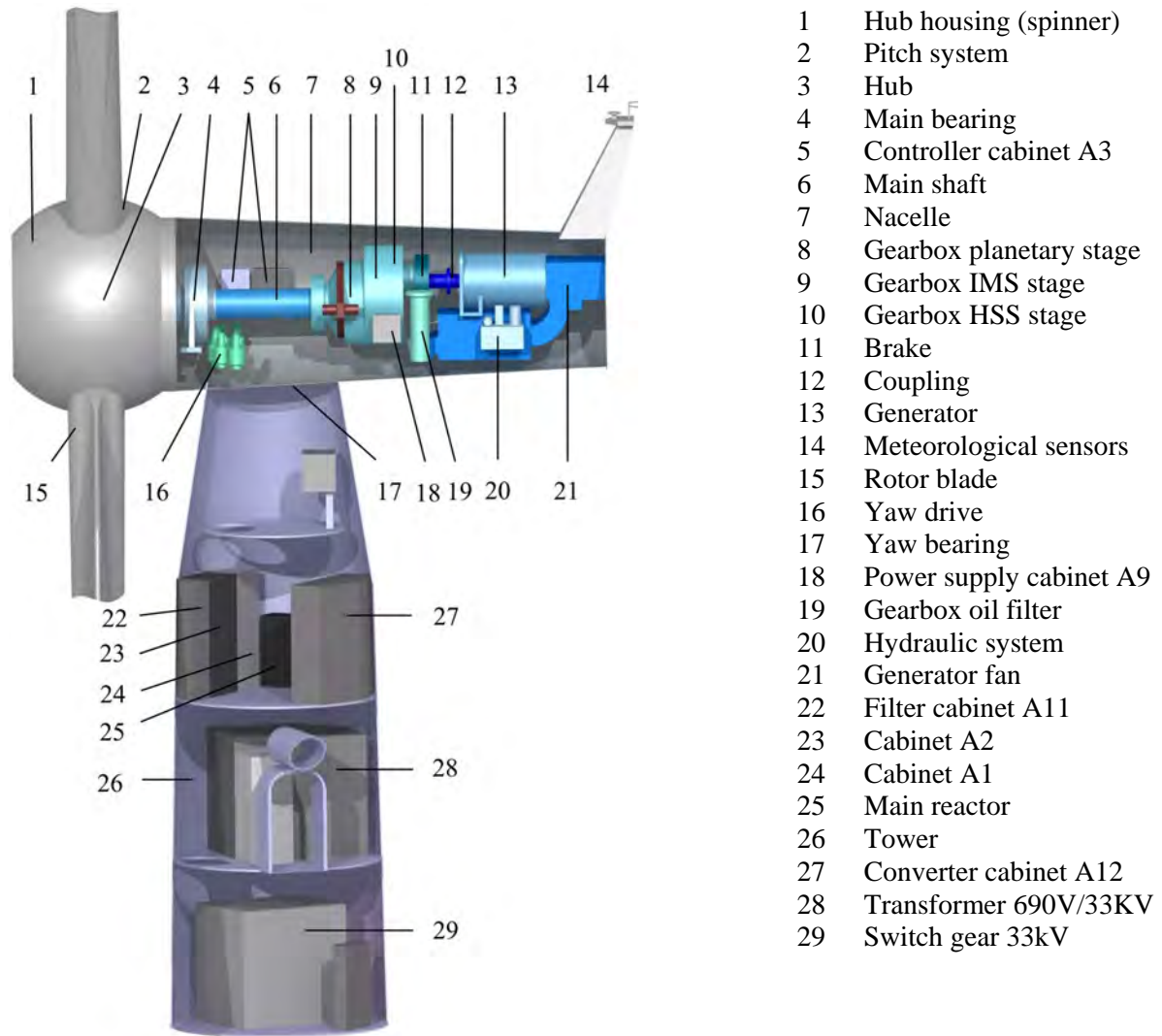


Figure 1.2: Turbine main components (power plant 2)

Table 1.2: SCADA data signal list of power plant 2

Name of variable	Unit	Sensor pos. (Fig.1.2)	Short description	NBM
Cabinet A1 ext. temp.	°C	24	Temp. outside cabinet A1	Yes
Cabinet A1 int. temp.	°C	24	Temp. inside cabinet A1	Yes
Cabinet A2 int. temp.	°C	23	Temp. inside cabinet A2	Yes
Cabinet A3 int. temp. left	°C	5	Temp. inside cabinet A3 at the left	Yes
Cabinet A3 int. temp. right	°C	5	Temp. inside cabinet A3 at the right	Yes
Cabinet A9 int. temp.	°C	18	Temp. inside cabinet A0	Yes
Ambient temp.	°C	14	Ambient temp.	Yes
Hub temp.	°C	3	Spinner temp. (in hub housing)	Yes
Pitch angle A	°	2	Blade pitch angle blade A	Yes
Pitch angle B	°	2	Blade pitch angle blade B	Yes
Pitch angle C	°	2	Blade pitch angle blade C	Yes
Pitch oil pres. A	bar	2	Oil pressure at pitch blade A	Yes
Pitch oil pres. B	bar	2	Oil pressure at pitch blade B	Yes

Name of variable	Unit	Sensor pos. (Fig.1.2)	Short description	NBM
Pitch oil pres. C	bar	2	Oil pressure at pitch blade C	Yes
Pitch oil bypass on time	s	-	Time pitch oil bypass on	No
Main bearing temp.	°C	4	Main bearing temp. (rotor support)	Yes
Rotor speed	rpm	4	Rotor speed	Yes
Hydraulic oil pres. pump	bar	20	Hydraulic oil pressure at pump	Yes
Hydraulic oil temp.	°C	20	Hydraulic oil temp.	Yes
Gearbox oil temp.	°C	8	Gearbox oil sump temp.	Yes
Gearbox oil vent on low	s	-	Time gearbox oil fan on (low)	No
Gearbox oil vent on high	s	-	Time gearbox oil fan on (high)	No
Gearbox bearing temp. HSS gen.	°C	10	Gearbox bearing temp. HSS (generator side)	Yes
Gearbox bearing temp. HSS rot.	°C	10	Gearbox bearing temp. HSS (rotor side)	Yes
Gearbox bearing temp. IMS gen.	°C	9	Gearbox bearing temp. IMS (generator side)	Yes
Gearbox bearing temp. IMS rot.	°C	9	Gearbox bearing temp. HSS (rotor side)	Yes
Gearbox oil pres.	bar	19	Gearbox oil pressure	Yes
Gearbox oil pres. after inline filter	bar	19	Gearbox oil pressure after filter	Yes
Gearbox oil pres. before inline filter	bar	19	Gearbox oil pressure before filter	Yes
Gearbox oil pres. diff. offline filter	bar	-	Gearbox oil pressure difference	Yes
Brake temp. gear	°C	11	Brake temp. (at the HSS)	Yes
Brake oil pres.	bar	11	Brake pressure	No
Generator stator temp. R	°C	13	Generator stator temp. phase R	Yes
Generator stator temp. S	°C	13	Generator stator temp. phase S	Yes
Generator stator temp. T	°C	13	Generator stator temp. phase T	Yes
Generator bearing temp. NDE	°C	13	Generator bearing temp. NDE	Yes
Generator bearing temp. DE	°C	13	Generator bearing temp. DE	Yes
Generator speed	rpm	13	Generator speed	Yes
Nacelle temp.	°C	5	Temp. in machine housing	Yes
Converter cool water temp.	°C	27	Converter cooling water temp.	Yes
Main reactor temp. R	°C	25	Converter reactor temp. phase R	Yes
Main reactor temp. S	°C	25	Converter reactor temp. phase S	Yes
Main reactor temp. T	°C	25	Converter reactor temp. phase T	Yes
Grid filter temp. B1	°C	22	Grid filter temp. 1 (used to filter ripple harmonics)	Yes
Grid filter temp. B2	°C	22	Grid filter temp. 1 (used to filter ripple harmonics)	Yes
Grid filter temp. B3	°C	22	Grid filter temp. 1 (used to filter ripple harmonics)	Yes
IO module 1 temp.	°C	23	IO module temp. 1	Yes
IO module 2 temp.	°C	23	IO module temp. 2	Yes
IO module 3 temp.	°C	5	IO module temp. 3	Yes
Grid module board temp.	°C	-	Temp. main board grid module	No
Voltage R	V	24	Voltage phase R	Yes
Voltage S	V	24	Voltage phase S	Yes
Voltage T	V	24	Voltage phase T	Yes
Power output	kW	24	Turbine power output	Yes
Current R	A	24	Current phase R	Yes
Current S	A	24	Current phase S	Yes
Current T	A	24	Current phase T	Yes
Grid frequency	Hz	24	Grid frequency	No

Name of variable	Unit	Sensor pos. (Fig.1.2)	Short description	NBM
Reactive power	kVAr	24	Reactive power	No
Wind speed 1	m/s	14	Wind speed measurement 1	Yes
Wind speed 2	m/s	14	Wind speed measurement 2	Yes
Nacelle position angle	°	17	Nacelle position angle	No
Humidity tower	%	26	Relative humidity in tower	Yes
Transformer oil temp. filtered	°C	29	Transformer oil temp. after filter	Yes
Transformer room temp.	°C	29	Transformer room temp.	Yes

At power plant 2, also vibration measurements of the main components are available. The data acquisition and processing system (M-system) periodically records time series from the different sensor positions visible in Figure 1.3.

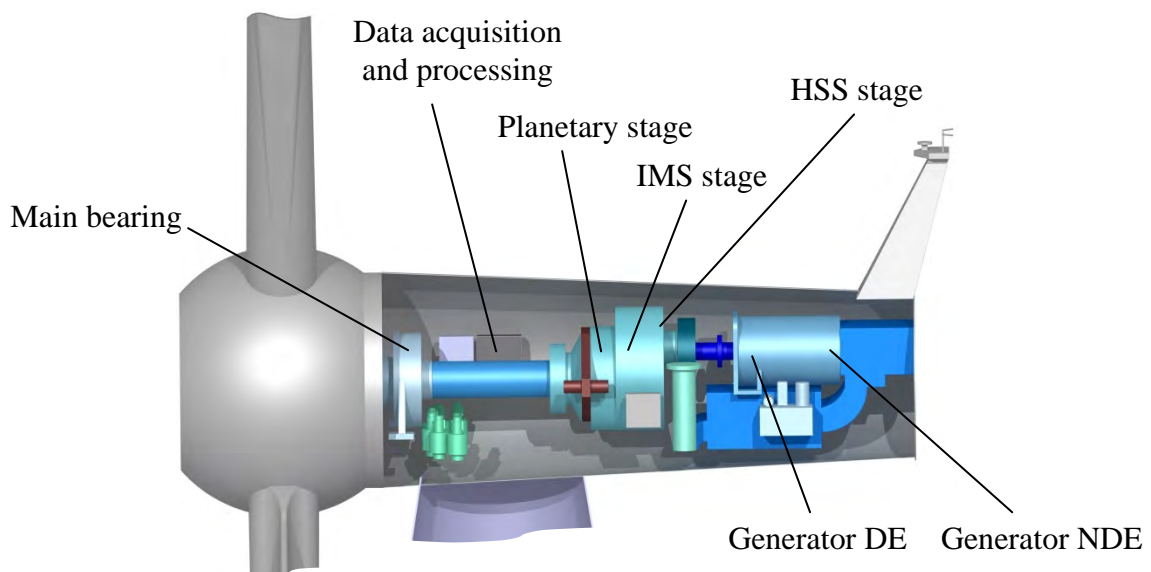


Figure 1.3: Accelerometer positions

In the M-system, various measurements and analysis can be specified, each with a specific sampling frequency and sample length as well as recording or analysis condition (e.g. turbine power output). Data analysis is frequently performed by the unit to allow fast response to significant changes in condition, e.g. fast propagating tooth cracks. In those cases, direct access to the turbine controller allows turbine shutdown. However, raw data are stored only every 14 days and only of the full load bin, i.e. when the turbine power output is between 2.18 and 2.42 MW. If the trigger runs off, and the turbine is not operating in this power range, data is collected as soon as this condition is matched. A new period begins, once data is collected. This causes the time between measurements to be 14 days or more.

In this thesis, only raw data are processed in order to evaluate state of the art algorithms. Table 1.3 supplies an overview about the available measurements for each sensor position, together with the corresponding sampling frequency and sample length.

Table 1.3: Vibration measurement overview

Name	Sampling freq. [kHz]	Sample length [s]	Num. of samples	Num. of meas. per turbine ²⁾	Main bearing	Planetary stage	IMS stage	HSS stage	Generator DE	Generator NDE
62_Tr	0.159	100.79 s	16000	26	x					
127	0.320	100.00 s	32000	35		x				
127_Tr	0.320	100.00 s	32000	36		x				
250_Tr	0.640	93.78 s	60000	26	x					
300_Tr	0.768	41.66 s	32000	35			x	x		
1000	2.558	25.06 s	64100	38	x	x	x	x	x	x
1000_Tr	2.558	25.06 s	64100	39	x	x	x	x	x	x
6000_Tr	15.385	4.17 s	64100	39	x	x	x	x	x	x
16000	41.667	2.40 s	64100 ¹⁾	39	x	x	x	x	x	x

¹⁾ Number of samples for main bearing 100000 instead of 64100; ²⁾ Since commissioning

The notation “Tr” at the end of the time series name indicates tracking. By tracking speed variations are encountered by resampling the signal synchronous to the shaft speed. Here these data are resampled to the nominal speed. The resampling is done autonomously by the M-system. The necessary speed encoder information for resampling is not stored. The time between the record values is no longer constant for these time series, which is why the stated sampling frequency in Table 1.3 can be interpreted the average sampling frequency of this time series.

2 General system setup

2.1. Description

In this chapter, the general CMS concept is briefly described in order to further outline the thesis structure and sharpen understanding of the sub sequential chapters and the CMS in general. In Figure 2.1, the general CMS concept is illustrated.

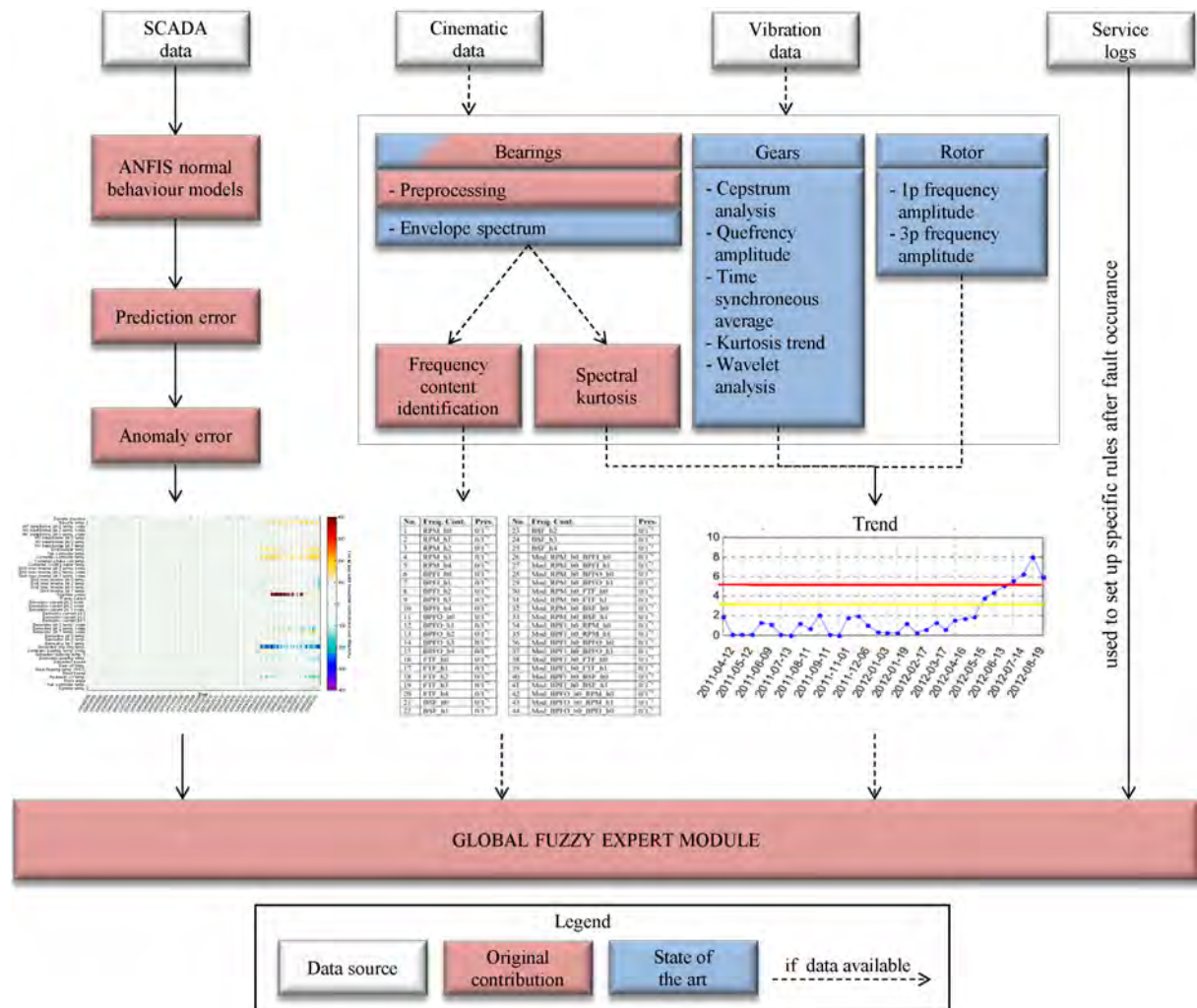


Figure 2.1: General CMS concept

Next to the measurements (SCADA and vibration data), also kinematic data as well as service logs are available as information sources. Due to their different nature and time scale, SCADA data must be treated in a different way than vibration data. Never the less this is a

consequence of the broad variety of signals available in the SCADA data sets (see Table 1.2), with different signal characteristics. For this reason, each data type has to be processed using individual algorithms.

Despite the used processing technique, the available information is condensed to a level that allows analysis by the global fuzzy expert system. In this sense, the output of each processing technique is the information if the signal behaves as expected or if an anomaly is present in the data. This is achieved by learning the signals normal behavior using a set of reference values (from a defect free period), to distinguish between normal and abnormal or calculate the parameter statistics.

For vibration measurements, diagnostic information is also contained in the frequency content. Pin pointing the faulty component is often possible if a characteristic defect frequency occurs in the spectrum. For this reason the kinematic data i.e. typical defect and meshing frequencies originating from the system kinematics is another important information source aiding diagnosis in the global fuzzy expert module.

At the level of the global fuzzy expert module, the signal behavior is independent of the information root categorizable via the linguistic terms:

- Very low
- Low
- Normal
- High
- Very high
- Present
- Not present

The latter two are currently only used for identification of the frequency content (defect frequencies). The rules build with the linguistic terms are defined by an expert, who may consider service logs containing useful information about the fault root cause or performed investigations. After rule establishment, the fault can be identified and prioritized automatically, next time of fault occurrence.

3 SCADA Data analysis

This chapter is based on four journal papers documenting the research performed to systematically utilize the information contained in the SCADA data, and its usage for condition monitoring.

3.1. Publication [P1] (published)

From the initial description of the state of the art, it becomes obvious that different methods are described in literature how SCADA data can be processed to extract information for condition monitoring. For large-scale practical application, the methods used must be robust, handle the individual turbine characteristics and show abnormalities in signal behavior clearly. A frequently found method in literature (e.g. [14] [19] [20] [21] [17] [22] [23]) is building Normal Behavior Models (NBMs) from data with NNs. This promising method even works in the absence of prior knowledge about the signal behavior, since it is purely data mining based.

However, the major drawbacks of NNs are their black-box data processing structure and slow convergence speed [24]. NN globally fit a single function into the data and thereby losing insight into the problem [25]. By using regression based methods to build the NBM this insight is maintained. Therefore research is carried out, to identify the benefits of using NNs in terms of fault detection capabilities in comparison to rather traditional regression based approaches.



Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/jnlabr/ymssp

Review

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

Meik Schlechtingen, Ilmar Ferreira Santos*

Department of Mechanical Engineering, Section of Solid Mechanics, Technical University of Denmark, Denmark

ARTICLE INFO

Article history:

Received 3 March 2010

Received in revised form

2 December 2010

Accepted 11 December 2010

Available online 24 December 2010

Keywords:

Condition monitoring

Neural networks

SCADA data

Fault detection

Wind turbine

Normal behavior models

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, which 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 led to the highest confidence level.

© 2010 Elsevier Ltd. All rights reserved.

Contents

1. Introduction	1850
2. Time series containing anomalies	1851
2.1. Generator bearing anomaly	1851
2.2. Gearbox bearing temperature anomaly I	1852
2.3. Gearbox bearing temperature anomaly II	1854
2.4. Generator stator temperature anomaly I	1854
2.5. Generator stator temperature anomaly II	1854
3. Regression model development	1855
4. Neural network model setup and training	1858
4.1. Input signals	1858

* Corresponding author. Tel.: +45 45256269; fax: +45 45931475.
E-mail address: ifs@mek.dtu.dk (I. Ferreira Santos).

4.2.	Network type and transfer/activation function.....	1858
4.3.	Input data pre-processing and validity check.....	1858
4.4.	Number of training patterns	1859
4.5.	Network structure	1859
4.6.	Training method.....	1860
4.7.	Weight initialization	1860
4.8.	FSRC neural network model	1860
4.9.	Autoregressive neural network model	1861
5.	Fault identification and comparison	1862
5.1.	Regression model.....	1862
5.2.	FSRC neural network model	1863
5.3.	Autoregressive neural network model	1864
5.4.	Comparison	1865
6.	Further analysis results.....	1866
6.1.	Gearbox bearing damage I.....	1866
6.1.1.	Model development.....	1866
6.1.2.	FSRC neural network model.....	1867
6.1.3.	Autoregressive neural network model.....	1867
6.1.4.	Comparison.....	1868
6.2.	Gearbox bearing damage II	1869
6.2.1.	Model development.....	1869
6.2.2.	FSRC neural network model.....	1869
6.2.3.	Autoregressive neural network model.....	1869
6.2.4.	Comparison.....	1869
6.3.	Generator stator temperature anomaly I.....	1870
6.3.1.	Model development.....	1870
6.3.2.	FSRC neural network model.....	1871
6.3.3.	Autoregressive neural network model.....	1871
6.3.4.	Comparison.....	1872
6.4.	Generator stator temperature anomaly II	1872
6.4.1.	Model development.....	1872
6.4.2.	FSRC neural network model.....	1872
6.4.3.	Autoregressive neural network model.....	1873
6.4.4.	Comparison.....	1874
7.	Conclusion	1874
	References	1875

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 often does 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 on 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. Here 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 those 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 are 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 through 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 presented here.

In 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 specifically occurring faults, laying the basis for statements on 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 2 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 high accuracy and 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 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 low autocorrelation, e.g. the turbine power output, and are thus 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 how the NN is set up and how it is trained to achieve a good performance is briefly described. In Sections 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 in Fig. 1.

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. The basis of this analysis is a time series containing two catastrophic generator bearing damages of a 2 MW offshore wind turbine shown in Fig. 2.

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 NNs and the regression model development the period after the first bearing damage shown in Fig. 3 is used.

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.

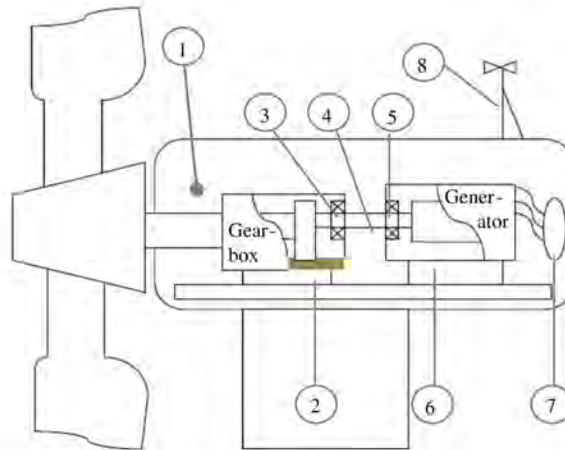


Fig. 1. Schematic of sensors in a wind turbine [6]: (1) nacelle temperature, (2) gearbox oil sump temperature, (3) gearbox bearing temperature, (4) high speed shaft revolution speed, (5) generator bearing temperature, (6) stator temperature (3 phases), (7) power output and (8) ambient temperature.

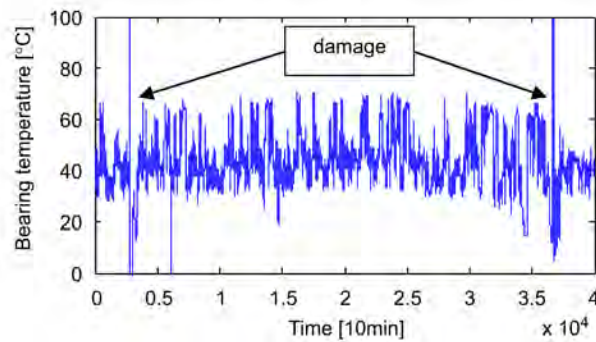


Fig. 2. Time series containing two generator bearing damages measured on WTG A.

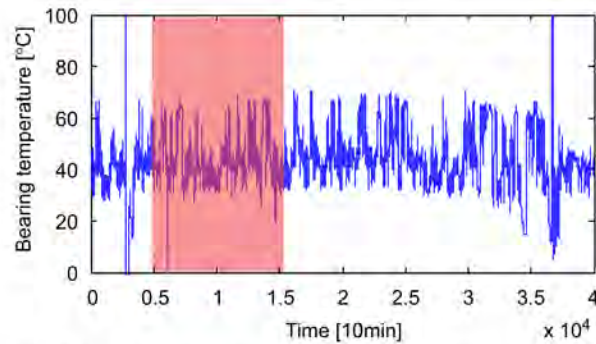


Fig. 3. Period used for training/regression model development WTG A.

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 Fig. 4 illustrates the gearbox high speed shaft bearing temperature measured at another turbine.

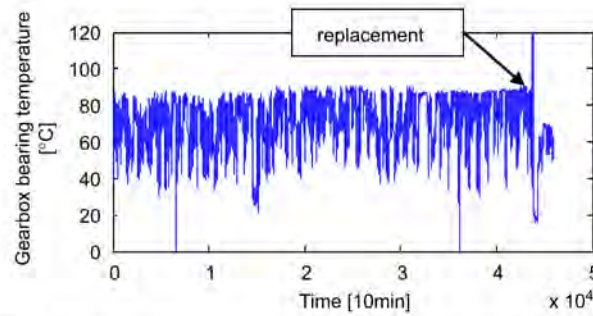


Fig. 4. Time series containing a gearbox bearing damage measured on WTG B.

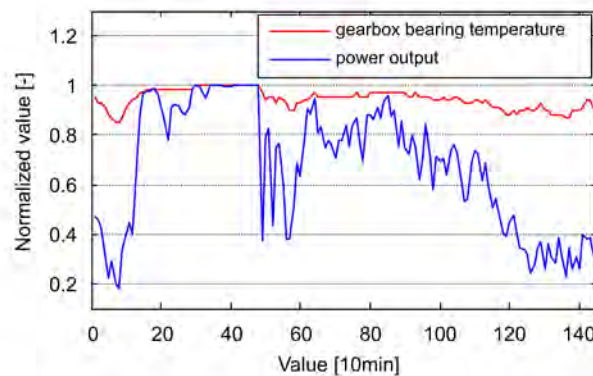


Fig. 5. Exemplary power and gearbox bearing temperature fluctuations over time.

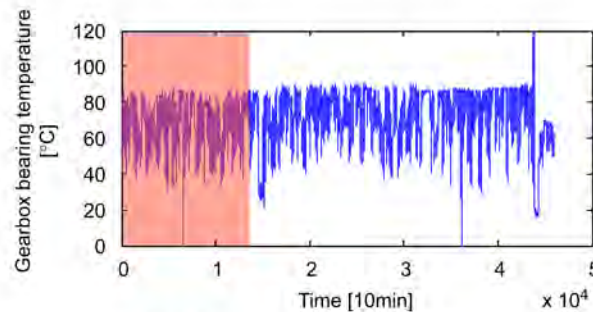


Fig. 6. Period used for training WTG B.

Although the time series in Fig. 4 shows a gradual temperature increase over time, identification of temperature increase based on thresholds is possible only 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 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. Fig. 5 shows an example of the power fluctuations over time and the related gearbox bearing temperature.

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

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 on 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 Fig. 7.

The time series shows a trend in maximum temperature measured right at the beginning of the time series. After 20,000 values the maximum temperature measured stabilizes. 7000 values prior to the bearing damage the maximum power output is restricted to 50% of the nominal power that led to a drop in maximum temperature level.

Although the time series shows a trend in the first 5 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 20,000 values would not leave enough data for fault analysis. Therefore the influence on performance is accepted, due to lag of earlier data being accessible. The training period is highlighted in Fig. 8.

All data currently available for this research cover a period of 14 months recorded after 2–3 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 other types of faults are also investigated in this research. Fig. 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, even if the turbine and the grid are okay and the wind speed is above the cut in wind speed.

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 as shown in Fig. 10.

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 Fig. 11.

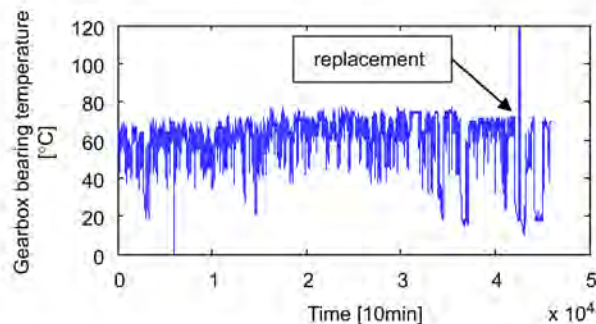


Fig. 7. Time series containing a gearbox bearing damage measured on WTG C.

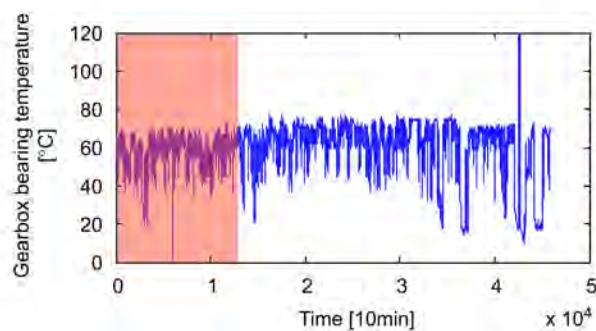


Fig. 8. Period used for training WTG C.

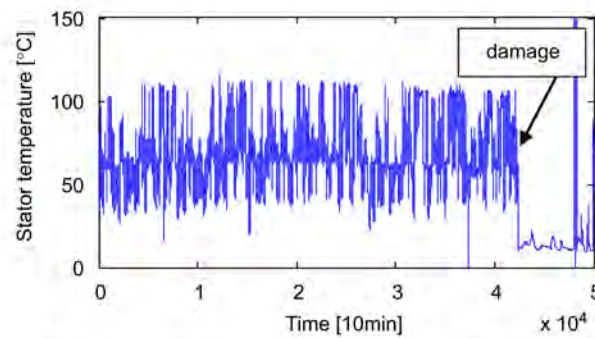


Fig. 9. Time series of the stator temperature measured on WTG D.

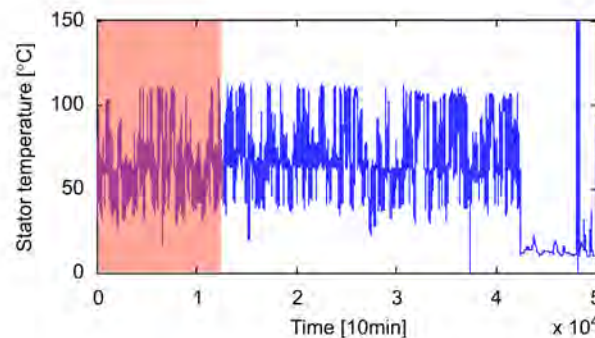


Fig. 10. Period used for training WTG D.

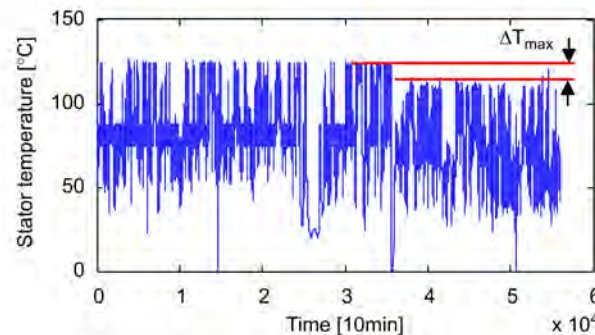


Fig. 11. Time series of the stator temperature measured on WTG E.

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 finally cannot be proved by the data supplied.

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

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 with the desired signal can be found. The actual lag removal is an important step to reduce scatter and improve prediction.

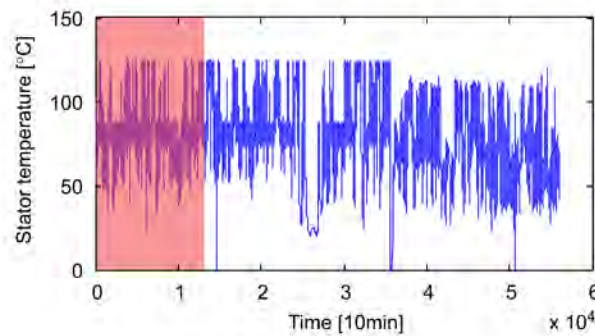


Fig. 12. Period used for training WTG E.

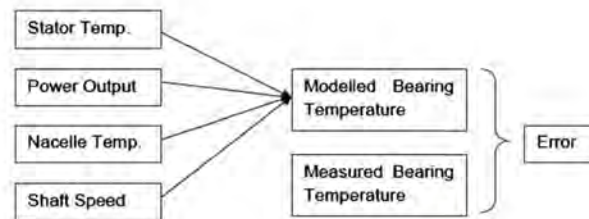


Fig. 13. Regression model schematic [6].

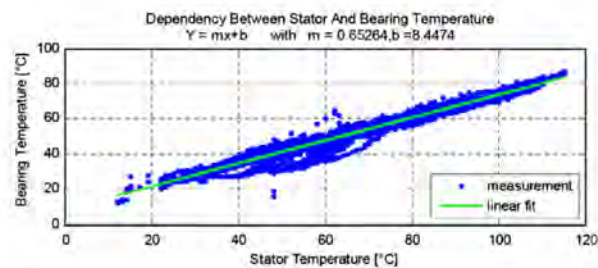


Fig. 14. Correlation between the stator and the bearing temperature [6].

The identified correlated signals can then be used to build the regression model as shown in Fig. 13.

The stator temperature was found to be especially highly correlated to the generator bearing temperature as is visible in Fig. 14.

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 Fig. 15:

Apart from the few outliers, the regression model is capable of predicting the bearing temperature in the range of $\pm 5^\circ\text{C}$. The threshold value is five times the standard deviation. The remaining outliers are caused by extreme transient situations such as turbine runup or shutdown. 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 the 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 at a rate that is proportional to the area of contact between the body and the environment [7].

The high ratio of stator heat capacity to 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

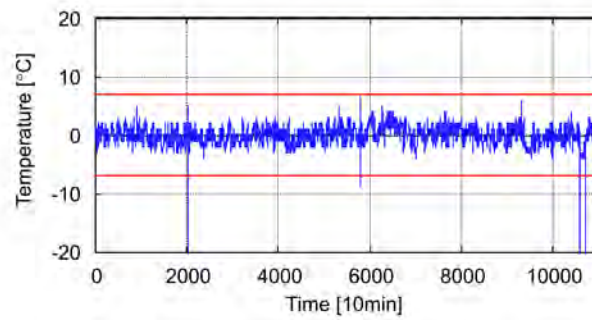


Fig. 15. Estimation error for the full regression model.

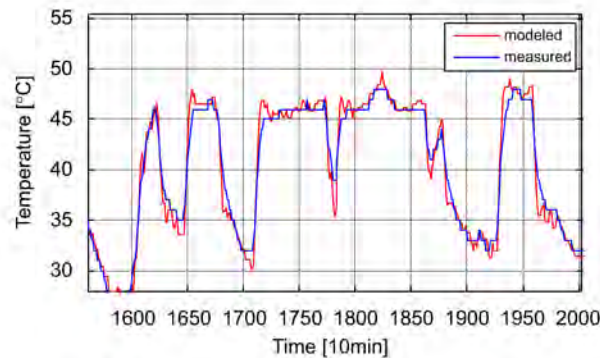


Fig. 16. Estimated and measured bearing temperature [6].

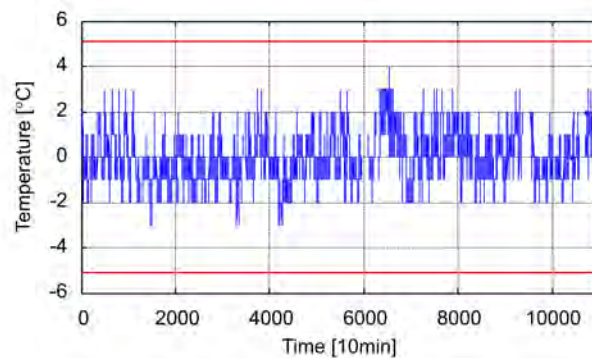


Fig. 17. Bearing temperature estimation error.

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 true 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 Fig. 16.

By the 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 is shown in Fig. 17. The prediction error is finally down to ± 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 the 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 are 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 [8–11].

Generally one of the major concerns is to identify the suitable network setup 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. 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 so as to independently reconstruct the signal. The model input signals are (see Fig. 1):

- el. power output ($t - \text{lag}$) {pos.7},
- nacelle temperature ($t - \text{lag}$) {pos.1},
- generator stator temperature ($t - \text{lag}$) {pos.6},
- generator speed ($t - \text{lag}$) {pos.4},

The second is to identify changes in the autoregression 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 ($t - \text{lag}$) {pos.7},
- nacelle temperature ($t - \text{lag}$) {pos.1},
- generator stator temperature ($t - \text{lag}$) {pos.6},
- generator speed ($t - \text{lag}$) {pos.4},

The autoregressive approach is used by [3,1] to identify gearbox failures. Both kinds of models are compared to the regression model.

4.2. Network type and transfer/activation function

In this research multilayer 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.

4.3. 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 particularly 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 and (d) lag removal

(a) *Validity check*—For a smoother network generalization it may sometimes be necessary to remove some of the data outliers [9]. 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 50–400 would be expected. A value of –10 or 900, is clearly wrong [11]. 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 [11].

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:

$$S = \frac{V - V_{\min}}{V_{\max} - V_{\min}} \quad (1)$$

where V is a variable and S is the normalized 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.

(d) *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 the use of 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 data set.

4.4. 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 [9]. 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 degradation from the first hour of operation.

4.5. 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 [11] 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 that gives the best generalization should be chosen. A similar method is also recommended by [9] 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 [1,3] 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 shown in Fig. 18. The network architecture for the FSRC generator bearing temperature model is similar, but has only five neurons in the hidden layer.

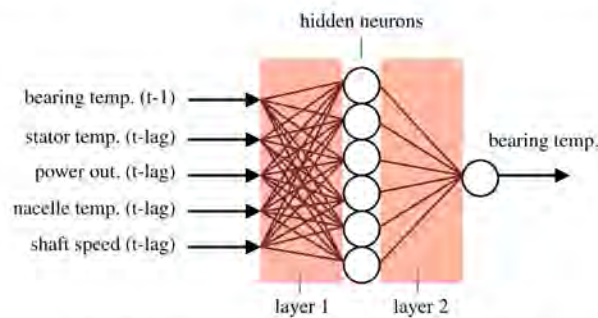


Fig. 18. Network architecture—feed forward network [6].

4.6. Training method

Generally there are two different modes of training NNs: batch mode and pattern mode (sequential learning) [11]. In [9] using batch mode is recommended 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.

4.7. 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 $+0.01$ [11]. 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 computational efficiency.

4.8. FSRC neural network model

The setup network is capable of predicting the generator bearing temperature when the stator and nacelle temperatures as well as the shaft speed and the power output are given. 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 Fig. 19 the prediction errors for the FSRC NN and the simple regression model are shown.

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 the 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 is possible for the simple regression model. In SIMAP [3] a fuzzy expert system is implemented to identify the abnormalities. The more the input signals 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 Giebhardt [12]. If more than five subsequent values exceed the limit, an alarm is triggered [12]. For the developed models in this paper, it is sufficient using a threshold of three subsequent exceedings to prevent false alarms.

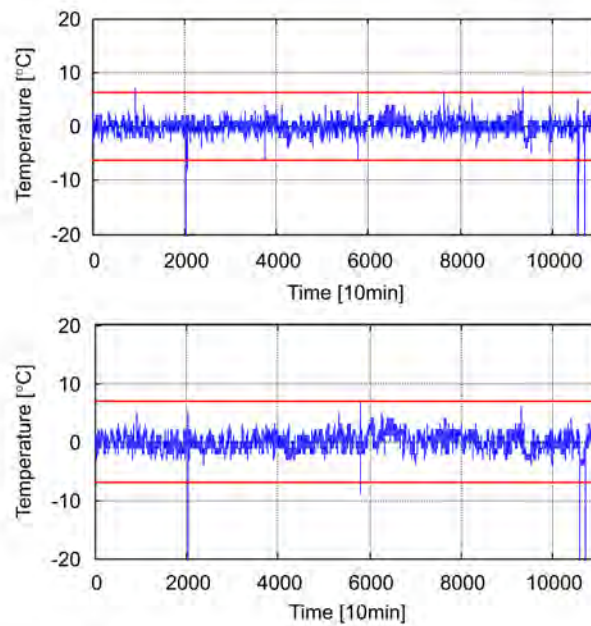


Fig. 19. Prediction error for the training data set: FSRC NN model (top) and regression model (bottom).

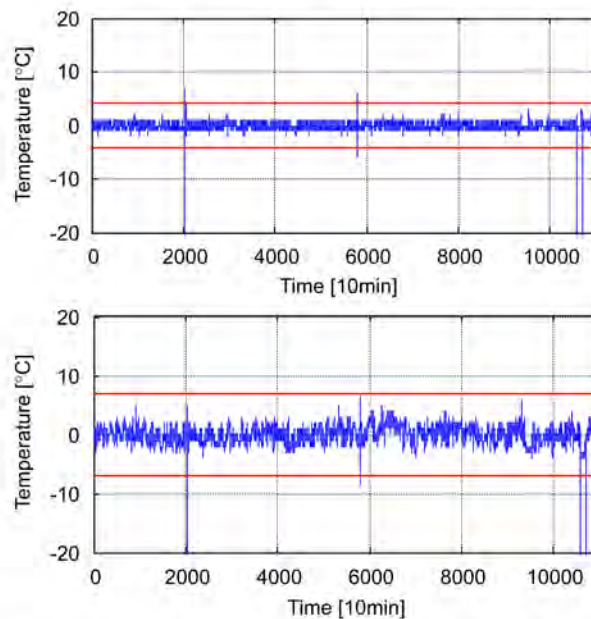


Fig. 20. Prediction error for the training data set: autoregressive NN model (top) and regression model (bottom).

4.9. 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. Fig. 20 shows the prediction error for the training data set and a comparison to the simple regression model.

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 Fig. 1 pos. 6.

The prediction error is in the range of $\pm 1^\circ\text{C}$ 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 Fig. 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. Fig. 21 shows the evolution of the prediction error from the time series exceeding the one used for model development until the second bearing damage. 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 to the damage.

Fig. 22 shows the one day average values of the prediction error for the period under consideration. The averaged signal eases the trend identification. 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.

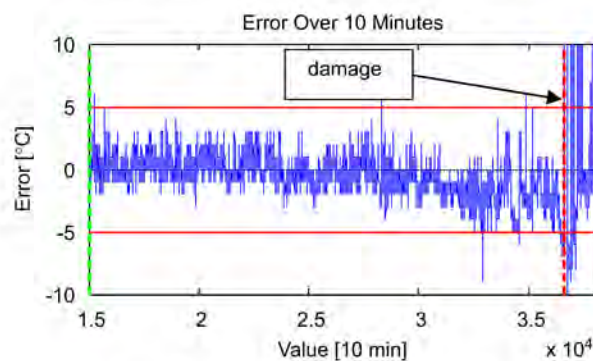


Fig. 21. Prediction error of the regression model.

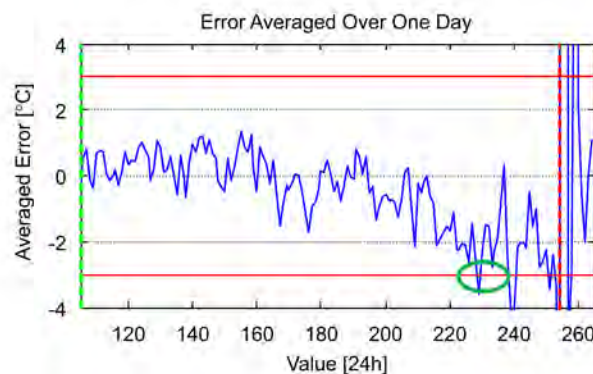


Fig. 22. Averaged prediction error of the regression model.

The amplitude increases 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 among the bearing, nacelle and ambient temperatures 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. Fig. 23 shows the one week average 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 previously unseen input signals are given. The prediction strongly depends on the weight initialization. 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.

Fig. 24 shows the prediction error development prior to the damage. Visually an anomaly can be detected, beginning 4000 values (≈ 27 days) prior to the damage.

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

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. Fig. 26 shows the long term prediction error development averaged over one week.

The seasonal effect contained in the signal is $\pm 0.3^\circ\text{C}$ in comparison with $\pm 0.5^\circ\text{C}$ 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

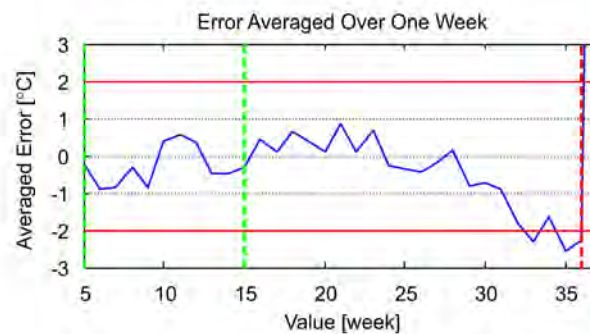


Fig. 23. Averaged prediction error including the model development period.

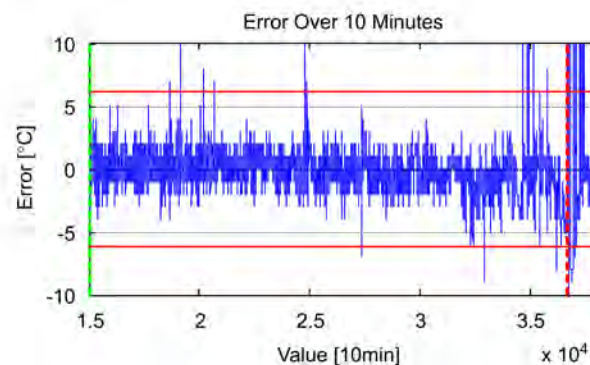


Fig. 24. Prediction error of the full signal reconstruction neural network model.

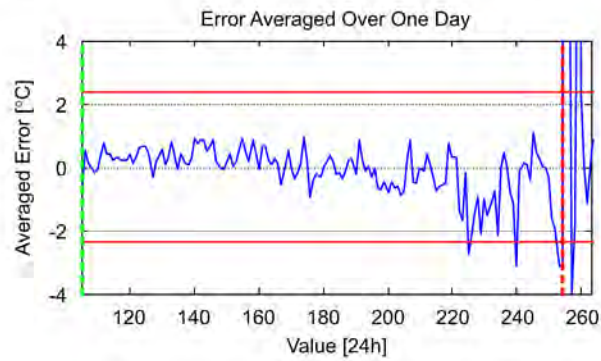


Fig. 25. Averaged prediction error of the full signal reconstruction neural network model.

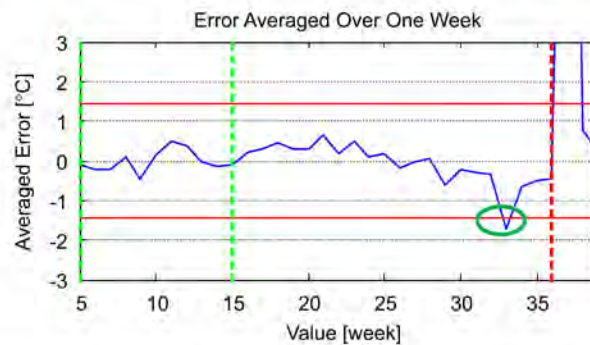


Fig. 26. Averaged prediction error of the FSRC neural network model including the training period.

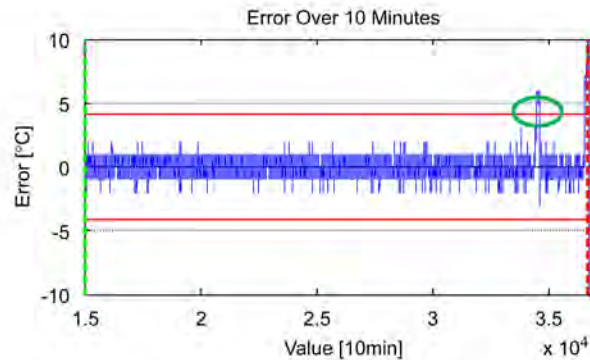


Fig. 27. Prediction error of the autoregressive neural network model.

term trend. An anomaly in the 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 (Fig. 27). This could be for instance a bearing damage, where the bearing internal heat increases. On 10 min bases, the prediction accuracy is in the range of ± 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 ± 2 °C.

Fig. 28 shows the averaged prediction error over one day. 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 to 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 Fig. 29. 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 with the other two model setups, 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. Table 1 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. 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.

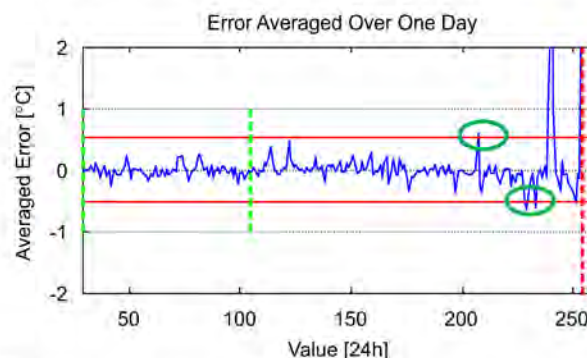


Fig. 28. Averaged prediction error of the autoregressive neural network model.

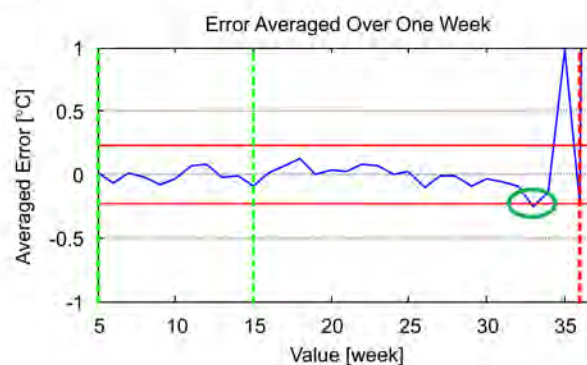


Fig. 29. Averaged prediction error of the autoregressive neural network model including the training period.

Table 1

Comparison of the fault visibility between the three different normal behavior model type approaches.

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

During regression model development, it was found that each turbine must be treated as individual. The cause can be found to be the 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 to iterate 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 complexity model that has implemented the main signal correlations and has a low correlation of the prediction error to any of the input signals.

A fully 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 setup, 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 the use of FSRC and the autoregressive NN model only. The corresponding time series are shown in Figs. 4–12.

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 Fig. 1)

- gearbox oil sump temperature ($t - \text{lag}$) {pos.2},
- high speed shaft revolution speed ($t - \text{lag}$) {pos.4},
- power output ($t - \text{lag}$) {pos.7},
- nacelle temperature ($t - \text{lag}$) {pos.1},
- ambient temperature ($t - \text{lag}$) {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 ($t - \text{lag}$) {pos.7},
- nacelle temperature ($t - \text{lag}$) {pos.1},
- ambient temperature ($t - \text{lag}$) {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 and McArthur [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. 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 Fig. 30 the model prediction error is shown for the entire time series, including the training period.

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, as illustrated in Fig. 31.

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 to 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 and McArthur [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.

Fig. 32 shows the prediction error for the autoregressive NN model including the training period. The alarm limit is $\pm 7.1\text{ }^{\circ}\text{C}$. An ALV is first visible 12,000 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 1 day average can be used (Fig. 33) to better identify the underlying trend and to increase the sensitivity by means of reduced alarm limits.

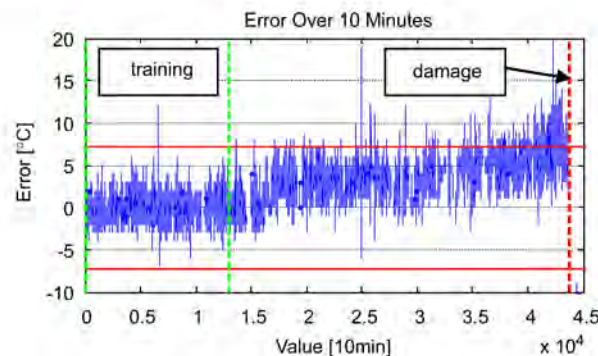


Fig. 30. Prediction error for the FSRC NN model.

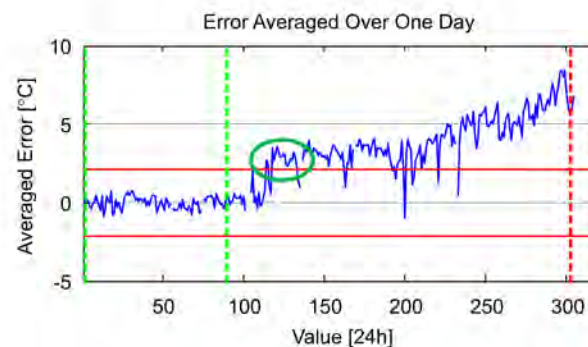


Fig. 31. Averaged prediction error for the FSRC NN model.

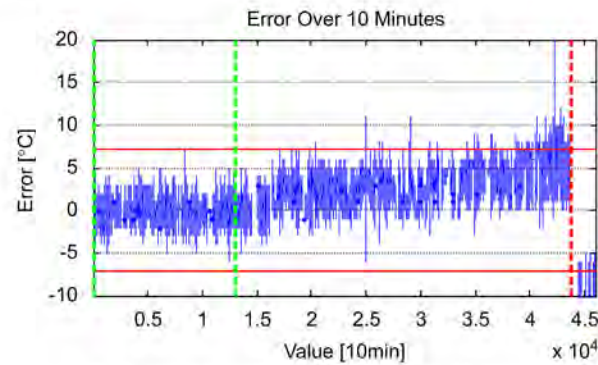


Fig. 32. Prediction error for the autoregressive NN model.

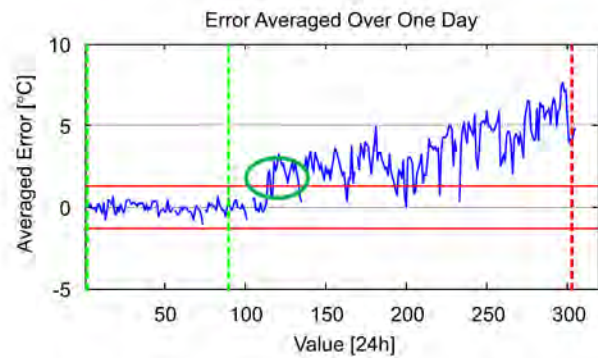


Fig. 33. Averaged prediction error for the autoregressive NN model.

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

	First ALV (days)	Second ALV/trend (days)
Full signal reconst. NN	190	184
Autoregressive NN	185	186

Due to the alarm limits being reduced to 1.3 °C, an ALV is raised 185 days prior to 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 progressing damage, the variations 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 1 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 1 day average base. Table 2 below summarizes the results of the comparison concerning anomaly detection. 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 as in Section 6.1.1.

6.2.2. FSRC neural network model

The corresponding time series showing the temperature evolution is shown in Fig. 7. Below the prediction error is illustrated on 10 min average (Fig. 34) and on 1 day average base (Fig. 35).

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 to 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 Figs. 36 and 37 below. 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 1 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 the case of fault occurrence.

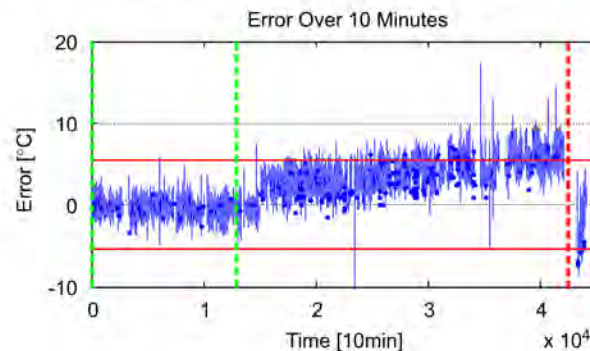


Fig. 34. Prediction error for the FSRC NN model.

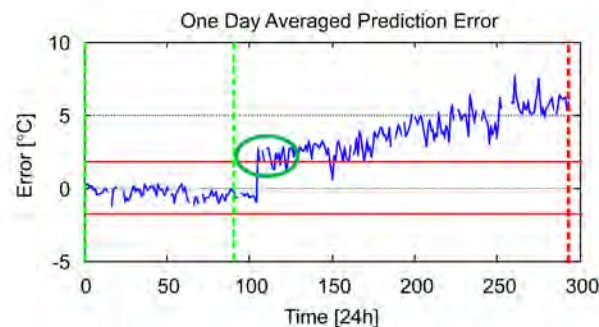


Fig. 35. Averaged prediction error for the FSRC NN model.

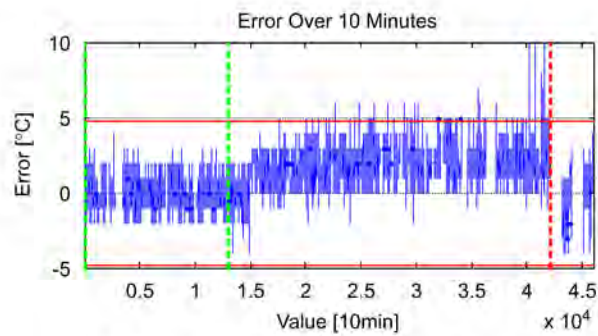


Fig. 36. Prediction error for the autoregressive NN model.

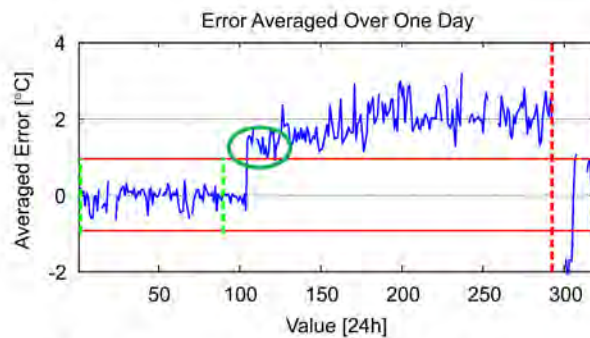


Fig. 37. Averaged prediction error for the autoregressive NN model.

Table 3

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

	First ALV (days)	Second ALV/trend (days)
Full signal reconst. NN	187	186
Autoregressive NN	187	186

Table 3 summarizes the results gained concerning anomaly detection for the two approaches. 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.

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 Fig. 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 ($t-\text{lag}$) {pos.7},
- nacelle temperature ($t-\text{lag}$) {pos.1},
- ambient temperature ($t-\text{lag}$) {pos.8}.

6.3.2. FSRC neural network model

The corresponding time series showing the generator stator temperature is shown in Fig. 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 Figs. 38 and 39. 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 4 days, leading to a second ALV be raised 1 day later. After further 4 days the prediction error falls back to the normal operational range.

Although by this time the ALV might be considered as an outlier, 5 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

The corresponding prediction errors for the autoregressive NN model are shown in Figs. 40 and 41. The autoregressive NN model shows an increase in amplitude starting 30 days after training is completed. On 1 day average base the alarm limits are down to the measurement accuracy of $\pm 1^\circ\text{C}$. 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 to the generator damage. The maximum model deviation is 2.6°C .

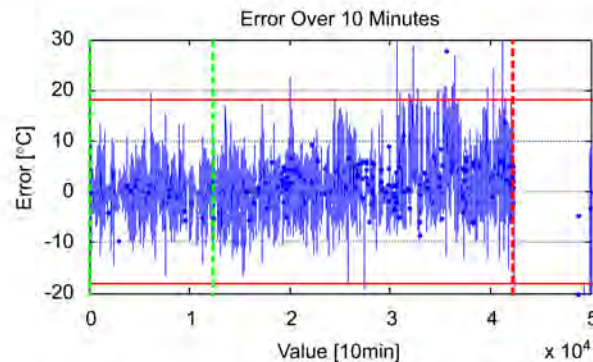


Fig. 38. Prediction error for the FSRC NN model.

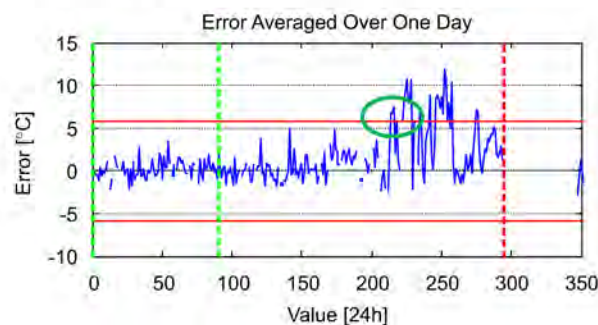


Fig. 39. Averaged prediction error for the FSRC NN model.

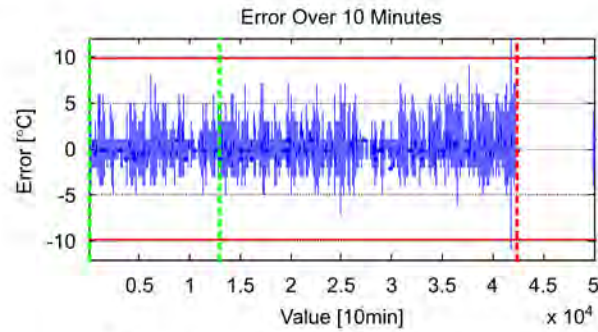


Fig. 40. Prediction error for the autoregressive NN model.

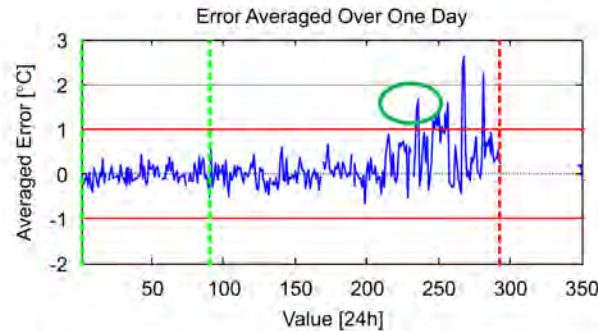


Fig. 41. Averaged prediction error for the autoregressive NN model.

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

	First ALV (days)	Second ALV/trend (days)
Full signal reconst. NN	81	70
Autoregressive NN	59	48

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. 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 as in 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 Fig. 11. The prediction of the FSRC NN model is shown in Figs. 42 and 43 below. 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.

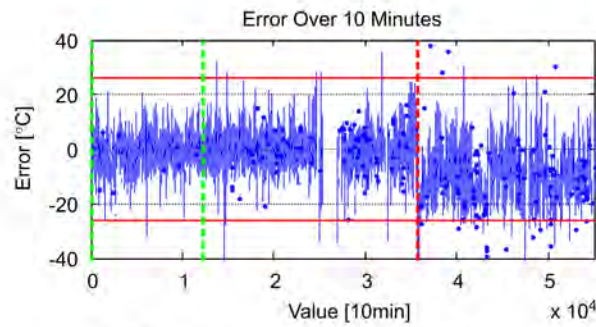


Fig. 42. Prediction error for the FSRC NN model.

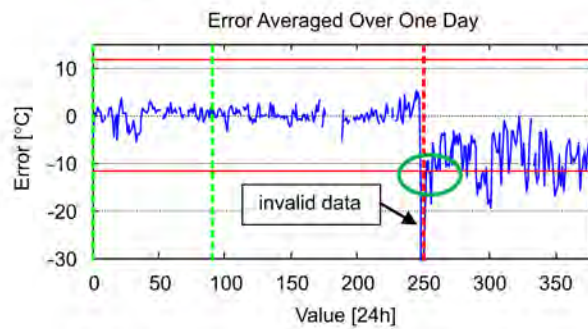


Fig. 43. Averaged prediction error for the FSRC NN model.

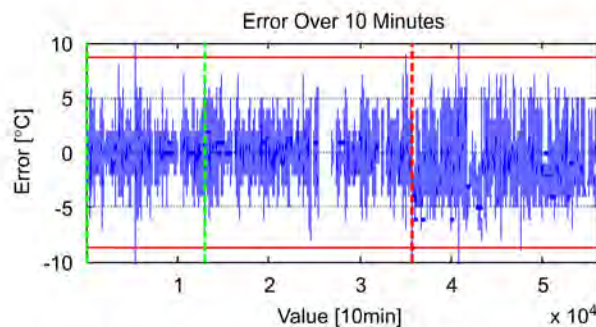


Fig. 44. Prediction error for the autoregressive NN model.

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 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 personnel 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

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

Here the invalid data generated during service are visible, which emphasizes 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

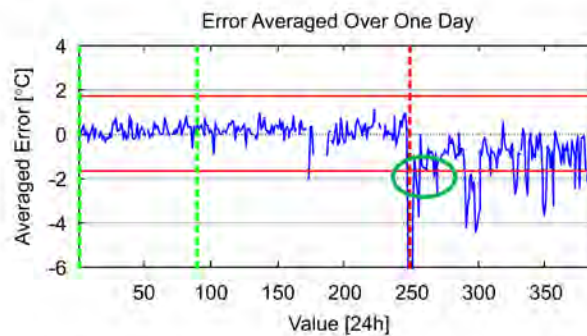


Fig. 45. Averaged prediction error for the autoregressive NN model.

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 of 1:1 with the visual shift in the raw temperature data, the autoregressive model has a smaller shift. Similar to the previously discussed stator temperature anomaly the autoregressive model has its advantage in having larger amplitudes in the 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

Comparisons of 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 gives only an introduction to possible applications, the procedure in favor 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 the 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 the case of 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 months (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. slowly 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 of previous measurements. For this reason, not only can the relative signal changes 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 the case of 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] A. Zaher, S.D.J. McArthur, D.G. Infield, Online wind turbine fault detection through automated SCADA data analysis, *Wind Energy* 12 (2009) 574–593.
- [2] T.W. Verbruggen, Wind turbine operation & maintenance based on condition monitoring, ECN-C-03-047, 2003.
- [3] M.A. Sanz-Bobi, J. del Pico, M.C. Garcia, SIMAP: intelligent system for predictive maintenance application to the health condition monitoring of a windturbine gearbox, *Computers in Industry* 57 (2006) 552–568.
- [4] A.S. Zaher, S.D.J. McArthur, A multi-agent fault detection system for wind turbine defect recognition and diagnosis, *Proceedings of IEEE Lausanne Powertech* (2007) 22–27.
- [5] L. Shuhui, et al., Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation, *Journal of Solar Energy Engineering* 123 (2001) 327–332.
- [6] M. Süttmann (Schlechtingen), Condition monitoring in wind turbines—a drive train monitoring system, Master Thesis, 2010.
- [7] S. Blundell, K.M. Blundell, *Concepts in Thermal Physics*, 2006, 978-0198567691.
- [8] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, ISBN 0132733501, 1998.
- [9] M.Y. Rafiq, G. Bugmann, D.J. Easterbrook, Neural network design for engineering applications, *Computers and Structures* 79 (2001) 1541–1552.
- [10] K. Swingler, *Applying Neural Networks—A Practical Guide*, Elsevier, ISBN 0340705892, 1996.
- [11] L. Tarassenko, *Guide to Neural Computing Applications*, Elsevier, ISBN 0340705892, 1998.
- [12] P. Caselitz, J. Giehardt, *Advanced Maintenance and Repair for Offshore Wind Farms using Fault Prediction Techniques*, Institut für Solare Energieversorgungstechnik e.V.—Department of Control Engineering, 2001.

3.2. Publication [P2] (published)

The FSRC NN as set up to for the signals

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

showed good performance both in terms of the Standard Deviation (SD or σ) and the capability to detect anomalies in the signals [26]. Due to the number of hidden neurons to be generally unknown, several trainings must be performed to find a NN architecture giving good results [27]. Moreover, Tarrasenko [27] and Rafiq [28] recommend reducing the risk of local minima by performing several runs with random weight initialization. In total, this leads to a large number of trials, before a good performing network may be found.

Considering the large amount of models requiring training in the practical application at wind turbines, further research was performed to identify more efficient model types.

In 1993 Jang [29] described an adaptive network based fuzzy interference system and showed that in comparison to NN, fewer parameters must be trained leading generally to faster training [29]. More detailed information about this model type can be found in Jang, Sun and Mizutani [30].

Especially for monitoring of the wind turbine power output many model based approaches have been under research in the past. Using the most popular model types a further performance comparison was carried out, including the rather new Adaptive Neuro Fuzzy Interference System (ANFIS) models. Performance is compared in terms of the Mean Absolute Error (MAE), the Root Mean Squared error (RMS), the Mean Absolute Percentage Error (MAPE), the Standard Deviation (SD) and the ability to highlight power output anomalies. The performance evaluation parameters are defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_{m,i} - x_{p,i}| \quad (1) \quad RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{m,i} - x_{p,i})^2} \quad (2)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{x_{m,i} - x_{p,i}}{x_{m,i}} \right| \quad (3)$$

With $x_{m,i}$ being the i th measurement, $x_{p,i}$ the i th prediction and n the number of measurements.

Using data-mining approaches for wind turbine power curve monitoring: A comparative study

Meik Schlechtingen¹⁾, Ilmar Ferreira Santos²⁾, Sofiane Achiche³⁾

Abstract—Four data-mining approaches for wind turbine power curve monitoring are compared. Power curve monitoring can be applied to evaluate the turbine power output and detect deviations, causing financial loss. In this research cluster center fuzzy logic, neural network and k-nearest neighbor models are build and their performance compared against literature. Recently developed adaptive neuro-fuzzy-interference system models are set up and their performance compared with the other models, using the same data.

Literature models often neglect the influence of the ambient temperature and the wind direction. The ambient temperature can influence the power output up to 20%. Nearby obstacles can lower the power output for certain wind directions. The approaches proposed in literature and the ANFIS models are compared by using wind speed only and two additional inputs.

The comparison is based on the mean absolute error, root mean squared error, mean absolute percentage error and standard deviation using data coming from three pitch regulated turbines rating 2 MW each. The ability to highlight performance deviations is investigated by use of real measurements. The comparison shows the decrease of error rates and of the ANFIS models when taking into account the two additional inputs and the ability to detect faults earlier.

Index Terms—Wind energy, power generation, signal analysis, condition monitoring, machine learning, fuzzy neural networks, neural networks, power system faults, data mining

I. NOMENCLATURE

ANFIS	Adaptive Neuro-Fuzzy Interference System
CCFL	Cluster Centre Fuzzy Logic
k-NN	k-Nearest Neighbor
M5P	Quinlan' M5 algorithm for including trees
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MF	Membership Function
MLP	Multi-Layer-Perceptron
NN	Neural Network
NR	Normal Range

REP	Representative
RMS	Root Mean Squared error
SCADA	Supervisory Control And Data Acquisition
SD	Standard Deviation
WEC	Wind Energy Converter

II. INTRODUCTION

In the past decades the cumulated worldwide installed capacity of Wind Energy Converters (WECs) grew exponentially. One of the reasons is the achieved reduction in cost of energy that has nowadays reached a level where it is almost comparable to conventionally generated power from coal and gas fired power plants. More and more wind turbine operators decide to trade their energy directly on the electricity market. For this purpose and to keep the cost of energy down and increase profit margins, operators need to be able to prognosticate the performance of their turbines more accurately. In case of decreased turbine performance, operators may be unable to deliver their traded amount of energy and consequentially have to pay fines. Furthermore, financial loss is generated as the power output of the turbine is lower than expected and the revenue is hence missing on the balance sheet. Here, power curve monitoring can serve as an effective method to evaluate the performance as power curves for WECs describe the essential relation between wind speed and electrical power output [1]. Detected decrease allows the operator to take action to identify the root cause and improve performance.

Different models were proposed in the past to estimate wind turbine power curves for performance evaluation. The basic idea of all model approaches in this context is to identify closely related signals (e.g. the wind speed) to use them to build a model of the power output. After model training (learning the model the input-output e.g. wind speed – power output relation) the model is kept fixed and applied in the following using the inputs to obtain an expectation of the output. The prediction error can then be an indicator for anomaly - the prediction error is defined here as the difference between the model's output (expectation) and the real measurement.

In 1997, Li *et al.* [2] presented a method using Multi-Layer Perceptron (MLP) Neural Networks (NNs) to predict wind power generation of stall regulated wind turbines. NNs can learn non-linear relationships between input and output data sets by use of activation functions within the hidden

¹⁾ Department of Technical Operation Wind Offshore, EnBW Erneuerbare Energien GmbH, Germany Admiralitätsstr.4; 20459 Hamburg; Email: m.schlechtingen@enbw.com; Tel.: +49 40533268134

²⁾ Department of Mechanical Engineering, Section of Solid Mechanics, Technical University of Denmark, Denmark

³⁾ Department of Mechanical Engineering, Machines Design Section, Ecole Polytechnique de Montréal, Canada

neurons. However, it uses a black box approach to globally fit a single function, to the data and thereby losing insight into the problem [3].

Four years later the same authors published a comparative analysis of regression and artificial NN models for wind turbine power curve estimation [4]. The comparison showed that NN models perform better than regression models. In 2008 Üstütaş and Duran Şahin applied Cluster Centre Fuzzy Logic (CCFL) modelling for power curve estimation [5]. In this method the number of cluster centres is determined and their initial values for initializing iterative optimization are based on clustering algorithms. [5]. The CCFL model is compared to a polynomial model estimated with the least squares algorithm. It is concluded that the Root Mean Squared error (RMS) is much lower when using the fuzzy logic methodology [5]. However, the clustering approach puts limits on the type of MF used and thus the accuracy of the model depends mainly on the number of cluster centres. A probabilistic (Copula) model of wind turbine performance was proposed by Stephen et al. [6] in 2011, but the performance was not quantified in terms of the metrics used here.

It is worth noting that the modelling methodologies presented in [1] [2] [4] [5] [6] were applied to stall regulated (fixed pitch angle) turbines only. The pitch angle describes the angle of the rotor blade, which adjustment can be used to control the aerodynamic lift. Nowadays turbines are pitch regulated (variable pitch angle). Stall regulated wind turbines have decreased power output above rated wind speed (stall region), while pitch regulated wind turbines maintain constant power in this region up to the maximum operating wind speed (cut-out wind speed).

In 2009 Kusiak, Zheng, and Song compared five non-parametric models for monitoring wind farm power and compared their performances [7]. These were: MLP NNs, Quinlan's M5 algorithm for including trees (M5P), Representative (REP) tree, Bagging tree and k-Nearest Neighbour (k-NN) models. Model trees are a type of decision tree with linear regression functions at the leaves, and form the basis of a recent successful technique for predicting continuous numeric values. They can be applied to classification problems by employing a standard method of transforming a classification problem into a problem of function approximation [8]. In the k-NN algorithm each new instance is compared with existing ones using a distance metric and the closest existing instances are used to assign the class to the new one [9].

The models were applied, tested and evaluated at a wind power plant consisting of pitch regulated wind turbines with the aim of abnormal power output detection. The k-NN models performed best. The authors also compared the performance of the models to least squares and maximum likelihood parametric models. The least squares parametric model showed good performances [10], but the performance of those models is not quantified in terms of the Mean Absolute Error (MAE) or the RMS.

Among the several comparative studies found in literature the best performing and recommended models for wind turbine power curve monitoring are:

- CCFL models [5]
- NN models [4] [7] [10]
- k-NN models [7] [10]

This paper presents the findings of a power curve modelling performance study using data from modern pitch regulated wind turbines. For this purpose CCFL, NN and k-NN models are set up in accordance with [4] [5] [7] and [10], and their performance is compared to the performance studies reported in these references.

In addition Adaptive Neuro-Fuzzy Interference System (ANFIS) [11] models are developed and their performance evaluated and compared. Earlier computational studies by Kusiak and Wei [12] have proven that the model for methane production prediction built by the ANFIS algorithm offered the best performance [12]. ANFIS models can learn nonlinear signal relations by setting up a set of fuzzy rules and tuning the Membership Function (MF) parameters in a training phase. In [11] it is shown that in comparison to NN models, fewer parameters must be trained in ANFIS models, leading generally to faster training. Furthermore, the output of NN's is difficult to backtrack (black box model) and a-priori knowledge about the system is difficult to be incorporated. Here ANFIS models have a major advantage, due to their output being based on linguistic rules and tuneable MF's.

Modelling approaches proposed in literature tend to neglect the influence of the ambient temperature and the wind direction as input parameters. In section III it is shown that their influence is of non-negligible order and is therefore including them is expected to decrease the prediction error. The performance of the initially developed models is therefore also compared to models considering the two additional inputs.

In this framework the main original contribution of this paper lies in the direct comparison of different model approaches found in literature based on data sets reflecting modern turbine behaviour (pitch regulated turbines) as well as a comparison to newly developed ANFIS models. Moreover, the models are enhanced by considering the ambient temperature and the wind direction to further improve the prediction performance. Performance is measured in terms of the MAE and RMS. Furthermore, the performance of the extended models is additionally evaluated using the Mean Absolute Percentage Error (MAPE) and the Standard Deviation (SD) of the prediction error. A further performance measure used is the ability to highlight abnormal turbine performance. For this scope the trained models are applied to a data set containing real measured performance deviations.

In section III the general power curve properties are discussed and the influencing factors identified based on theory. Moreover it is mentioned where deviations from theory may stem from. In section IV the data sets and the data pre-processes are described, used to filter invalid data and non-operational periods as well as remove signal lags. Section V introduces the different model types and training algorithms. Optimal settings for the different models (e.g. number of neurons or cluster centres) are selected and discussed. Section VI summarizes the performance comparison based on

Finally conclusions are drawn in section VII.

III. GENERAL POWER CURVE PROPERTIES

In this section, the general power curve properties are discussed in order to identify the most influential factors on the power output.

The theoretical power output of a wind turbine can be expressed by the following equation:

$$P = 0.5 \rho A C_p V^3 \quad (1)$$

Where P is the power output, C_p the power coefficient, ρ the air density, V the wind speed and A the swept rotor area. The variable having the largest effect on the power output is the wind speed as it is powered by three.

The power coefficient C_p is a function of the blade pitch angle which is controlled by the turbine controller to regulate the power output. Large pitch angles lead to lower power coefficients and thus lower power output [13].

The air density ρ is defined as $\rho = p / (R \cdot T)$ with p being the absolute air pressure, R the specific gas constant and T the absolute air temperature. " p " is a function of the altitude of the wind turbine location and the fluctuations due to weather phenomena. The pressure is constant when considered in relation to the altitude at the turbine location whereas it varies of about 10% as a function of weather phenomena. The specific gas constant is as function of the humidity. Wet air has a specific gas constant that is about 2% larger than dry air. The air temperature may vary in the range of -20 to 30°C (20%). Hence the temperature has the largest influence on the air density and consequentially also influences the power output. This effect is often neglected in literature when developing models for the power curve estimation, leading to larger prediction errors. Furthermore, the wind direction may have an influence due to terrain and shading effects of neighbouring turbines or nearby obstacles.

A typical power curve of a pitch regulated wind turbine is shown in Fig. 1.

As illustrated in Fig. 1, the power curve can be divided into five regions:

- wind speed < cut in speed – the cut in speed is the wind speed value at which the turbine starts operating; power output zero or negative; pitch angle about 90°
- wind speed > cut in speed < rated speed; power output according to eq.1; Pitch angle between 0...-5°
- wind speed around rated speed; power output according to eq.1; Pitch angle between -5...5°
- wind speed > rated speed; power output according to eq.1; Pitch angle between 5...25°
- wind speed > cut-out speed; power output zero or negative; pitch angle about 90°

In general, the power curve measured can deviate from the theoretical value due to:

- Inaccuracies in wind speed measurement, for example the point of measurement can be situated on top of the

nacelle, whereas the power output depends on the wind speed across the rotor area

- Air density fluctuations
- Yaw and pitch misalignments
- Issues related to period averaging, as the power output has a cubic dependency of the wind speed; strong wind speed fluctuations in the averaging period influences more the power output
- Power output restrictions manually set in the controller
- Shading effects from neighbouring turbines or nearby obstacles

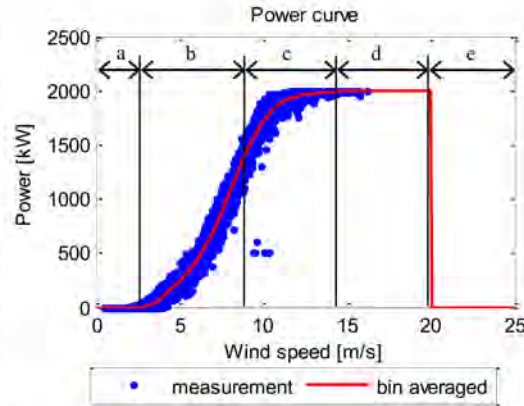


Fig. 1. Power curve properties

Among others these influences cause the variance in the measured power curve visible in Fig. 1. Influences caused by the measurement procedure and the used equipment are addressed in an IEC standard (61400-12-1) [14]. However, in practice these recommendations are seldom met, and mostly used for initial power performance measurements performed by turbine manufacturers [6]. Models which only consider the wind speed as input will not be able to model any of the variance and are thus expected to have a larger prediction error. In this research, the air density influences as well as the shading effect are considered in order to reduce the prediction error.

IV. DATA SETS AND DATA PRE-PROCESSING

The data sets, used in this paper, come from three wind turbines of the 2 MW class and contain 149 Supervisory Control And Data Acquisition (SCADA) signals, including the power output, wind speed, wind direction and the ambient temperature.

229 days of data are selected for model training and validation (32976 10 min. average values in total). The data is collected from December 2010 until August 2011 and thus covers both a summer and a winter period.

The data sets are randomly divided into training and validation data sets with a ratio of 60/40%. The validation data sets are excluded from model training and are only used for performance evaluation in terms of the MAE, RMS, MAPE and SD.

Before training the models, the data are pre-processed according to the methodology proposed in [15] which encompasses the three following steps: 1) validity check, 2) data range check and 3) missing data processing. Moreover, only the data where the turbine is operating (power output > 25 kW) is used. The data set sizes after pre-processing are given in Table I.

TABLE I
Remaining data sizes after data pre-processing

Turbine	Training [10 min avg.]	Validation [10 min avg.]
WEC A	14430	9621
WEC B	13535	9024
WEC C	14327	9552

The power curves of the three wind turbines considered are illustrated in Fig. 2 together with the remaining data after pre-processing.

In addition to these, a time series from a later period exists which contains an anomaly in terms of a reduction in turbine power output, caused by a wrong controller setting after a controller software update (Fig. 3).

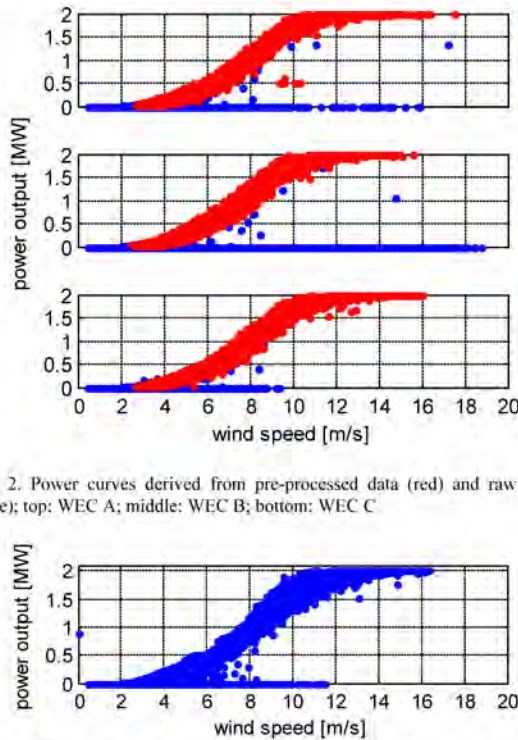


Fig. 2. Power curves derived from pre-processed data (red) and raw data (blue); top: WEC A; middle: WEC B; bottom: WEC C

Fig. 3. Power curve during abnormal turbine performance (WEC A)

In comparison to Fig. 2 the power curve in Fig. 3 is broader especially for wind speeds above 9 m/s. Without the comparison to the normal power curve this type of deviation is difficult to identify by a visual check.

V. MODELS

For each approach two model types are set up. Type one is the one commonly used in literature (e.g. in [4] [5] [7] [10]), whereas with type two, the ambient temperature and the wind direction are also considered. The two types are illustrated Fig. 4.

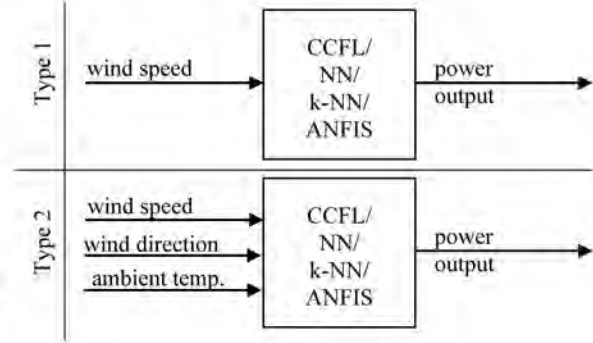


Fig. 1. Model types used for power output modelling

In order to evaluate the contribution of each of the additional inputs of type 2 an exhaustive search is performed, by training ANFIS models with different input combinations (using the pre-processed data of WEC A). Computational time can be saved, by knowing that the input combination 1-2-3 gives the same result as 1-3-2, since the order of inputs presented to ANFIS can be arbitrary. As a measure of importance the correlation coefficient of the predictions and the measurements is used. The result is summarized in Table II.

TABLE II
Input signal importance in terms of the correlation coefficient

Input no.	Input name	1 Input	2 Inputs	3 Inputs
1	Wind speed	0.9939	0.9947	0.9949
2	Ambient temp			
3	Wind direction			
1	Wind speed	0.9939	0.9941	0.9949
2	Wind direction			
3	Ambient temp.			
Computational time ¹⁾		5.85s	14.85s	35.81

¹⁾ Using a standard 2.6 GHz processor

The ambient temperature contributes to the overall explained variance of the target signal with 0.0008 (0.08%), while the wind direction contributes 0.0002 (0.02%).

A. CCFL model

The cluster centre fuzzy logic (CCFL) model is set up in accordance to [5] using the algorithm proposed by [16]. Additional information can be found in [17]. The main free parameter is the number of cluster centres. In [5] eight cluster centres are used. In Fig. 5 the relation between the MAE and the number of clusters is shown for the training data set

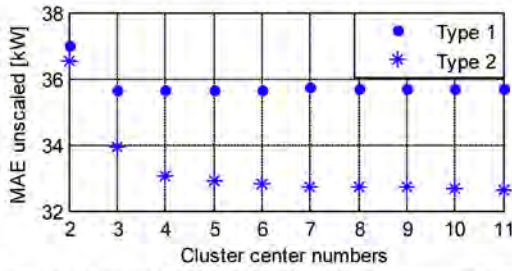


Fig. 5. Relation between the MAE and numbers of cluster centres (WEC A)

Based on Fig. 5 three cluster centres are used for model type one and six for type two. For model type two the MAE does not decrease significantly for larger numbers of cluster centres.

B. NN model

In [10] no information about the setup of the MLP NN is given.

Next to the number of hidden layers, the number of hidden neurons is an important parameter characterizing the NN structure [15]. The number of hidden layers chosen is two, as with three layers, no major performance improvement was observed, but the computational time increased drastically. In [18] 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. This procedure is also followed here. The number of hidden neurons in the hidden layer is $N \cdot n_i$ with n_i being the number of input signals and N an integer multiplier. In accordance with Fig. 6, N is chosen to be three for both model types since for N larger three the MAE does not significantly change.

The training method used is gradient decent with momentum. The advantage of this method is that not only the local gradient of the error function is calculated, but so is the general trend. Thus local minima may be survived and generalization is improved [15].

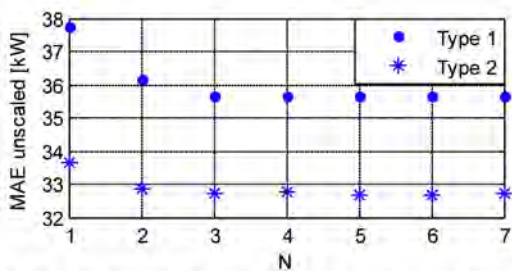


Fig. 6. Relation between the MAE and numbers of hidden neurons N (WEC A)

Sigmoid transfer functions are used for the hidden layer and linear transfer functions for the output layer as this proved to give better results. Next to the Sigmoid transfer functions, logarithmic and linear functions were evaluated. More information about NN models can be found in [19] [20] [21].

C. k-NN model

The k-nearest neighbour model is set up according to the methodology stated in [22]. In [10] the number of neighbor's k was set to 150. Fig. 7 shows the relation between the MAE and the number of neighbours for the training data set coming from WEC A.

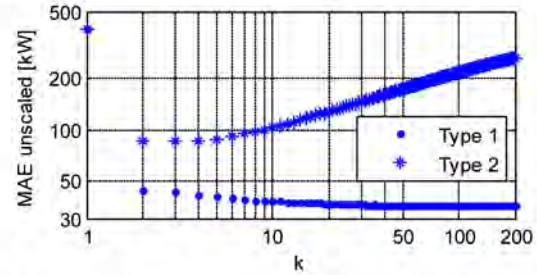


Fig. 7. Relation between the MAE and number of neighbour's k (WEC A)

The number of k used for modelling of type one is 150 and three for type two. Although the performance of model type one does not change much for $k > 30$, k is chosen to be 150 because of this value being used in the reference study allowing a better comparison. More information about the k-NN method can be found in [9] and [22].

D. ANFIS model

The adaptive neuro-fuzzy interference system model build is described in more detail, since these models are comparably young and the application to the specific problem presents a novelty. Additional information about ANFIS models as well as theory can be found in [11].

1) FIS structure

There are two common types of fuzzy inference: the Mamdani [23] and Sugeno [24]. The main difference between the two types is in the consequent part, which can be a non-fuzzy equation (Sugeno) instead of a fuzzy linguistic value (Mamdani) [25]. In this work the Sugeno type is used as it has fuzzy sets only involved in the premise part. However, the consequent part can be a non-fuzzy equation. Due to the qualifiers on the premise parts, each fuzzy if-then rule can be viewed as a local description of the system under consideration [11]. Moreover the Sugeno type is known to be computationally more efficient and has guaranteed continuity of the output surface, which makes it well suited for power curve prediction.

2) Membership functions

Membership functions can be represented by an arbitrary curve whose shape is defined as a function that suits the application. In this paper, generalized normal distribution MF's are employed for the input space and linear membership functions for the output space. The generalized normal distribution functions have the advantage of giving a broad flexibility with regards to the function shape, depending on their parameters. Furthermore these functions assure

smoothness of the transitions in the input space. The free parameters are: location, scale and shape [25].

3) Number of membership functions / rules:

In a conventional Fuzzy Inference System (FIS) the number of rules is decided by an expert who is familiar with the system to be modelled [11]. Generally, the number of MF's can be defined for each input separately [25]. In this research the expert for this particular task is assumed to be unavailable and the number of MF's h per input for both model types is set to three, in agreement with Fig. 8.

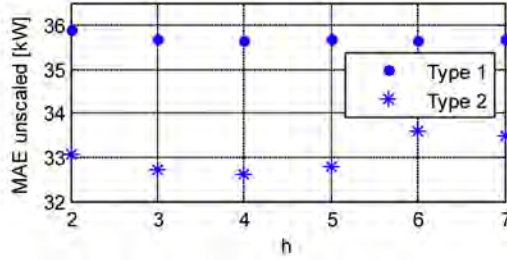


Fig. 8. Relation between the MAE and number of MF's h (WEC A)

Although the MAE for type two and h equal four is smaller, the computational time drastically increases with increasing h , but the performance improvements are not significant.

4) Training method

The training algorithm used is a hybrid learning rule as proposed in [11], consisting of a combination of gradient decent and least squares estimation. Not only can this hybrid learning approach decrease the dimension of the search space in the gradient method, but in general, it will also cut down substantially the convergence time [11].

VI. PERFORMANCE COMPARISON

In this section the performance of the models is compared using MAE, RMS and the ability to identify abnormal turbine power output. The values coming from own calculations are averaged among the model performance of the three turbines considered.

In order to compare the achieved performance to values stated in the literature, rescaling is required. The performance values stated in [10] and [5] are scaled to the 0...100 kW range. Rescaling here is done by dividing the result by the turbine rated power and multiplying by 100 kW. The performance values in this paper will be given both as scaled and non-scaled values.

A. Type 1 model performance comparison

1) Performance based on MAE and RMS

Table III summarizes the results of the performance evaluation.

Type 1 model performance comparison

*	Un-scaled	Scaled to 0...100 kW			
		New models		Literature models	
	MAE	MAE	RMS	MAE	RMS
CCFL	38.72	1.94	2.78	-	0.49 ¹⁾ [5]
NN	38.74	1.94	2.78	5.3 ¹⁾ [4] 3.56 ²⁾ [7] 5.50 [10]	-
k-NN	39.03	1.95	2.80	2.15 ²⁾ [7] 3.19 [10]	-
ANFIS	38.71	1.94	2.78	-	-

¹⁾ Rescaled from 500kW; ²⁾ Rescaled from total plant output;

³⁾ Calculated from bin averaged data set; *Values given in kW

The performance of all models is comparable. The chosen free parameters allow all models to map the existing relation between the power output and the wind speed. Fig. 9 shows the measured and the estimated power curve for the validation data set of WEC A using the set up ANFIS model.

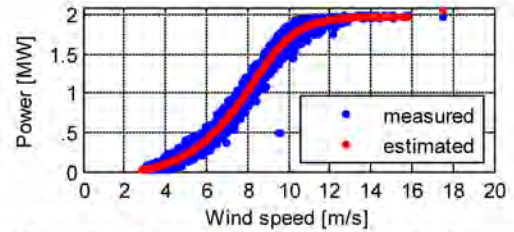


Fig. 9. Power curve of WEC A for the validation period; measured and estimated by the ANFIS model (type 1)

The model is incapable of representing the variance of the measurements, but the general power curve shape appears to be well modelled.

The performance of the CCFL, NN and k-NN models is better in terms of the MAE than the values reported in literature (compare Table III). The RMS value coming from [5] is difficult to compare as here an averaged power curve is modelled. Using a similar procedure the achieved RMS with the set up CCFL model is: 0.21 kW.

In [4] and [5] data from stall regulated wind turbines are modelled. Power curves of those turbines have a broader scatter above rated wind speeds, which is caused by the stall effect. Furthermore the power output in this region cannot be modelled as constant as the power output usually drops in this region. Hence modelling of power curves from stall regulated wind turbines is more difficult and a larger prediction error is expected.

One explanation for the poorer performance of the models set up by [7] and [10] can be the fact that they used smaller data sets for their experiments. In both studies data sets containing 3460 observations are used from which 2589 observations were used for training and 871 for validation (compare values in Table III). Furthermore, the settings for the NN structure in [7] and [10] appear to be not ideally set, since the performance of the k-NN model is found to be

TABLE III

better, which is not consistent with the results obtained in this paper.

2) Performance based on abnormal power output identification

Fig. 10a shows the prediction error of the ANFIS model over time and Fig. 10b the corresponding prediction error probability distribution for the validation data set of WEC A.

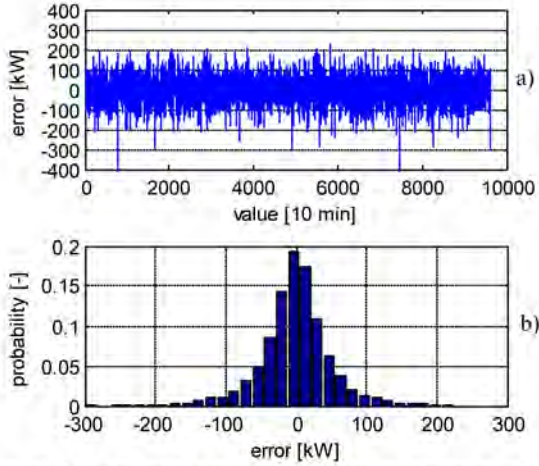


Fig.10. a) Prediction error over time; b) Probability distribution of the prediction error (WEC A; type 1)

In [15] it is shown that prediction error averaging can drastically decrease the variance of the prediction and thus increase the fault visibility by allowing tighter alarm limits. This is because of the prediction errors of successfully trained models usually being normally distributed (compare Fig. 10b). Depending on the application, evaluating the turbine performance once a day can be sufficient. For those applications averaging is recommended.

Fig. 11 shows the performance of the different models both for normal turbine behaviour and a period where the turbine power output is lower than expected due to a wrong controller setting after a software update. Red lines mark the normal operational range being three times the standard deviation of the prediction error in the normal period.

All models show similar performances and are capable in detecting the decrease of the turbine performance. For all models the prediction error leaves the normal performance range, on day 109 of the considered time series.

B. Type 2 model performance comparison

1) Performance based on MAE, RMS, MAPE and SD

The performances of the different models are summarized in Table IV. Again, the given values are averages of the three turbines considered.

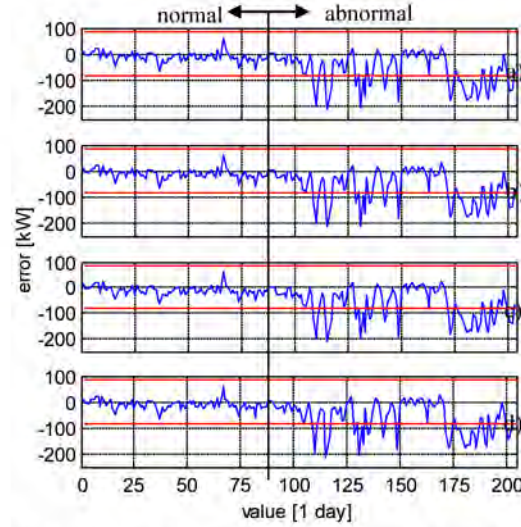


Fig. 11. Averaged prediction error of the different models in case of normal and abnormal turbine power performance (WEC A); a) CCFL; b) NN; c) k-NN; d) ANFIS

TABLE IV
Type 2 model performance comparison

	Un-scaled	Scaled to 0...100kW			Normalized
	MAE [kW]	MAE [kW]	RMS [kW]	SD [kW]	MAPE [%]
CCFL	33.53	1.68	2.42	2.41	8.91
NN	32.49	1.62	2.34	2.34	8.38
k-NN	85.35	4.27	6.02	5.96	25.48
ANFIS	32.01	1.60	2.30	2.30	8.25

The ANFIS models and the NN models perform best. Both model types are known to be well suited for non-linear modelling due to the large flexibility when given enough neurons/ MF's. The k-NN approach can also be used to model non-linear relations as it was shown earlier. However, the existing data is organized in a d -dimensional space, where d is the dimension of the input space, i.e. the number of variables. By adding two more input variables, the number of considered neighbour's k must decrease (compare Fig. 7). As a consequence, the predictions also become more sensitive to outliers in the reference data set due to reduced number of considered and existing neighbours. In fact, the distances between the neighbours may increase since the data is spread among a larger space. Hence the number of observations should be larger in order to achieve better model performance. The CCFL algorithm also organizes the data in a d dimensional space. Depending on the number of clusters centres and radii set, the cluster centres are identified for each dimension separately. For each cluster centre a MF is initialized and linked via fuzzy logic rules. The MF shapes are Gaussian normal distributions. The performance of the CCFL is somewhat lower than the performance of the NN and ANFIS, which is potentially caused by the MF shapes.

The measured and estimated power curve of the validation data set from WEC A is visualized in Fig. 12 using the ANFIS model.

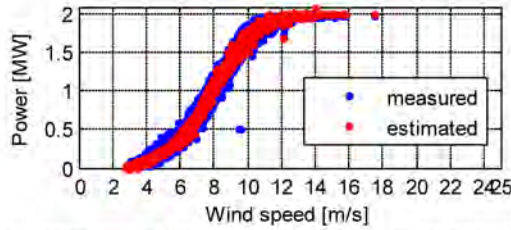


Fig. 12. Power curve of WEC A for the validation period measured and estimated by the ANFIS model (type 2)

In comparison to Fig. 9 the modelled power curve has a broader scatter and some of the variance in the measurements is modelled, and incorporated in the model.

2) Performance based on abnormal power output identification

Fig. 13 shows the prediction error of the ANFIS model for the validation data set of WEC A. The corresponding prediction error probability distribution is compared to the distribution of type one in Fig. 14.

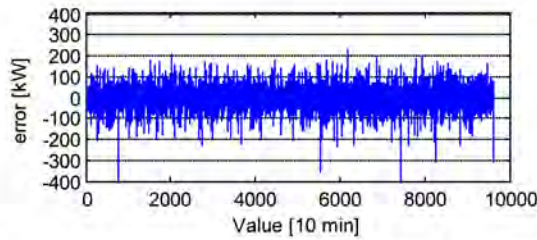


Fig. 13. Prediction error of the ANFIS model for the validation data set (WEC A; type 2)

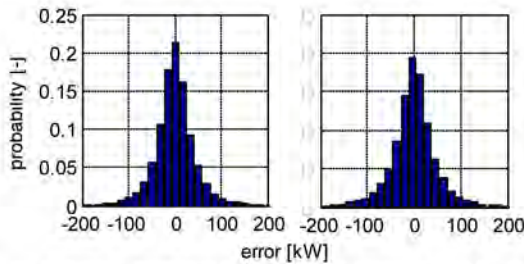


Fig. 14. Probability distribution of the prediction error (WEC A); left: type 2; right: type 1

It is important to highlight here, that when moving from model type one to model type two the prediction error magnitudes decreased.

The probability density distribution becomes slender and the probability of obtaining a low prediction error increases. The capability of detecting abnormal turbine performance of the different models is visualized in Fig. 15, using the same

definition of the normal operational range as in the previous section and the same time series.

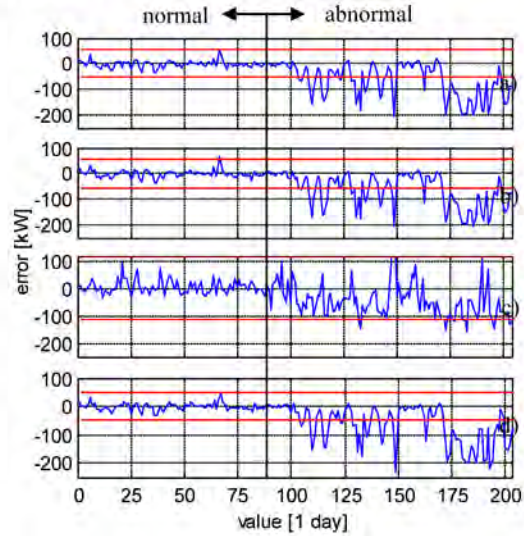


Fig. 15. Averaged prediction error of the different models in case of normal and abnormal turbine power performance (WEC A); a) CCFL; b) NN; c) k-NN; d) ANFIS

Table V gives an overview about the determined Normal Ranges (NR) for the different models and states the day the NR is violated, together with the model deviation at this day.

TABLE V
Anomaly detection capabilities of the different models

	Day of NR violation	Model deviation [kW]	NR [kW]
CCFL	104	-54.98	± 53.60
NN	105	-64.19	± 54.99
k-NN	132	-124.50	± 115.10
ANFIS	104	-54.57	± 50.17

The performance of the k-NN model is poor. Not only is the SD of the prediction error large, but also the anomaly is detected late in comparison to the other models. Based on the day of NR violation, the ANFIS and the CCFL model perform equally well and highlight the anomaly one day earlier than the NN model. Considering also the difference between the NR and the model deviation at the point of NR violation as an indicator for the fault visibility, the ANFIS model performs better with a difference of -4.40 kW in comparison to -1.38 kW for the CCFL model.

In comparison to model type one, the consideration of the ambient temperature and the wind direction in model type two lead to more accurate models, i.e. lower variance in the prediction error and enabled anomaly detection five days earlier.

VII. CONCLUSION

Next to the wind speed, the ambient temperature and the wind direction were proven to be important parameters when setting up data-mining models for wind turbine power curve monitoring. The results show that when only the wind speed is used as input, the performance of the tested models is of comparable magnitude. Using data from pitch regulated modern wind turbines all models set up perform better than the ones described in literature. No major differences exist in the ability to detect abnormal turbine performance.

The results further show that when enhancing the CCFL, NN and the ANFIS models by using the ambient temperature and the wind direction as additional inputs, earlier detection of abnormal turbine performance is possible due to reduced variance in the prediction errors. For the available data set, the anomaly is detected up to five days earlier than with the models using only the wind speed.

Using the enhanced models, the differences between ANFIS, CCFL and the NN in RMS, MAE, MAPE and SD are rather small. The ANFIS model not only showed the best performance in terms of this metric, but also in terms of abnormal power output detection, whereas the k-NN model performed worst. One explanation for the poor performance of the k-NN model can be the higher dimensionality of the space where the data are distributed when adding more input variables. By adding more input variables the number of considered neighbours must decrease. In the consequence the predictions become more sensible to outliers in the reference data set.

Depending on the wind resources at the turbine site it can take a longer time before data up to the cut-out speed are collected. Hence the available data sets for model training may not include all wind speeds. For pitch regulated wind turbines this does not endanger the success of the proposed procedure. In case no high wind speeds measurements are available, the power curve characteristics allow modelling the power output in this region as constant rated power. Thus the power curve can be split into two parts: the first is modelled by the set up model and the second is constant rated power for high wind speeds up to the cut-out speed.

Even with the enhanced models it still remains a challenge to identify small performance deviations in the 10 min average period. Future research should therefore focus on how to further reduce the variance in the prediction errors. This could for instance be achieved by considering the air pressure or other parameters influencing the turbine power output.

In order to avoid financial loss, turbine operators are strongly recommended to monitor the performance of their turbines with either of the methods proposed. Ideally this would be done using the enhanced ANFIS models. Those models not only showed best performance but also their results can be, if necessary, back-tracked and both the rules and the MF can be added or further modified by an expert to fine-tune them.

In case no advanced modelling tools are available which would allow ANFIS models to be set up, easy to program methods such as the k-NN can be used for power curve monitoring when considering only the wind speed. The

drawback is, however, a larger prediction error and delayed anomaly detection.

Once performance decrease is detected, operators can take action to improve the performance, dependent on the root cause identified (here the problem was fixed by a further software update).

- [1] Wächter, M.; Milan, P.; Mücke, T.; Peinke, J. (2011). "Power performance of wind energy converters characterized as stochastic process: applications of the Langevin power curve." *Wind Energy Vol. 14* pp.711-717. Wind Energy Vol. 14 pp. 711-717.
- [2] Li, S.; O'Hair, E. and Giesselmann, M.G. (1997). "Using neural networks to predict wind power generation." *Proceedings of the International Solar Energy Conference, Washington D.C.* pp.415-420.
- [3] Haykin, S. (1998). "Neural Networks: A Comprehensive Foundation." ISBN: 0132733501. Prentice-Hall.
- [4] Li S.; Wunsch, D.C.; O'Hair, E.; Giesselmann, M.G. (2001). "Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation." *Journal of Solar Energy Engineering Vol. 123* pp.327-331.
- [5] Üstütaş, T. and Duran Şahin, A. (2008). "Wind turbine power curve estimation based on cluster center fuzzy logic modeling." *Journal of Wind Engineering and Industrial Aerodynamics Vol. 96* pp.611-620.
- [6] Stephen, B.; Galloway, S.J.; McMillan, D.; Hill, D.C.; Infield, D.G. (2011). "A Copula model of wind turbine performance." *IEEE Transactions on Power Systems Vol. 26* pp.965-966.
- [7] Kusiak, A., Zheng, H. and Song, Z. (2009). "Models for monitoring wind farm power." *Renewable Energy Vol.34* pp.583-590.
- [8] Frank, E.; Wang, Y.; Inglis, S.; Holmes, G.; Witten, I.H. (1998). "Using model trees for classification." *Maschine Learning Vol. 32* No.1, pp. 63-76.
- [9] Witten, I.H. and Frank, E. (1999). "Data mining: practical machine learning tools and techniques." ISBN: 978-1-558-60552-7.
- [10] Kusiak, A., Zheng, H. and Song, Z. (2009). "On-line monitoring of power curves." *Renewable Energy Vol.34* pp.1487-1493.
- [11] Jang, J.-S. R. (1993). "ANFIS: Adaptive-Neuro-Fuzzy Inference System." *Transactions on Systems, Man and Cybernetics Vol.23* No.3. pp.665-685.
- [12] Kusiak, A. and Wei, X. (2011). "Prediction of methane production in wastewater treatment facility: a data-mining approach." *Annals of Operations Research*.
- [13] Hansen, M.O.L. (2008). "Aerodynamics of wind turbines." ISBN: 978-1-84407-438-9.
- [14] International Electrotechnical Commission. (2005). "Wind turbines - Part 12-1: Power performance measurements of electricity producing wind turbines." *International Standard, IEC 61400-12-1*.
- [15] Schlechtingen, M. and Santos, I.F. (2011). "Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection." *Mechanical Systems and Signal Processing Vol.25* No.5. pp.1849-1875.
- [16] Chiu, S.L. (1994). "Fuzzy model identification based on cluster estimation." *Journal of Intelligent Fuzzy Systems Vol. 2* pp.267-278.
- [17] Ross, T. J. (2010). "Fuzzy logic with engineering applications." Wiley. 3rd.
- [18] Tarassenko, L. (1998). "Guide to Neural Computing Applications." ISBN:0340705892. Elsevier.
- [19] Rafiq, M.Y., Bugmann, G. and Easterbrook, D.J. (2001). "Neural network design for engineering applications." *Computers and Structures*. 79 pp.1541-1552.
- [20] Swingler, K. (1996). "Applying Neural Networks - A Practical Guide." ISBN: 0340705892. Elsevier.
- [21] Tarassenko, L. (1998). "Guide to Neural Computing Applications." Elsevier. ISBN:0340705892.
- [22] Bishop, C.M. (2009). "Pattern recognition and machine learning." ISBN: 978-0-387-31073.
- [23] Mamdani, E. H. (1977). "Application of fuzzy logic to approximate reasoning using linguistic synthesis." *Transactions on Computers Vol.23* No.12. pp.1182-1191.
- [24] Sugeno, M. (1985). "Industrial applications of fuzzy control." Elsevier Science Publishers.

- [25] **Schlechtingen, M., Santos, I.F. and Achiche, S.** (2013). "Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description. " *Journal of Applied Soft Computing* Vol.13, pp.259-270.

VIII. BIOGRAPHIES



study he worked at a department.

Meik Schlechtingen received a diploma from the University of Applied Science Gießen-Friedberg (Germany) in 2008 and Master of Science in Engineering degree from the Technical University of Denmark (Denmark) in 2010. Since 2010 he is industrial Ph.D. student at the Technical University of Denmark and employed by the EnBW Erneuerbare Energien GmbH. His research focuses on wind turbine condition monitoring using SCADA data and vibration measurements. Prior his Ph.D.



interests focus upon machinery dynamics, tribology, control and mechatronics.

Ilmar Ferreira Santos, Dr.-Ing. dr. techn. livre-docente, is full professor at Technical University of Denmark, Department of Mechanical Engineering. He received his "Dr.-Ing." from the Technical University of Munich, in Germany, his "dr. techn." from Technical University of Denmark, and his "livre-docente" from State University of Campinas, Brazil. He works in the field of multi-physics design, optimization, monitoring and control of electro-mechanical machine elements. His research

Sofiane Achiche, Ing. M.Sc.A, Ph.D., is a professor at École Polytechnique de Montréal, Mechanical Engineering Department, Design of Machinery Section. He received his Ph.D. from École Polytechnique de Montréal, Canada. His research interests focus upon evolutionary computational intelligence applied to engineering problems such as condition monitoring. Furthermore, he works in the field of mechatronics design as well as understanding and modeling activities of new product development processes for decision support purposes.



3.3. Publication [P3] (published) & [P4] (submitted)

Publication [P3] and [P4] is a two part paper series giving an overview about the research findings and the resulting CMS development to monitor wind turbine SCADA data. The publication is based on the data set coming from power plant 1. The earlier research (see publication [P2]) showed good performance of the ANFIS models. Publication [P3] includes a further comparison between NN and ANFIS models concerning the training speed and the prediction error SD, but its focus lies on the system description and the involved algorithms and recommendations.

Publication [P4] gives application examples to show the working principle of the CMS.



Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description

Meik Schlechtingen^{a,*}, Ilmar Ferreira Santos^b, Sofiane Achiche^c

^a Department of Technical Operation Wind Offshore, EnBW Erneuerbare Energien GmbH, Admiralitätsstr. 4, 20459 Hamburg, Germany

^b Department of Mechanical Engineering, Section of Solid Mechanics, Technical University of Denmark, Denmark

^c Department of Mechanical Engineering, Machine Design Section, École Polytechnique de Montréal, Canada

ARTICLE INFO

Article history:

Received 9 December 2011

Received in revised form 3 May 2012

Accepted 8 August 2012

Available online 23 August 2012

Keywords:

ANFIS models

Condition monitoring

Wind turbine

SCADA data

Normal behavior models

ABSTRACT

This paper proposes a system for wind turbine condition monitoring using Adaptive Neuro-Fuzzy Interference Systems (ANFIS). For this purpose: (1) ANFIS normal behavior models for common Supervisory Control And Data Acquisition (SCADA) data are developed in order to detect abnormal behavior of the captured signals and indicate component malfunctions or faults using the prediction error. 33 different standard SCADA signals are used and described, for which 45 normal behavior models are developed. The performance of these models is evaluated in terms of the prediction error standard deviations to show the applicability of ANFIS models for monitoring wind turbine SCADA signals. The computational time needed for model training is compared to Neural Network (NN) models showing the strength of ANFIS in training speed. (2) For automation of fault diagnosis Fuzzy Interference Systems (FIS) are used to analyze the prediction errors for fault patterns. The outputs are both the condition of the component and a possible root cause for the anomaly. The output is generated by the aid of rules that capture the existing expert knowledge linking observed prediction error patterns to specific faults. The work is based on continuously measured wind turbine SCADA data from 18 turbines of the 2 MW class covering a period of 30 months.

The system proposed in this paper shows a novelty approach with regard to the usage of ANFIS models in this context and the application of the proposed procedure to a wide range of SCADA signals. The applicability of the set up ANFIS models for anomaly detection is proved by the achieved performance of the models. In combination with the FIS the prediction errors can provide information about the condition of the monitored components.

In this paper the condition monitoring system is described. Part two will entirely focus on application examples and further efficiency evaluation of the system.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Condition monitoring of wind turbines is of increasing importance as the size and remote locations of wind turbines used nowadays makes the technical availability of the turbine very crucial. Unexpected faults, especially of large components, can lead to excessive downtime offshore due to lack of suitable crane ships or other specialized vessels. However, also smaller issues and faults of auxiliary equipment like pumps or fans can cause expensive turbine downtime due to restricted turbine accessibility. From an operator's point of view it is therefore worth increasing the effort spent to monitor the turbine condition in order to reduce unscheduled downtime and thus operational costs.

The available CMS mostly require high level knowledge about the system to be monitored. However, this knowledge is difficult to access and does often not exist. Physical models of the system to monitor its condition and predict failures can thus seldom be built with high accuracy due to its complex interaction among several dynamical subsystems. Moreover the available CMS mainly focus on vibrations. Vibration analysis is by far the most prevalent method for machine condition monitoring [2]. However, vibration sensors are not installed on all turbines and components due to their high costs. This causes a large number of turbines not being condition monitored at all or vibration sensors being installed at the main components only.

On the other hand, there is a large amount of operational (SCADA) data available, which can be used to give an indication of the turbine condition. This fact is also stressed by Yang and Jiang [3] who additionally point out that these data are the cheapest resource for developing a CMS for wind turbines. The operational data can be either turbine status information or measurements of signals such

* Corresponding author. Tel.: +49 40533268134.
E-mail address: m.schlechtingen@enbw.com (M. Schlechtingen).

as temperatures, currents or pressures. Using turbine status information Kusiak and Wenyan [4] and Kusiak and Verma [5] show that faults can be predicted 5–60 min in advance. In order to perform predictive maintenance this prediction period is too short, since it does not leave the operator with enough time to take maintenance actions. By applying 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 [6]. Using Neural Network (NN) model based approaches Sanz-Bobi et al. [7], Zaher et al. [1] as well as Schlechtingen and Santos [8] show that changes in signal behavior can be detected, days, weeks and in some cases months in advance. These methods are therefore better suited to allow operators to take measures to improve the condition before the component finally fails. In the model based approaches historical operational data is used to develop normal behavior models capable of predicting a certain output signal, when given one or more input signals. For wind turbine signals these attempts are well suited, since many signals can be found to be correlated to other signals measured simultaneously, e.g. the wind speed or the power output.

One advantage of using normal behavior models to monitor wind turbine signals is that no prior knowledge about the signal behavior is needed. Another important property is that with normal behavior models the possibility of monitoring the signal is widely decoupled from the operational mode as reported by Zaher et al. [1] and Sanz-Bobi et al. [7]. The normal behavior models are developed in periods where the turbine components can be considered healthy (normally operating), usually the period at the beginning of the component lifetime. Afterwards, the trained model is used to predict a specific signal where the prediction error gives an indication of changes in signal behavior and thus incipient faults.

This approach is of large interest in research. In [9] the applicability of a system identification approach using an AutoRegressive with eXogenous input (ARX) model to monitor the condition of a wind turbine generator bearing using SCADA data is shown. However, this approach requires human intervention for parameter selection to find a good performing model. Due to the large amount of signals and turbines to be monitored, human intervention should be avoided. Most activities take advantage of artificial intelligence algorithms (learning capabilities) and among the most advanced systems using this approach is SIMAP [7] and a Multi Agent System (MAS) [1,11]. Both systems use artificial NN to set up the normal behavior models of SCADA data. This line is also followed by Xiang et al. [13], where a NN model based method to monitor the condition of a wind turbine generator bearing is developed and the good performance of NN in this context shown. Additionally, earlier research activities by Schlechtingen and Santos [8] and Schlechtingen [10] proposed a condition monitoring system for wind turbine drive train components using NNs.

Instead of using NNs, ANFIS is used in this paper, which presents a novelty, in this field of application. ANFIS models can learn nonlinear signal relations by setting up a set of fuzzy rules and tuning the Membership Function (MF) parameters in a training phase. Jang [14] shows that in comparison to NN models, fewer parameters must be trained in ANFIS models, leading generally to faster training. The major drawbacks of NNs are their black-box data processing structure and slow convergence speed [15]. Furthermore, a priori knowledge about the system is difficult to be incorporated. Here ANFIS models have a major advantage, due to their output being based on linguistic rules and tuneable MFs.

The monitoring systems developed by the different researchers can furthermore be classified according to the input signals used for modeling. While Sanz-Bobi et al. [7], Zaher et al. [1] as well as Zaher and McArthur [11] use autoregressive approaches to set up the normal behavior models, the research presented by Schlechtingen and Santos [8] showed that it is advantageous using a Full Signal

ReConstruction (FSRC) approach, since this approach additionally allows for monitoring the signal value magnitude. The different approaches mainly concern the question of the used input signals to model the target signals. This approach is also followed by the research presented in this paper.

To interpret the prediction error of the normal behavior models, a fuzzy expert system that outputs a diagnosis, the condition and the certainty of statement based on rules that were established by an expert is employed by Sanz-Bobi et al. [7]. With regard to the surveyed gearbox faults in their research, this approach proved to be successful, since the rules established with fault patterns from one turbine were also applicable to predict the same faults on other turbines.

The research presented in this paper is the result of two years investigations carried out to develop a CMS that uses wind turbine SCADA data available to wind turbine operators. Thereby the advances made by Sanz-Bobi et al. [7] as well as Zaher and McArthur [11] are combined and further developed.

The development of the aforementioned CMS follows the three step strategy below:

1. Normal behavior models of the relevant SCADA data were developed in order to monitor and detect anomalies by considering the prediction error.
2. Occurred anomalies within the prediction error were related to reported faults.
3. The identified relations were implemented in knowledge data bases which are used by FIS to automatically analyze these patterns and output a diagnosis.

The SCADA data used in this research is obtained from eighteen different operating onshore turbines of the 2MW class, where continuous operational data from April 2009 to March 2012 were gathered.

In Section 2 of this article the general concept of the system developed is described giving details about how the information (SCADA data) is processed to finally output condition statements. A brief description of the advantages of ANFIS models over NN models for the given application and a short description of the used model setup and structure is supplied in Section 3. Section 4 focuses on the performance of the developed normal behavior models and describes the identified input–output sets. An example for the prediction error analysis and the model interactions is given in Section 5. In Section 6 the concept of anomaly detection is shown and a definition of abnormal signal behavior is given. Section 7 clarifies how the prediction error together with the information about detected anomalies is processed by the Fuzzy Inference System (FIS) working as fuzzy expert. The output is a diagnosis about the component condition and a potential root cause of the anomaly. Results are discussed in Section 8 and conclusions are drawn in Section 9.

2. Description of the general CMS concept

The CMS developed in this research aims at detecting trends and patterns in SCADA data in order to predict possible failures, giving wind turbine operators enough time to adapt the maintenance schedule or take further measures to prevent unexpected system downtime. For this purpose 10 min averaged SCADA data are used that are commonly available to operators. The general architecture of the CMS developed is shown in Fig. 1.

In the following, the function of the different CMS modules (see Fig. 1) is briefly described.

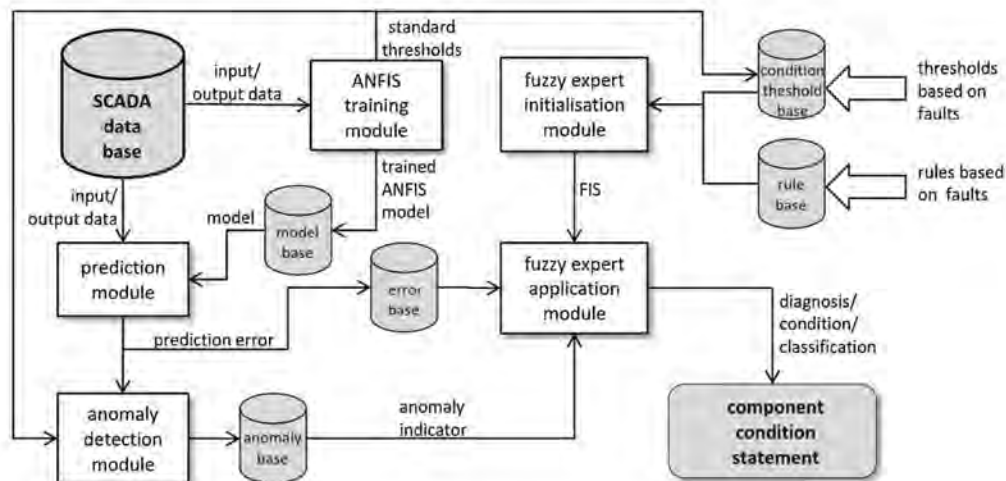


Fig. 1. CMS overview.

2.1. Training module

Within the training module the normal behavior model is trained if a model is not yet available or new training is required. The latter is true if a component is replaced and the signal relations change as a consequence. Before training the model, the data are preprocessed according to the methodology proposed in [8] which includes: (1) a validity check, (2) data range check, (3) missing data processing and (4) lag removal.

In the training module different training levels are implemented to allow early monitoring. A first model training is performed after one month of operational data collection. Further trainings are performed when three, six and nine months of data are available. The output of the training module is the trained ANFIS model and the standard thresholds marking the normal operational range of the wind turbines using the prediction error.

2.2. Prediction module

The prediction module is active once a trained model of the processed signal is available in the model base. The developed normal-behavior model is applied and the prediction error calculated and stored.

2.3. Anomaly detection module

In this module the anomalies in the prediction errors are identified. This is done on the basis of the determined normal-behavior thresholds by the training module or expert defined values. The output is an anomaly matrix containing information about the frequency and date of occurrence, as well as the duration of the current anomaly in days.

2.4. Fuzzy expert initialization module

Here the FIS structures used for diagnosis of the anomalies and component condition statements are initialized with regard to the number of inputs and outputs as well as their MFs. Each component to be monitored has its own FIS structure.

While the inputs solely depend on the component or subsystem to be diagnosed, each FIS structure has the following outputs:

- diagnosis (information about the abnormal behaving signal)
- condition (classification in green, yellow and red color code)
- potential root cause

2.5. Fuzzy expert application module

Within this module the initialized FIS structure is evaluated, given the prediction errors and the information about present anomalies. The output is stored in text format and is visualized to give the analyst a comprehensive summary of the turbine condition.

3. Model setup and structure

3.1. Data set description

The available SCADA data sets from the operating wind turbines contain more than 150 different signals, ranging from hour counters, calculated values, digital indicators of switch positions and set points, to continuous measurements of temperatures, currents, voltages, etc. For some of the continuous measurements the mean, maximum, minimum and standard deviation of the 10 min averaging period is available. In this research only normal behavior models for mean values were considered for three reasons: (1) The peaks stored in the min.–max. values can be caused by transient situations e.g. sudden changes in wind speed that are not in the scope of this research; (2) stochastic effects in the signals are averaged out in the 10 min period, making the prediction less sensitive to stochastic variations; (3) modeling the standard deviations requires representative signals, i.e. other standard deviations or higher resolution time series, that are not accessible.

The remaining signals considered for the CMS are given in Table 1 and a schematic of the turbine and of the sensor location are illustrated in Fig. 2.

In addition to these signals, information about service and non-operational periods is extracted from the corresponding hour

Table 1
SCADA data signal list used for normal behavior model development.

Name of variable	Unit	Sensor location (Fig. 2)	Short description	Normal behavior modeled
Spinner temp.	°C	1	Spinner temperature (in hub housing)	Yes
Hub controller temp.	°C	1	Pitch controller temperature	Yes
Pitch angle		1	Blade pitch angle (avg. over all 3 blades)	Yes
Hydraulic oil temp.	°C	10	Hydraulic oil temperature (used for pitching blades)	Yes
Rotor speed	rpm	2	Rotor speed (low speed shaft)	Yes
Gear bearing temp. (HSS)	°C	3	High speed shaft bearing temperature	Yes
Gear oil temp.	°C	3	Gearbox oil temperature	Yes
Generator speed	rpm	4	Generator speed (high speed shaft)	Yes
Generator bearing temp. 1	°C	5	Generator bearing temperature gearbox end	Yes
Generator bearing temp. 2	°C	5	Generator bearing temperature transformer end	Yes
Generator slip ring temp.	°C	5	Generator slip ring temperature	Yes
Generator ph.1 temp.	°C	5	Generator stator temperature phase 1	Yes
Generator ph.2 temp.	°C	5	Generator stator temperature phase 2	Yes
Generator ph.3 temp.	°C	5	Generator stator temperature phase 3	Yes
Generator current ph.1	A	6	Generator current phase 1	Yes
Generator current ph.2	A	6	Generator current phase 2	Yes
Generator current ph.3	A	6	Generator current phase 3	Yes
Power output	kW	6	Turbine power output	Yes
Reactive power	kVAr	6	Turbine reactive power consumption	Yes
Grid inverter ph.1 temp.	°C	6	Inverter temperature grid end	Yes
Grid rotor inverter ph.1 temp.	°C	6	Inverter temperature phase 1 generator end	Yes
Grid rotor inverter ph.2 temp.	°C	6	Inverter temperature phase 2 generator end	Yes
Grid rotor inverter ph.3 temp.	°C	6	Inverter temperature phase 3 generator end	Yes
Converter cooling water temp.	°C	6	Converter cooling water temperature	Yes
Converter choke coil temp.	°C	6	Converter choke coil temperature	Yes
Converter controller temp.	°C	6	Converter controller temperature	Yes
Top controller temp.	°C	6	Turbine controller temperature	Yes
Grid busbar temp.	°C	8	Busbar temperature	Yes
HV transformer ph.1 temp.	°C	8	High voltage transformer temperature phase 1	Yes
HV transformer ph.2 temp.	°C	8	High voltage transformer temperature phase 2	Yes
HV transformer ph.3 temp.	°C	8	High voltage transformer temperature phase 3	Yes
Nacelle temp.	°C	7	Temperature in nacelle (housing) of the turbine	Yes
Wind speed	m/s	9	Wind speed	Yes
Wind direction		9	Wind direction	No
Ambient temp.	°C	9	Outdoor temperature	No

counters. This information is used for preprocessing the data according to the methodology presented in [8] (Fig. 3).

3.2. Model type

Several researchers applied NN to set up normal behavior models in the field of condition monitoring. Some of the key features of NN are their high processing speeds which are due to their massive parallelism, their proven ability to be trained, to produce instantaneous and correct responses from noisy or partially incomplete data, and their ability to generalize information over a wide range [16]. In earlier publications of the authors [8,10] it was shown that NN are indeed capable of accurately finding an existing input–output mapping for different wind turbine SCADA signals. However, it was found that NN have a high likelihood of becoming trapped in local minima. In Fig. 4 three examples of NN performances

over the number of runs are shown. For this example two hidden neurons per input signal are used.

The variations in performances are up to 20%, stressing the risk of local minima. Furthermore the number of hidden neurons is generally unknown. Tarrasenko [17] and Rafiq et al. [18] suggest performing several runs with random weight initializations and runs with a varying number of hidden neurons to overcome this shortfall. Finally the network that performs best should be chosen. This leads to a large number of different trials to obtain an acceptable solution.

Fuzzy systems are very useful in two general contexts: (1) in situations involving highly complex systems whose behaviors are not well understood and (2) in situations where an approximate, but fast, solution is desired [19]. A further advantage of fuzzy systems is that existing expert knowledge can be implemented to improve

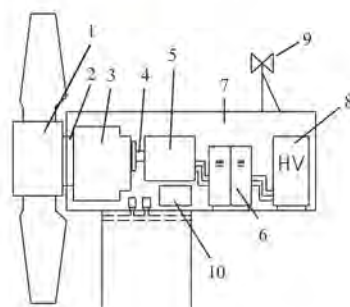


Fig. 2. Wind turbine schematic: sensor positions.

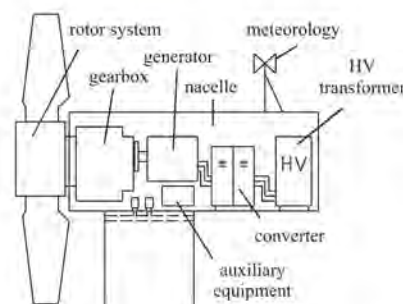


Fig. 3. Wind turbine schematic: components/subsystems in the considered wind turbines.

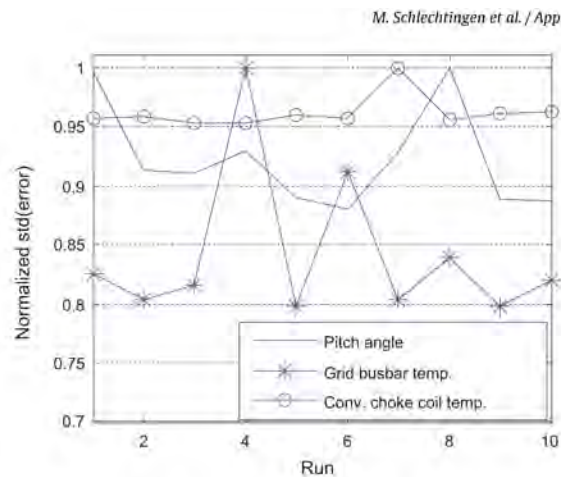


Fig. 4. Exemplary performances of trained NN's over different runs.

the approximation by tuning, removing or adding of membership functions and rules.

Considerable work has been carried out to integrate the learning capability of NN with FIS for deriving the initial rules of a fuzzy system and tuning the membership functions [20]. Fuzzy neural networks have shown to be very advantageous in dealing with real-world problems [21]. These neuro-fuzzy systems combine the benefits of these two powerful paradigms into a single capsule. This gives the ability to accommodate both data and existing expert knowledge about the problem under consideration [20].

In ANFIS the advantages of NN are combined with FIS. Thereby the FIS is used to set up a set of rules whose membership functions parameters are tuned in a training phase.

In the following sections the setup and structure of the ANFIS models are described.

3.2.1. FIS structure

There are two common types of fuzzy inference: the Mamdani [22] and Sugeno [23]. The main difference between the two types is in the consequent part, which can be a nonfuzzy equation (Sugeno) instead of a fuzzy linguistic value (Mamdani). In this work the Sugeno type is used as it has fuzzy sets only involved in the premise part. However, the consequent part can be a non-fuzzy equation. Due to the qualifiers on the premise parts, each fuzzy if-then rule can be viewed as a local description of the system under consideration [14]. Moreover the Sugeno type is known to be more computationally efficient and has guaranteed continuity of the output surface, which makes it well suited for time series prediction.

3.2.2. Membership functions

Membership functions can be represented by an arbitrary curve whose shape is defined as a function that suits the application. In most fuzzy applications straight-line membership functions are used (triangular/trapezoidal). In this paper generalized normal distribution membership functions are employed for the input space and linear MFs for the output space. The generalized normal distribution functions have the advantage of giving a broad flexibility with regard to the function shape, depending on their parameters. Furthermore these functions assure smoothness of the transitions in the input space. The free parameters are: location, scale and shape.

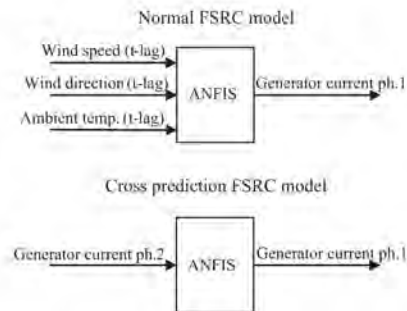


Fig. 5. Comparison between a normal FSRC model and the cross prediction FSRC model.

3.2.3. Number of membership functions/rules

In a conventional FIS, the number of rules is decided by an expert who is familiar with the system to be modeled [14]. Generally, the number of MFs can be defined for each input separately. In this research the expert for this particular task is not available, and the number of membership functions per input is fixed to two, as this, after preliminary tests, was found to be a good compromise between computational efficiency and model performance.

3.2.4. Training method

The learning algorithm used is a hybrid learning rule as proposed in [14], consisting of a combination of gradient decent and least squares estimation. Not only can this hybrid learning approach decrease the dimension of the search space in the gradient method, but in general, it will also cut down substantially the convergence time [14].

3.2.5. Input signals

The number of input signals differs for each signal to be monitored. The choice of relevant input signals is not trivial and requires a combination of physical understanding of the system to be modeled and advanced data reduction techniques. In this research a genetic algorithm combined with a partial least squares regression (GAPLS) [24] is used to detect potential input signals from the large pool of SCADA data. For this purpose the genetic algorithm toolbox supplied by Leardi [25] is applied. However, GAPLS points to the input signals containing most of the target signal features, which may not be the best ones to choose from a condition monitoring point of view if one considers the physical understanding of the phenomenon at hand. An example for possible difficulties with automatic input signal selections is the generator phase currents for which the GAPLS suggests using the generator current ph.1 as input to model generator current ph.2. A model like this is very accurate but by using this input-output set, monitoring of the absolute current level is impossible. Instead only differences between the phases can be detected.

Hence, a combination of data reduction techniques and the understanding of the physical process to select suitable input sets is the preferred strategy used in this paper.

Since both, the information about relative changes between signal readings of the same type as well as their absolute level is important, two types of models are developed. The first is a FSRC (Full Signal ReConstruction) approach, where the target signal is modeled by fully reconstructing it through correlated signals of different types and the second is by fully reconstructing it through correlated signals of the same type. The difference is illustrated in Fig. 5 for the generator phase current.

With the cross prediction FSRC models, asymmetries (for instance differences in generator phase currents or temperatures of

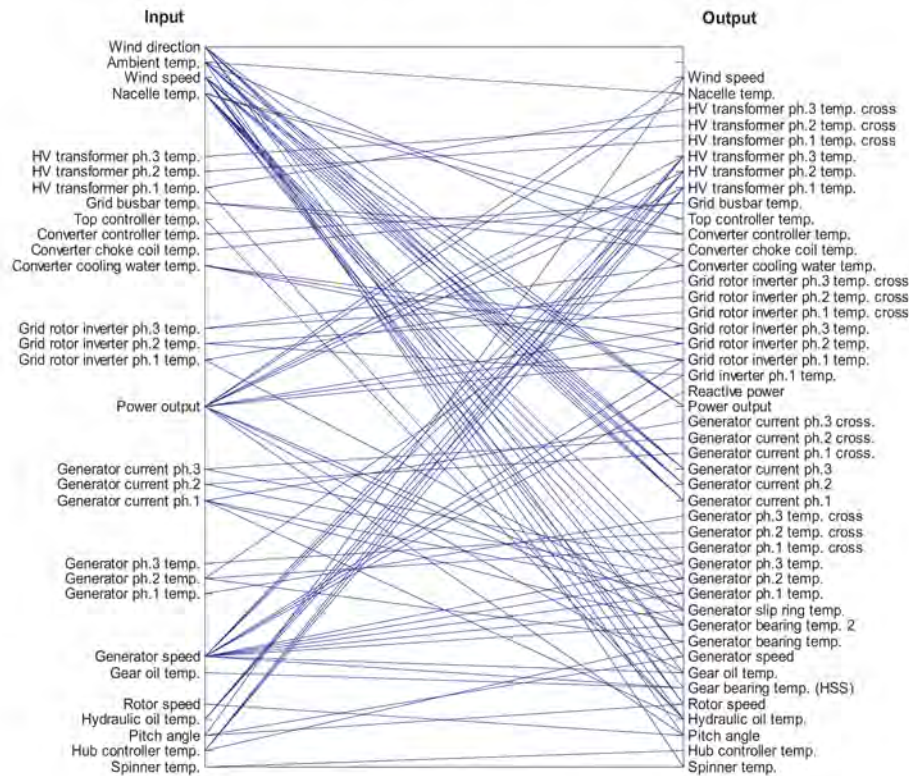


Fig. 6. Input-output sets of the developed models.

the three phases) can be identified very accurately whereas the normal FSRC model is used to monitor the absolute level and relative changes in comparison to other signals.

However, the model accuracy is not the only criterion for choosing input signals when developing normal behavior models for condition monitoring purposes. A further requisite is the fault visibility, as being the visibility of the developing fault in the prediction error. The relation between the fault visibility and the input signal choice is difficult to define and not exactly known, which is why expert knowledge is required to estimate this influence. A general requirement of normal behavior models is that they should have a low prediction error in case of normal behavior and a high prediction error otherwise. For the given problem the combination of the genetic algorithm with engineering knowledge about the component or subsystem lead to the input-output sets illustrated in Fig. 6 for the 45 models developed.

4. Model performance

A good model performance is the direct result of the correct choice of input signals and successful training. There are several ways the performance can be measured. Due to the different nature of the modeled signals (power output, currents, temperatures) a performance comparison between them is not meaningful. What is more of interest is the prediction error variation around its mean, giving an indication of how good the prediction is, by keeping the physical unit. Here the standard deviation is used. Since the prediction errors after training are close to being normal distributed and centered about zero this choice is acceptable (compare Fig. 9). The

values are calculated based on test data sets that were not used for training.

$$E = x_m - x_p \quad (1)$$

$$P = \sigma(E) \quad (2)$$

E is the prediction errors; x_m is the measured values; x_p is the predicted values; P is the performance measure; σ is the standard deviation.

In order to assess the performance information about the normal operational range of the signals is essential (see Table 2). For instance a model performance of 1 °C is good when the normal operational range is –100 to 100 °C, but rather poor if the normal operational range is 25–30 °C.

4.1. Comparison between NN and ANFIS models

To stress the advances of ANFIS models the results of a brief performance and training speed comparison is shown in Table 3 for ANFIS and NN training. The chosen NN training procedure is analog to the one described in [8], but with the difference that here only five runs with random weight initializations are performed. The number of hidden neurons and the number of MFs, per input and output signal is set to two (Table 4).

The training is performed with nine month of operational data (29,513 10 min average values).

The comparison shows that the performance in terms of the standard deviation of both approaches is similar. However, the time required for training is generally smaller with ANFIS models, which is due to the necessary trial and error procedure when training NN.

Table 2
Normal operational ranges of the SCADA data.

Signal	Normal range	Signal	Normal range
Spinner temp.	–10 to 45 °C	Power output	–50 to 2100 kW
Hub controller temp.	–5 to 50 °C	Reactive power	–25 to 300 kVAR
Pitch angle (avg. over all blades)	–5° to 95°	Grid inverter ph.1 temp.	25 to 50 °C
Hydraulic oil temp.	10 to 65 °C	Grid rotor inverter ph.1 temp.	25 to 55 °C
Rotor speed	0 to 16 rpm	Grid rotor inverter ph.2 temp.	25 to 55 °C
Gear bearing temp. (HSS)	30 to 80 °C	Grid rotor inverter ph.3 temp.	25 to 55 °C
Gear oil temp.	30 to 65 °C	Converter cooling water temp.	5 to 50 °C
Generator speed	0 to 1700 rpm	Converter choke coil temp.	5 to 120 °C
Generator bearing temp. 1	10 to 85 °C	Converter controller temp.	5 to 60 °C
Generator bearing temp. 2	10 to 85 °C	Top controller temp.	10 to 50 °C
Generator slip ring temp.	5 to 45 °C	Grid busbar temp.	5 to 60 °C
Generator ph.1 temp.	10 to 140 °C	HV transformer ph.1 temp.	20 to 95 °C
Generator ph.2 temp.	10 to 140 °C	HV transformer ph.2 temp.	20 to 95 °C
Generator ph.3 temp.	10 to 140 °C	HV transformer ph.3 temp.	20 to 95 °C
Generator current ph.1	0 to 1700 A	Nacelle temp.	5 to 50 °C
Generator current ph.2	0 to 1700 A	Wind speed	0 to 35 m/s
Generator current ph.3	0 to 1700 A		

Jang [14] showed that in ANFIS models fewer parameters must be trained, leading generally to faster training.

The performance of ANFIS models does not vary among different runs, which is due to the initialization of the MFs, involving a grid partition algorithm [14,26]. What is important to notice is that with NN and the suggested number of different trials, a better performance can be achieved (see Fig. 7) potentially. However, finding an optimal solution is computationally expensive. Fig. 7 further shows that with ANFIS models a good compromise between training speed and model accuracy can be found.

NN are good in identifying non-linear relations, but their output is difficult to back-track (black box model). A priori knowledge about the system is difficult to incorporate. Here ANFIS models have a major advantage. Not only can a priori knowledge be used in terms of rules or shape of MFs, the output can also be back-tracked by using the information about the rule that fired. When using the grid partitioning method, the number of rules is equal to: nMF^{n_i} with nMF being the number of MFs and n_i the number of inputs. The number of inputs is smaller or equal to four and the number of MFs is equal to two for the models developed, leading to a maximum of 16 rules. Especially for small numbers of inputs this leads to very fast to train and simple to back-track models. However, the model complexity (number of rules) scales exponentially with the number of inputs, which causes training speed also increase exponentially.

Table 3
Performance and speed of ANFIS training.

SCADA signal	ANFIS	
	Elapsed time for training	P
Hydraulic oil temp.	125.2 s	2.00 °C
Generator bearing temp.	37.3 s	1.30 °C
Power output	93.6 s	45.78 kW
Top controller temp.	14.1 s	1.24 °C

Table 4
Performance and speed of NN training.

SCADA signal	NN	
	Elapsed time for training	P
Hydraulic oil temp.	452.3 s	2.06 °C
Generator bearing temp.	373.2 s	1.33 °C
Power output	284.9 s	45.99 kW
Top controller temp.	369.3 s	1.57 °C

4.2. Performance of the set up ANFIS models

Schlechtingen and Santos [8] showed that averaging the prediction error decreases the variance of the system and consequently increases the sensitivity against anomalies. In fact it was shown that by averaging, upcoming faults and anomalies can be detected earlier and with a higher certainty (fault patterns are stronger pronounced). Generating average values can be advantageous, because daily averaged signals of several turbines can easily be compared within a single plot and turbine operators are most often interested in information that can support their day-to-day planning of necessary actions, for which not too much details (such as hourly data) are required [27]. In this paper all analysis is based on averaged prediction errors. The averaging period used is one day, i.e. the average for 144 individual prediction errors is calculated, giving one single averaged prediction error per day.

$$P_{avg} = \sigma(\bar{E}) \quad (3)$$

\bar{E} is the averaged prediction errors; P_{avg} is the performance measure errors; $\sigma(\bar{E})$ is the standard deviation of \bar{E} .

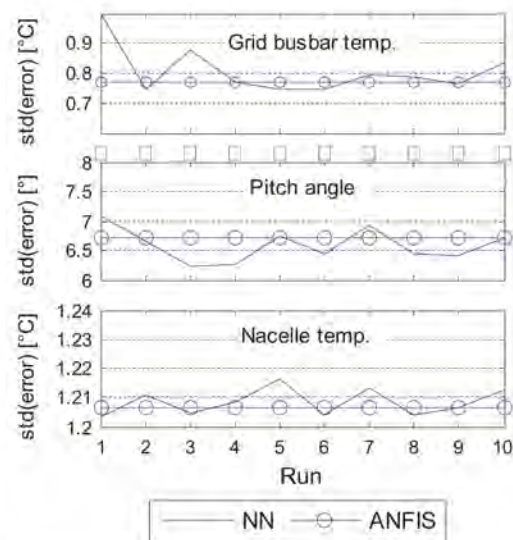


Fig. 7. Performance of different NN and ANFIS trainings over the number of runs.

Table 5
Example model performance in terms of standard deviation of the prediction error after training.

Model	P	P _{avg}	Unit	Model	P	P _{avg}	Unit
Spinner temp.	1.55	0.63	°C	Power output	47.00	16.34	kW
Hub controller temp.	0.55	0.18	°C	Reactive power	0.26	0.03	kWAr
Pitch angle (avg. over all blades)	0.44	0.08	°	Grid inverter ph.1 temp.	0.45	0.13	°C
Hydraulic oil temp.	2.59	1.35	°C	Grid rotor inverter ph.1 temp.	0.42	0.14	°C
Rotor speed	0.21	0.09	rpm	Grid rotor inverter ph.2 temp.	0.42	0.15	°C
Gear bearing temp. (HSS)	0.79	0.35	°C	Grid rotor inverter ph.3 temp.	0.41	0.13	°C
Gear oil temp.	2.09	1.02	°C	Grid rotor inverter ph.1 temp. cross.	0.39	0.07	°C
Generator speed	23.46	9.32	rpm	Grid rotor inverter ph.2 temp. cross.	0.48	0.13	°C
Generator bearing temp. 1	1.44	0.73	°C	Grid rotor inverter ph.3 temp. cross.	0.37	0.08	°C
Generator bearing temp. 2	1.86	0.75	°C	Converter cooling water temp.	0.69	0.33	°C
Generator slip ring temp.	1.30	0.46	°C	Converter choke coil temp.	2.95	1.26	°C
Generator ph.1 temp.	4.98	1.90	°C	Converter controller temp.	0.54	0.22	°C
Generator ph.2 temp.	4.92	1.87	°C	Top controller temp.	1.15	0.48	°C
Generator ph.3 temp.	4.92	1.86	°C	Grid busbar temp.	0.57	0.19	°C
Generator ph.1 temp. cross.	0.51	0.11	°C	HV transformer ph.1 temp.	3.59	1.64	°C
Generator ph.2 temp. cross.	0.51	0.09	°C	HV transformer ph.2 temp.	2.88	1.32	°C
Generator ph.3 temp. cross.	0.66	0.15	°C	HV transformer ph.3 temp.	3.25	1.47	°C
Generator current ph.1	39.07	13.62	A	HV transformer ph.1 temp. cross.	1.59	0.87	°C
Generator current ph.2	38.63	13.29	A	HV transformer ph.2 temp. cross.	1.39	0.80	°C
Generator current ph.3	39.05	13.57	A	HV transformer ph.3 temp. cross.	1.35	0.91	°C
Generator current ph.1 cross.	2.32	1.45	A	Nacelle temp.	1.21	0.60	°C
Generator current ph.2 cross.	2.37	1.72	A	Wind speed	0.20	0.09	m/s
Generator current ph.3 cross.	1.60	0.99	A				

The performances are listed in Table 5 for a random Wind Turbine Generator (WTG) of the fleet (the individual performance of each model differs from turbine to turbine).

The standard deviations decrease almost by a factor of three, when the predictions error is averaged.

In some cases it may be difficult to identify suitable inputs to achieve a good model performance. This is for instance if a signal is not found to be well correlated to any of the other signals simultaneously measured and thus an accurate FSRC model cannot be set up. Although this was not the case for the models developed, it should be mentioned that for those signals an alternative model approach is to build autoregressive models as proposed by Zaher et al. [1] and Sanz-Bobi et al. [7].

5. Prediction errors

The interpretation of the prediction errors for fault diagnosis requires some information about the validity or the trustworthiness of the input signals to the models. Abnormal behaving input signals, consequentially cause the prediction to be abnormal. Hence it is of utmost importance to take this dependency into account. Some input signals are used by a large number of models. Examples for excessively used input signals are the power output or the nacelle temperature. In turn that means that once these signals behave abnormal, a number of models will give bad

predictions. This behavior is shown in Fig. 8, where a 2D waterfall plot of the normalized averaged percentage prediction errors over time is visualized. In this plot the colors indicate the prediction error amplitude. White areas mark periods where no prediction is available, e.g. due to missing data or non-operational periods. The prediction errors are normalized according to the following equation:

$$\bar{E}_{per} (\%) = \frac{\bar{E}}{3\sigma(\bar{E})} \times 100 \quad (4)$$

\bar{E}_{per} is the percentage error normalized by determined anomaly limit.

The color bar in Fig. 8 is set in a way that normally behaving signals are shown with a light green color. Only if the one day averaged prediction error exceeds the anomaly limit (see Section 6) other colors become visible. The color range indicates whether the prediction error is positive or negative as well as its amplitude. The prediction error of each model is shown in horizontal lines from the axis label on the left.

Fig. 8 shows that eight further models have abnormal prediction errors at the time the nacelle temperature increases. This effect influences the analysis of the root cause of the prediction error pattern. Furthermore it emphasizes the need of algorithms that take into account the context of the patterns. Here fuzzy logic is a powerful tool.

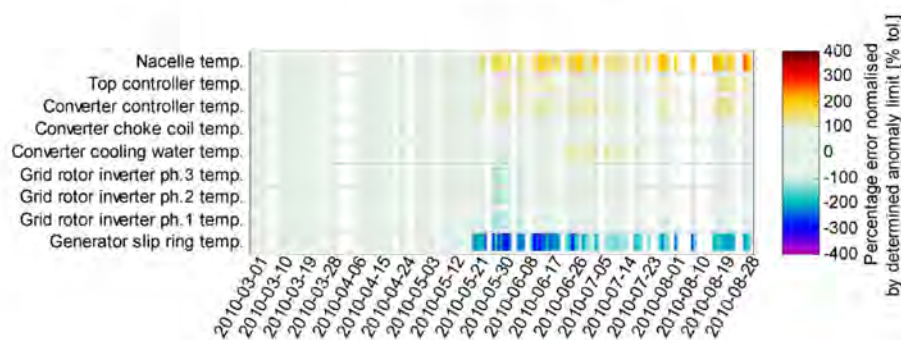


Fig. 8. 2D Waterfall plot of the normalized averaged prediction errors.

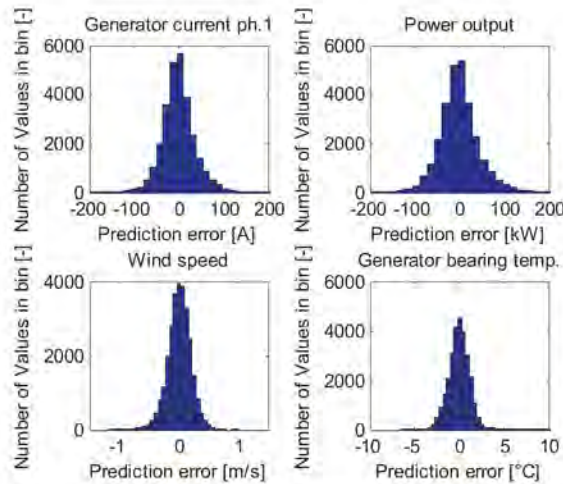


Fig. 9. Prediction error distributions after training.

6. Anomaly definition and detection

Anomaly detection refers to the problem of finding patterns in data that do not conform to expected behavior. These nonconforming patterns are among others often referred to as anomalies or outliers [28]. Detecting anomalies in the prediction error is one of the central issues of the condition monitoring approach addressed in this research. The task is to find an anomaly detection algorithm that highlights any unexpected patterns, once they exceed a certain limit. In 2009 Chandola et al. [28] reported a comprehensive review about anomaly detection algorithms. A majority of the reported techniques can be categorized into classification-based, nearest neighbor-based, clustering-based and statistical techniques [28]. In this paper the focus lies on parametric statistical techniques, because of nature of the prediction errors coming from successfully trained models that are usually normally distributed with a mean around zero. Statistical anomaly detection techniques are based on the assumption that normal data instances occur in high probability regions of a stochastic model, while anomalies occur in the low probability regions of the stochastic model [28].

Due to the availability of the prediction error in the training phase of the ANFIS models, the parameters of the underlying statistical process can be estimated. In the current CMS this is done by a maximum likelihood estimator, knowing that the parameters may be biased. Fig. 9 shows the prediction error distributions for a selection of models after training.

The estimated parameters are stored and used for estimation of the standard thresholds for anomaly detection. These thresholds form the upper and lower bounds of the normal range of the SCADA data. Instances that have a low probability of being generated from the trained model, based on the applied test statistic, are classified as anomalies [28]. For this purpose the standard thresholds are calculated according to the following probabilistic assumptions:

- Prediction errors that have a probability of being generated by the trained model larger or equal than 0.01% are considered normal
- Prediction errors that have a probability of being generated by the trained model smaller than 0.01% are considered an anomaly.

The choice of the probability influences the sensitivity of the system and can be adjusted if required. Here 0.01% was chosen to suppress false anomaly classification. Due to the large amount of turbines monitored later, the approach taken must not classify

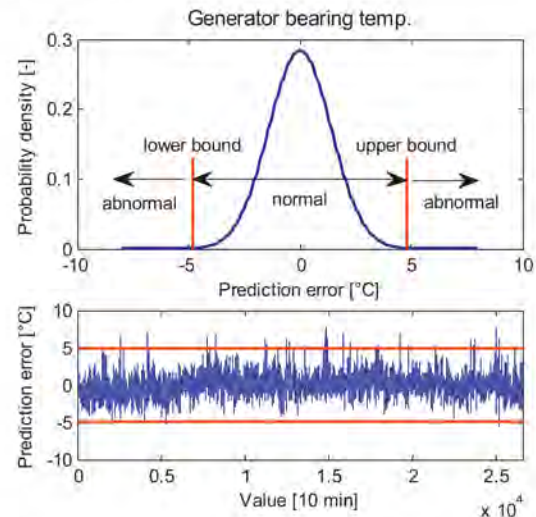


Fig. 10. Probability density distribution and evolution of the 10 min avg. prediction error (WTG 8).

normal data as abnormal. False alarms may cause extra unscheduled maintenance visits, or cause the operator to lose faith in the system and subsequently ignore indications of real problems. Misdiagnosis can result in replacement of the wrong part and additional maintenance to correct such errors [29]. The upper and lower bound is calculated with regard to the original data resolution (10 min average) and the averaged values (one day averages). Fig. 10 shows an example of the boundary between normal and abnormal behavior, both in terms of the probability density function and the prediction error in time domain.

When averaged over one day, the predictions have a reduced variance and the upper and lower bound move closer to zero. In turn this leads to less alarm limit violations as visible in Fig. 11.

To further reduce the risk of false classifications, an additional threshold is set, that requires at least three values violating the

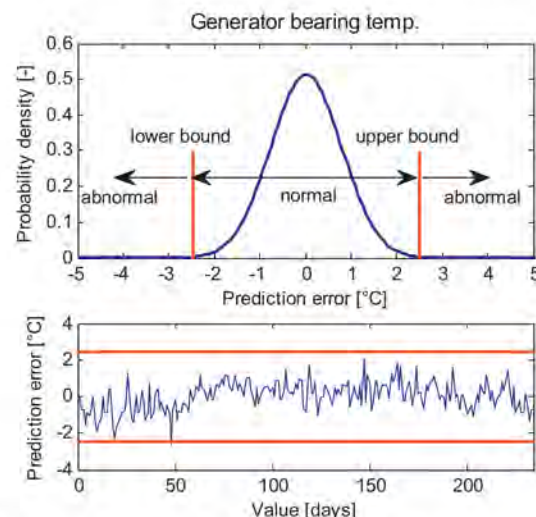


Fig. 11. Probability density distribution and evolution of the one day avg. prediction error (WTG 8).

probability threshold within a week, before the anomaly is analyzed by the fuzzy expert system.

7. Fuzzy expert system

The Fuzzy expert system consists of the fuzzy expert initialization module and the fuzzy expert application module which analyze patterns to classify the component conditions and output a potential root cause if a matching rule is found in the rule base. In total nine components or subsystems were defined, for each of which a FIS structure is developed. The components or subsystems are (see Fig. 3):

- Meteorology
- Rotor system
- Gearbox
- Generator
- Converter
- Transformer
- Auxiliary equipment
- Turbine performance
- Nacelle

7.1. FIS inputs

Only one FIS is developed per component or subsystem, hence the prediction error of several ANFIS models is processed by the FIS. The number of inputs to the FIS therefore depends on the number of models that exist to monitor the signals behavior of the component or subsystem. The relevant signals for the FIS are defined when setting up the FIS structure the first time, i.e. when the knowledge is implemented. However, the list of inputs is extendable, when rules implemented at a later stage require more inputs. The input signals are selected by considering the corresponding normal behavior model inputs and the physical understanding of the component or subsystem. For each input, a modified prediction error is passed to the input space of the FIS. The modified prediction error only contains the true prediction error value when an anomaly is present. The prediction error is set to a small number otherwise. Using the modified prediction error has the advantage that single values exceeding the anomaly limits do not cause false condition statements.

7.2. FIS membership functions

Considering which inputs are used, the fuzzy expert initialization module initializes the FIS with regard to the membership functions belonging to each modified prediction error input. Triangular and trapezoidal MFs are used in this research. A linear variation of the membership values is reasonable as long as no other dependency is known or desired. In 2008 Rodríguez and Arkio [30] proposed a diagnosis method for detection of stator winding faults in induction motors based on fuzzy logic. The system was tested with triangular, trapezoidal and Gaussian MFs. It was found that the combination of triangular and trapezoidal MFs is the most appropriated for fault diagnosis in induction motors [30]. In comparison to trapezoidal MFs, triangular MFs require one less parameter to be defined, which is why triangular MFs are preferred in this research.

It proved useful in expert knowledge implementation to specify two separate FIS inputs per relevant prediction error. The first FIS input type is with five membership functions (MFs) allowing the classification into the categories: very low, low, normal, high and very high. The second FIS input type is with three MFs allowing classification into the categories: low, normal, high, giving the expert full freedom to implement the rules. The membership

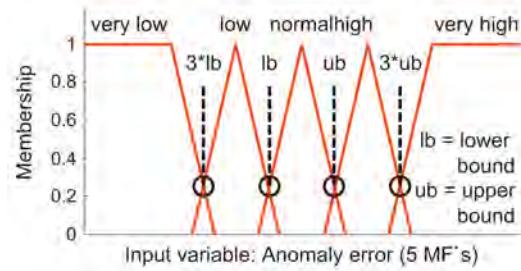


Fig. 12. Membership functions initialized for each input: FIS input type 1.

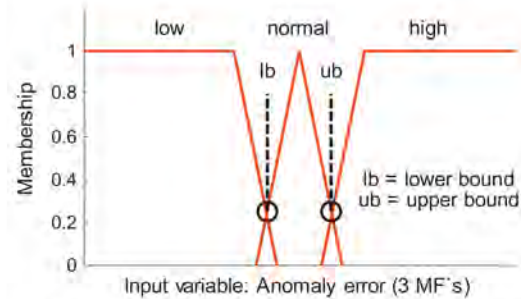


Fig. 13. Membership functions initialized for each input: FIS input type 2.

functions initialized for each of these inputs are visualized in Figs. 12 and 13.

The FIS structures are initialized each time an analysis is performed. This has the advantage that modifications of the structure inputs, rules and membership functions position and shape lead to fewer software conflicts. The definition of the statements, very low, low, normal, high and very high is given by the standard thresholds – the upper and lower bounds according to the 0.01% probability.

The standard thresholds do only serve as a first approximation for identification and classification of the component condition. These definitions are equal for all signals and turbines, but the values are individual, due to the anomaly definition given in Section 6.

In a master threshold table, the four bounds can be defined manually. This is usually done after fault occurrence, or after determination of the critical prediction error levels. Manual adoption is useful, since if for instance a model is very accurate, this consequently also leads to tight standard upper and lower bounds. Hence the prediction error may have a high membership in the MF very high, although the prediction error level is still considered to be uncritical. In this case the thresholds can be adopted in the master condition threshold file to match the real fault progression. This procedure is required, since the real bounds are unknown for most components and signals before fault occurrence. The thresholds defined as master thresholds are valid for all turbines and replace the individual ones. It is therefore possible to identify similar patterns in the whole fleet, since it is expected that the general fault characteristics are similar for turbines of the same type.

The outputs of the FIS are diverse with regard to their MFs. One of the three outputs are initialized for each FIS equally. This concerns the initialization of the output 'condition', visualized in Fig. 14.

The MF gray is reserved for periods, in which no diagnosis is possible due to missing data, whereas green, yellow and red indicate the condition status.

- Green: component/subsystem working as expected; condition good; state ok

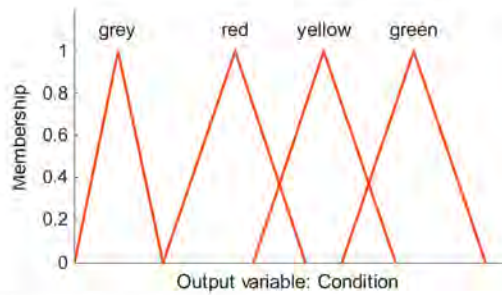


Fig. 14. Initialization of the membership functions of the variable condition.

- Yellow: service shall be scheduled for component/subsystem to improve the condition or investigate the issue; condition bad; state warning
- Red: service shall be performed at component/subsystem to improve the condition; condition very bad; state alarm

The other two outputs contain information about the diagnosis and the potential root cause. This information is individual to each FIS and is part of the knowledge that needs to be supplied by an expert. The setup is shown in Figs. 15 and 16.

7.3. FIS rules

The MFs are linked via rules that allow the FIS to output the diagnosis, condition and potential root cause. The rules are implemented by an expert manually. The concept of the rule base is that the base is constantly extended to allow the FIS to identify the various fault types accurately. The rule base is updated after fault occurrence and the validity of the already existing rules is checked. For each set of inputs, the FIS evaluates the input

memberships according to their MFs and selects the rule that is most applicable. The consecutives are then given according to this rule. All rules currently implemented employ the “and method”, taking the minimum value of the MF evaluation for decision on the rule applicability. The rules can be classified into two categories:

- generic rules
- specific rules

The generic rules are used to highlight present anomalies, in case no specific rule applies. This rule type gives no information about the specific condition or potential root cause, but highlights to the operator that an anomaly is present in the data. The specific rules on the other hand provide this level of information. Two examples of implemented specific rules are:

- (1) If (Anomaly error (5 MF's) Spinner temp.==high) & (Anomaly error (3 MF's) Nacelle temp.==ok) then (Diagnosis = Spinner temp. high) (Condition = yellow) (Pot. root cause = Sensor defect)
- (2) If (Anomaly error (5 MF's) Generator ph.1 temp. cross==high) & (Anomaly error (3 MF's) Generator ph.3 temp. cross==low) then (Diagnosis = Generator Ph.1 temp. to high) (Condition = yellow) (Pot. root cause = Generator phase insulation damage)

The rules are implemented to automate the analysis of the one day average prediction errors.

As a basis for certainty statements of the diagnosis, the result of evaluating the output values through the MFs for the best applicable rule is used. This value can be seen as a matching indicator. A high value indicates that the rule represents well the input pattern, giving a high certainty in the diagnosis.

8. Results and discussion

The developed ANFIS models show good performances at a variety of different SCADA data (see Table 5), which proves their general applicability in this context. Moreover it is shown that these models are faster to train than NN models by achieving similar or better performances. The differences in training speed are significant, when taking into account that the current development counts 45 models per turbine which require training. This is, because the failure modes of the turbines are not known beforehand. An anomaly in any of the monitored SCADA signals may indicate a potential fault or malfunction. For the eighteen turbines accessible during this research this leads to 810 different ANFIS models, which require training.

The focus in this work is on one day average predictions in order to reduce the fluctuations in the prediction error. This does not only increase the fault visibility, but it also is expected to lead to fewer false alarms. False alarms are of major concern for wind turbine operators, due to the large number of operating systems. An accurate representation of the signal normal behavior builds the basis for the subsequent analysis of occurring patterns by the FIS structure set up. The overall performance levels achieved here are expected to allow early anomaly detection.

In the fuzzy expert system the overall number of rules can be drastically reduced by specifying two separate FIS inputs per relevant prediction error as proposed in this research. The FIS input type 1 is used for the signals directly linked to the component monitored, whereas type 2 may be used for all other inputs where the fine differentiation of the prediction error magnitude is not necessary. If for instance the gearbox bearing condition shall be determined by a specific rule, it is useful using type 1 for the gearbox bearing temperature to be able to classify the condition according to

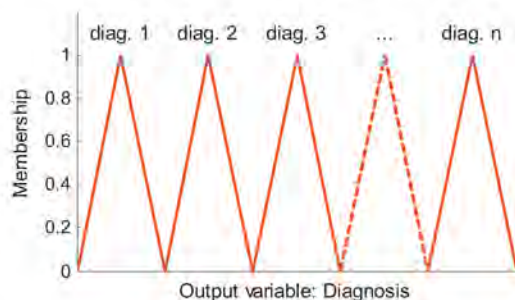


Fig. 15. Schematic of the membership function of the variables diagnosis.

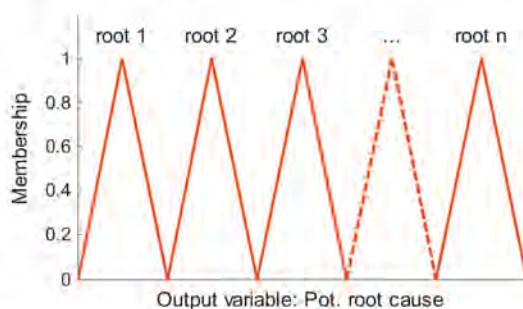


Fig. 16. Schematic of the membership function of the variables potential root cause.

the prediction error magnitude (high and very high). For all other inputs, the fine differentiation may not be necessary, making FIS type 2 for those inputs more useful. This reduces the number of possible combinations within the rule.

9. Conclusions

In this paper a method to monitor wind turbine SCADA data, via normal behavior models and fuzzy logic is proposed. The methodology gives the possibility of mining large amount of SCADA data available to wind turbine operators in a systematic manner, search for anomalies and use this information for condition statements. It allows not only monitoring of large components, but also auxiliary equipment that currently available wind turbine CMS do not cover.

Using fuzzy logic the existing expert knowledge in anomaly/prediction error pattern interpretation and root cause diagnosis can be implemented in an intuitive manner. This gives the possibility for automated fault diagnosis once specific rules are implemented.

However, the applicability of the proposed method and therefore with the achieved accuracies rely on the availability of a broad variety of different SCADA data to set up the ANFIS models. This criterion is usually fulfilled, which allows application of the proposed method to both existing older and new turbines. A further limitation concerns the availability of expert knowledge. In case no expert knowledge is available the methodology can only be used to highlight present anomalies in the SCADA data and only general statements about the condition and the diagnosis can be made.

Future research will focus on the implementation of other measurements, such as high resolution vibrational data or oil sample analysis in the CMS. This will open even more possibilities in implementing rules that capture the expert knowledge and will allow for larger diversity of the diagnoses. Furthermore turbines from a different type and brand will be included in the research.

In part two of the article the results of a field test will be presented by giving examples that show the applicability of the developed system based on real faults. It will be shown that the proposed method is capable of treating every turbine individually with regard to its characteristic. Additionally it will be shown that the FIS structures proposed allow generalizing rules that are valid for all turbines of the same type. This drastically reduces the development effort that is required to implement the expert knowledge.

References

- [1] A. Zaher, S.D.J. McArthur, D.G. Infield, Online wind turbine fault detection through automated SCADA data analysis, *Wind Energy* 12 (2009) 574–593.
- [2] R.B. Randall, *Vibration-based Condition Monitoring*, Wiley, Chichester, West Sussex, UK, 2011, ISBN 9780470747858.
- [3] W. Yang, J. Jiang, Wind turbine condition monitoring and reliability analysis by SCADA information, *IEEE* (2011) 1872–1875.
- [4] A. Kusiak, L. Wenyang, The prediction and diagnosis of wind turbine faults, *Renewable Energy* 36 (2011) 16–23.
- [5] A. Kusiak, A. Verma, A data-driven approach for monitoring blade pitch faults in wind turbines, *IEEE Transactions on Sustainable Energy* 2 (1) (2011) 87–96.
- [6] T.W. Verbruggen, *Wind Turbine Operation & Maintenance based on Condition Monitoring*, ECN-C-03-047, 2003.
- [7] M.A. Sanz-Bobi, J. del Pico, M.C. Garcia, SIMAP: Intelligent System for Predictive Maintenance Application to the health condition monitoring of a wind turbine gearbox, *Computers in Industry* 57 (2006) 552–568.
- [8] M. Schlechtingen, I.F. Santos, Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection, *Mechanical Systems and Signal Processing* 25 (5) (2011) 1849–1875.
- [9] W.G. Garlick, R. Dixon, S.J. Watson, A model-based approach to wind turbine condition monitoring using SCADA data, in: 20th International Conference on Systems Engineering, 8–10 September, Coventry University, Coventry, UK, 2009.
- [10] M. Süttmann (Schlechtingen), Master Thesis: Condition Monitoring in Wind Turbines – A Drive Train Monitoring System, 2010.
- [11] A.S. Zaher, S.D.J. McArthur, A multi-agent fault detection system for wind turbine defect recognition and diagnosis, in: *Proc. IEEE Lausanne POWERTECH*, 2007, pp. 22–27.
- [12] J. Xiang, S. Watson, Y. Liu, Smart monitoring of wind turbines using neural networks, *Sustainability in Energy and Buildings Part 1* (2009) 1–8.
- [13] J.S.R. Jang, ANFIS: adaptive-network-based fuzzy interference system, *Transactions on Systems, Man and Cybernetics* 23 (3) (1993) 665–685.
- [14] X.Z. Gao, S.J. Ovaska, Soft computing methods in motor fault diagnosis, *Applied Soft Computing* 1 (2001) 73–81.
- [15] K. Chandrasekhara, A. Chukwujekwu Okafor, Y.P. Jiang, Estimation of contact force on composite plates using impact-induced strain and neural networks, *Composite Part B* 29B (1998) 363–370.
- [16] L. Tarassenko, *Guide to Neural Computing Applications*, Elsevier, Jordan Hill, Oxford, USA, 1998, ISBN 0340705892.
- [17] M.Y. Rafiq, G. Bugmann, D.J. Easterbrook, Neural network design for engineering applications, *Computers and Structures* 79 (2001) 1541–1552.
- [18] T.J. Ross, *Fuzzy Logic with Engineering Applications*, Wiley, Chichester, West Sussex, UK, 2010, ISBN 9780470743768.
- [19] J. Kim, N. Kasabov, HyFIS: adaptive neuro-fuzzy interference systems and their application to nonlinear dynamical systems, *Neural Networks* 12 (1999) 1301–1319.
- [20] G.G. Yen, P. Meesias, An effective neuro-fuzzy paradigm for machinery condition health monitoring, *IEEE Transactions on Systems, Man and Cybernetics – Part B* 31 (4) (2001) 523–536.
- [21] E.H. Mamdani, Application of fuzzy logic to approximate reasoning using linguistic synthesis, *Transactions on Computers* 23 (12) (1977) 1182–1191.
- [22] M. Sugeno, *Industrial Applications of Fuzzy Control*, Elsevier Science Publishers, Upper Saddle River, New York, USA, 1985.
- [23] R. Leardi, A.L. González, Genetic algorithms applied to feature selection in PLS regression: how and when to use them, *Chemometrics and Intelligent Laboratory Systems* 41 (1998) 195–207.
- [24] R. Leardi, *The PLS-genetic Algorithm Toolbox for Matlab*, <http://www.models.life.ku.dk/GAPLS>
- [25] J.S.R. Jang, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, Upper Saddle River, New York, USA, 1997, ISBN 978-0132610667.
- [26] E.J. Wiggelinkhuizen, et al., *Conmow Final Report*, ECN-E-07-044, 2007.
- [27] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: a survey, *ACM Computing Surveys* 41 (3) (2009), 15:1–15:58.
- [28] R.W. Hyers, J.G. McGowan, K.L. Sullivan, J.F. Manwell, B.C. Syrett, Condition monitoring and prognosis of utility scale wind turbines, *Energy Materials* 1 (3) (2006) 187–203.
- [29] R.V.J. Rodrigues, A. Arkio, Detection of stator winding fault in induction motor using fuzzy logic, *Applied Soft Computing* 8 (2008) 1112–1120.

Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 2: Application examples

*Meik Schlechtingen

Department of Technical Operation Wind Offshore, EnBW Erneuerbare Energien GmbH, Germany
Admiralitätstr.4; 20459 Hamburg; Email: m.schlechtingen@enbw.com; Tel.: +49 40533268134

Ilmar Ferreira Santos

Department of Mechanical Engineering, Section of Solid Mechanics, Technical University of Denmark, Denmark

ARTICLE INFO

Article history:

Keywords:

ANFIS models
Condition monitoring
Wind turbine
SCADA data
Normal behavior models

ABSTRACT

This paper is the second part of a two part article series. The originality of part one was the proposal of a novelty approach for wind turbine Supervisory Control And Data Acquisition (SCADA) data mining for condition monitoring purposes. The novelty concerned the usage of Adaptive Neuro-Fuzzy Interference System (ANFIS) models in this context and the application of the proposed procedure to a wide range of different SCADA signals. The applicability of the set up ANFIS models for anomaly detection was proved by the achieved performance of the models. In combination with the Fuzzy Interference System (FIS) proposed the prediction errors can provide information about the condition of the monitored components.

The main contribution of part two lies in application examples showing the efficiency of the proposed method. The work is based on continuously measured wind turbine SCADA data from 18 modern type pitch regulated wind turbines of the 2 MW class covering a period of 35 months. Several real measured faults and issues in this data are analyzed and evaluated by the CMS and the results presented. It is shown that SCADA data contain crucial information for wind turbine operators worth extracting. Using Full Signal ReConstruction (FSRC) Adaptive Neuro-Fuzzy Interference System (ANFIS) Normal Behavior Models (NBM) in combination with Fuzzy Logic (FL) a setup is developed for mining of this information. A high degree of automation can be achieved. It is shown that FL rules established with a fault at one turbine can be applied to diagnose similar faults at other turbines automatically via the CMS proposed. A further focus in this paper lies in the process of rule optimization and adoption, allowing the expert to implement the gained knowledge in fault analysis. The fault types diagnosed here are: 1) a hydraulic oil leakage; 2) cooling system filter pollutions; 3) converter fan malfunctions; 4) anemometer offsets and 5) turbine controller malfunctions. Moreover the Graphical User Interface (GUI) developed to access, analyze and visualize the data and results is presented.

1. Introduction

Condition monitoring of wind turbines is of increasing importance as the size and remote locations of wind turbines used nowadays makes the technical availability of the turbine very crucial. Unexpected faults, especially of large components, can lead to excessive downtime offshore due to lack of suitable crane ships or other specialized vessels. However, also smaller issues and faults of auxiliary equipment like pumps or fans can cause expensive turbine downtime due to restricted turbine accessibility. From an operator's point of view it is therefore worth increasing the effort spent to monitor the turbine condition in order to reduce unscheduled downtime and thus operational costs.

The available Condition Monitoring Systems (CMS) mostly require high level knowledge about the system to be monitored. However, this knowledge is difficult to access and does often not exist. Physical models of the system to monitor its condition and predict failures can

thus seldom be built with high accuracy due to its complex interaction among several dynamical subsystems [1]. Moreover the available CMS mainly focus on vibrations. Vibration analysis is by far the most prevalent method for machine condition monitoring [2]. However, vibration sensors are not installed on all turbines and components due to their high costs. This causes a large number of turbines not being condition monitored at all or vibration sensors being installed at the main components only.

On the other hand, there is a large amount of operational (SCADA) data available, which can be used to give an indication of the turbine condition. This fact is also stressed by Yang and Jiang [3] who additionally point out that these data are the cheapest resource for developing a CMS for wind turbines.

In part one of this paper series [1], a CMS is proposed making use of Adaptive Neuro-Fuzzy Interference Systems (ANFIS) to build Normal Behavior Models (NBM) based on SCADA data. Afterwards Fuzzy Logic (FL) is used to treat the prediction error and extract information about the component conditions. Part two

will now entirely focus on application examples to show the performance of the proposed CMS. More information about the reasoning for choosing ANFIS models as well as FL can be found in part one [1].

The research presented in this paper is the result of two years investigations carried out to develop a CMS that uses wind turbine SCADA data available to wind turbine operators. The data available is obtained from 18 different operating onshore turbines of the 2MW class, where continuous operational data from April 2009 to March 2012 were gathered.

In section 2 of this paper the Graphical User Interface (GUI) developed to access and visualize the large amount of data is briefly introduced. Additionally general comments about the data volume are made and the result visualization is presented. In section 3, a flow chart is illustrated showing the information flow and fuzzy rule optimization process. Examples emphasizing the working principle and the effectiveness of the proposed method are given in section 4. Results are discussed in section 5. Section 6 focuses on future aspects and conclusions are drawn in section 7.

2. General comments, graphical user interface and visualization

The large amount of SCADA data available to wind turbine operators requires a systematic and standardized handling of data. For condition monitoring purposes not all SCADA data supplied are of interest. The first difficulty therefore arises from the selection of suitable signals for processing by the proposed CMS. The number of data tags available for a 10 min period can easily exceed 1000. This includes digital values, hour counters, calculated values, set point and statistical values (mean, minimum, maximum and standard deviation) for each 10 min period. For the reasons given in part one [1], the focus lies on 10 min average values in this research. Due to the number of data tags supplied by the turbine manufacturer 45 models are developed for each turbine in this research covering a broad variety of signals. Thus in total 810 (45×18) models are developed, trained and their prediction error monitored. For systematic handling a Graphical User Interface (GUI) was developed that eases data access, visualization and analysis. The GUI is presented in Figure 1. The functionality of the different features in the GUI are:

- A: Wind power plant selection: Allows the selection of different wind farms in the future (currently one implemented),
- B: Turbine selection
- C: Module selection (compare part one Figure 1):

- Training module: the ANFIS models are trained, or new data processed
 - Anomaly detection module: Anomalies in the prediction error are detected
 - Fuzzy expert module: evaluates existing rules and stores the results for graphical visualization
- D: Retrain panel: When the signal behavior changes due to service or replaced/ repaired components, retraining of the models may be required, which can be controlled in this panel.
- E: Advanced settings: Training and preprocessing settings can be edited to assure proper model performance
- Maximum signal lag to be removed before modeling
 - Minimum turbine power output for analysis
 - Filtration of transient situations, e.g. after turbine start
 - Number of data days and turbine operational days for the different training states
 - Probability value giving the alarm limits
 - Number of days the anomaly detection module shall wait for a further anomaly to occur before forgetting
 - Number of values to average (here 144 values equal to one day is chosen)
- F: Signal selection
- G: Normal Behavior Model (NBM) input information: Gives information about which input signals are used to build the particular model
- H: Training status information: Training states (one to four) according to the steps defined in the advanced settings panel
- I: One day average prediction error visualization: Visualization of the prediction error of the selected signal
- J: One day average model inputs prediction error visualization: Helping the expert to find the root cause of prediction error deviation from normal (only displayed, when NBM for specific input exists)
- K: Different graphic settings used for analysis:
- Anomaly error to get a quick overview about the prediction errors of the selected signals
 - 2-D prediction error to visualize the prediction error magnitude
 - Relative error as being the prediction error divided by the actual measurement in this period
 - Raw data
 - Wind turbine power curve

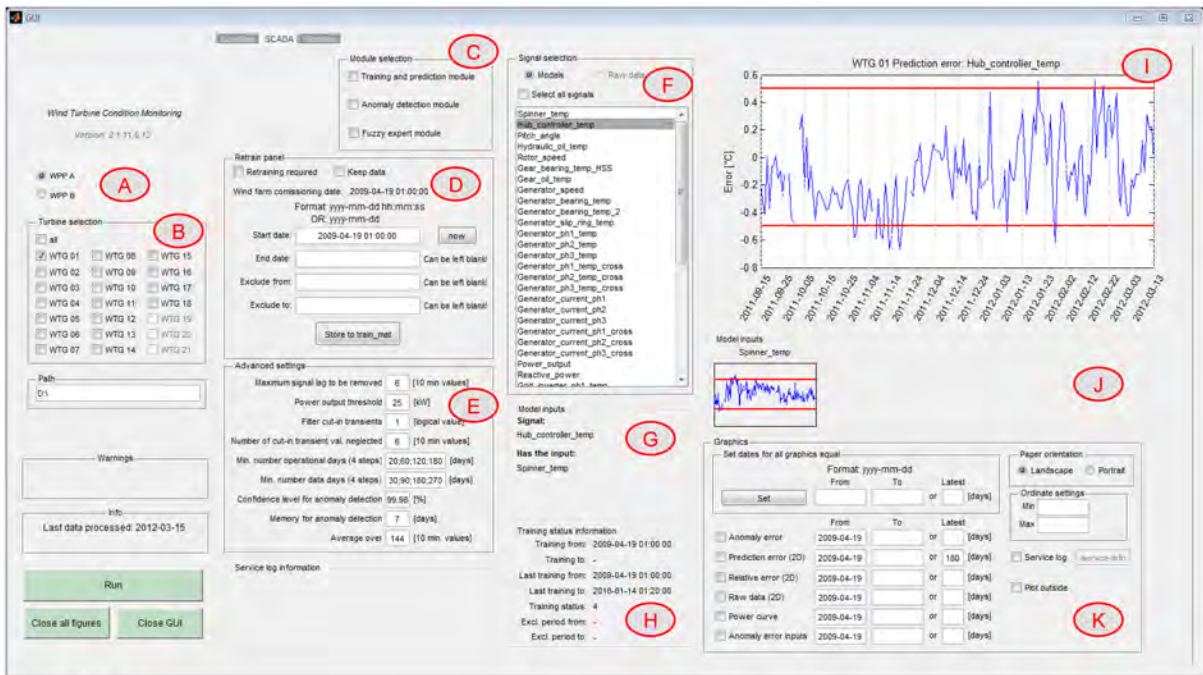


Figure 1: GUI

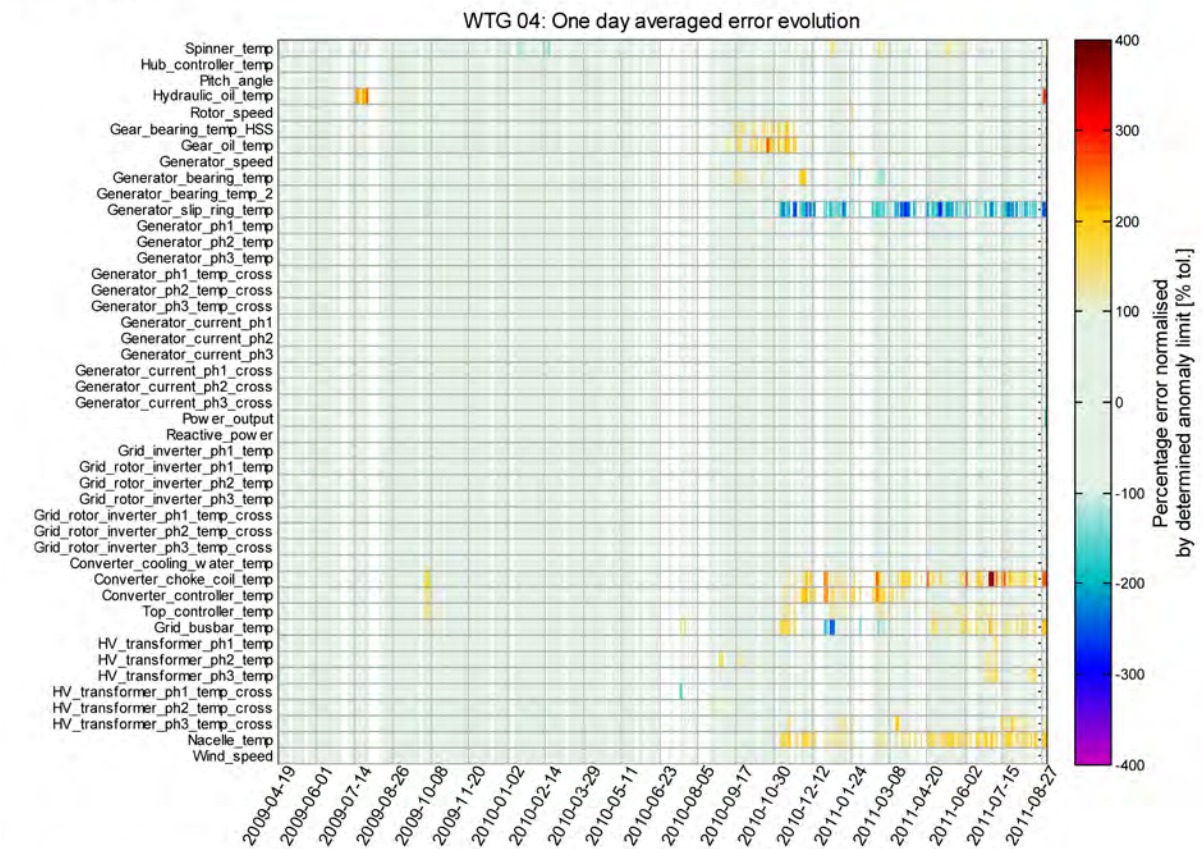


Figure 2: Percentage error normalized by the determined alarm limit (WTG 04)

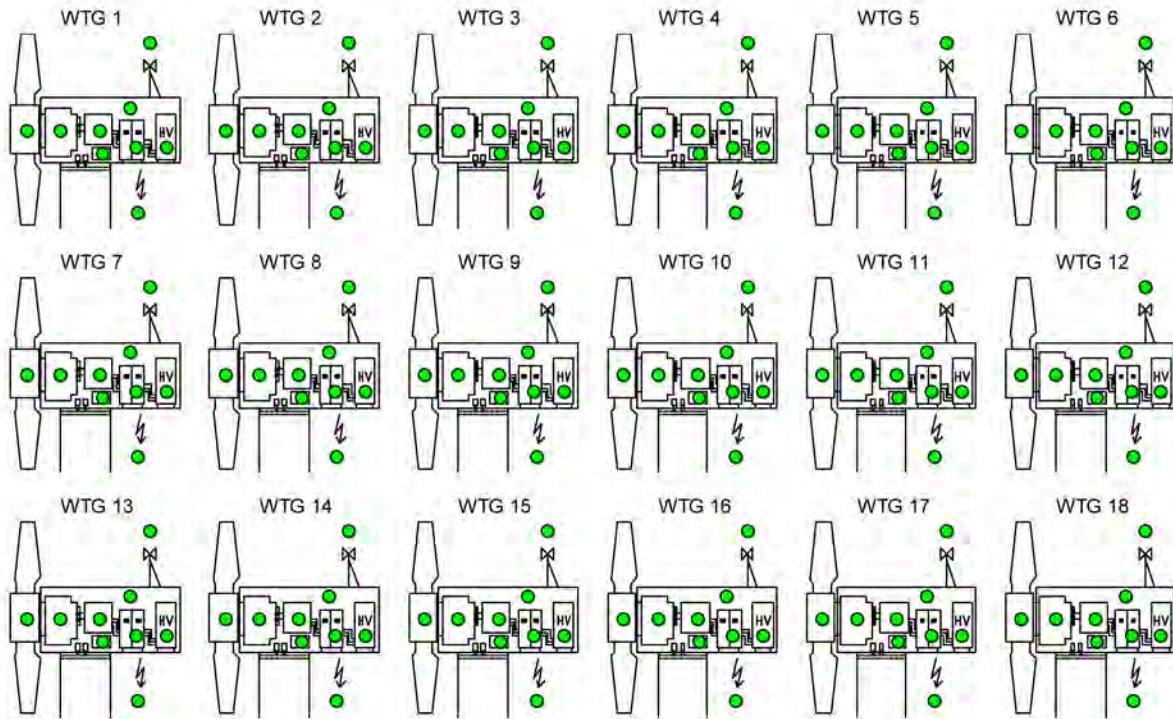


Figure 3: Condition overview

- Anomaly error inputs showing the prediction error evolution over time of the model inputs and the output

Using the anomaly error plot function of the GUI (see area K in Figure 1) the model error of all developed and monitored models per wind turbine can be visualized to get an overview of current model deviations. For this purpose the predictions errors are normalized by the determined anomaly limit. In Figure 2 a 2D waterfall plot of the normalized averaged percentage prediction errors over time is visualized for WTG 04 to emphasize the strength of this illustration in getting a one shot overview about present model deviations.

In this plot the colors indicate the prediction error amplitude. White areas mark periods where no prediction is available, e.g. due to missing data or non-operational periods

The final result of the condition evaluation (the wind power plant condition overview) via the proposed method is visualized in Figure 3 and the corresponding wind turbine schematic in Figure 4.

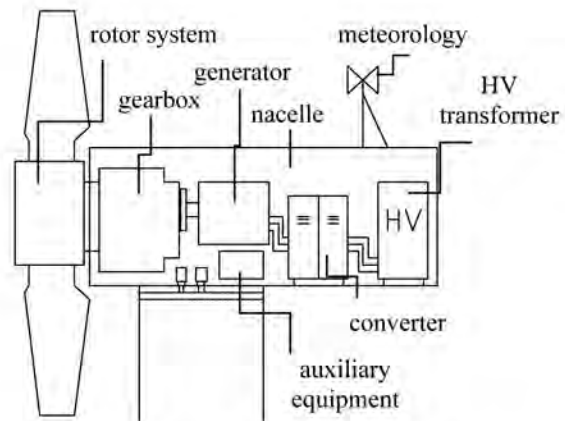


Figure 4: Wind turbine schematic: Components/subsystems in the considered wind turbines

3. Information flow

In order to arrive at component condition statements FL rules must be implemented by an expert. There are two types of rules

- generic rules
- specific rules

The generic rules are used to highlight present anomalies, in case no specific rule applies. This rule type gives no information about the specific condition or potential root cause, but highlights anomalies in the data. The specific rules on the other hand provide this level of information [1]. The overall information flow and diagnosis optimization loop is emphasized in Figure 5.

The expert analysis of the root cause is aided by the prediction errors, the raw data, service protocols, onsite inspections and personal experience. Once the root cause is identified, the generic rule can be updated, or a specific rule implemented that captures the fault pattern and leads to a correct diagnosis. Moreover the thresholds defining anomalies can be adjusted by the expert to give a better indication of the actual component condition, i.e. make the system more sensitive or insensitive. In this fashion, the expert system will become more diverse over time with regards to different diagnoses and more precise in condition classification. More information about the fuzzy expert system can be found in part one [1].

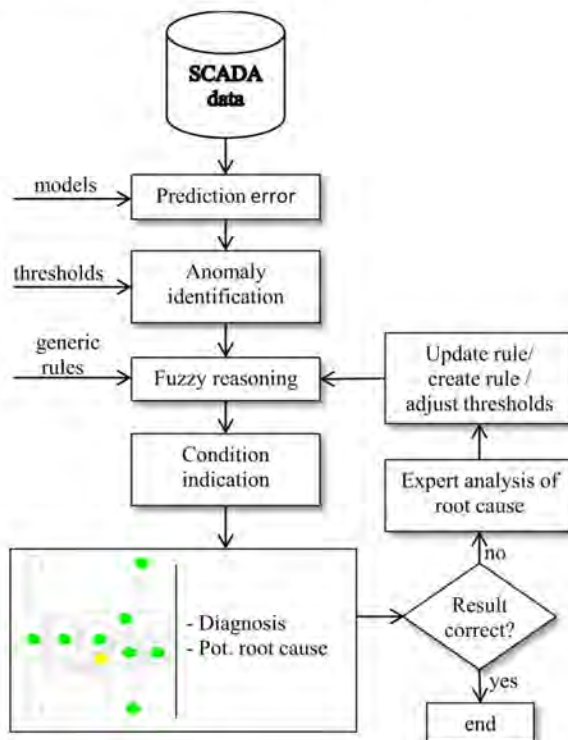


Figure 5: Information flow

Next, a number of examples will be presented which show the rule implementation process in more detail, followed by examples showing the system performance and general rule application.

4. Examples

The examples given in this section are real measured faults analyzed by the proposed CMS. Note that during CMS development no monitoring was performed, i.e. the examples given here are based on reanalysis.

4.1. Example 1: Hydraulic oil leakage

In the first example an anomaly in hydraulic oil temperature is highlighted by the CMS developed. In addition to this it is emphasized how information about the anomaly root cause can be implemented in the CMS in order to ease the analysis of future faults. Figure 6 shows a 2D waterfall plot of the normalized averaged percentage prediction errors of the hydraulic oil temperature and the related inputs of the NBM over time. The figure was generated by selecting the period between 2009-06-15 and 2009-08-14 and the anomaly error inputs graphic setting.

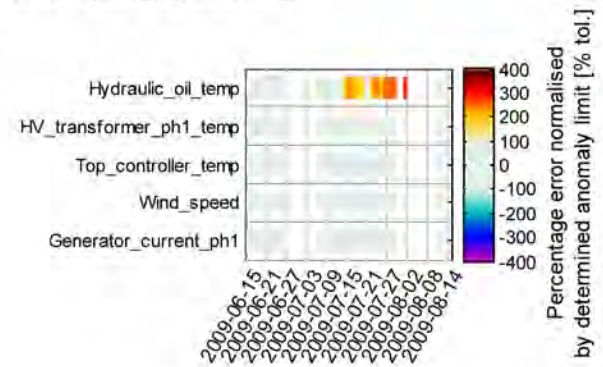


Figure 6: Normalized prediction errors of the hydraulic oil temperature and related inputs (WTG 04)

It is visible that beginning on 2009-07-16 an anomaly with rising amplitude occurs in the prediction error of the hydraulic oil temperature. At the same time the related inputs behave normal, which points to a real anomaly in the hydraulic oil temperature data.

Figure 7 shows the prediction error and Figure 8 the raw time series of the hydraulic oil temperature during anomaly occurrence over time to point out the amplitude height of the prediction error during fault occurrence.

The trend in the prediction error is clearly visible. The highest prediction error peak is around 15°C. Although the raw time series also indicates a trend, its visibility is dependent on the operational mode of the turbine.

The anomaly pattern of the normalized prediction error is analyzed by the fuzzy expert application module and a high hydraulic oil temperature is highlighted on 2009-07-16 by the CMS via the following generic rule:

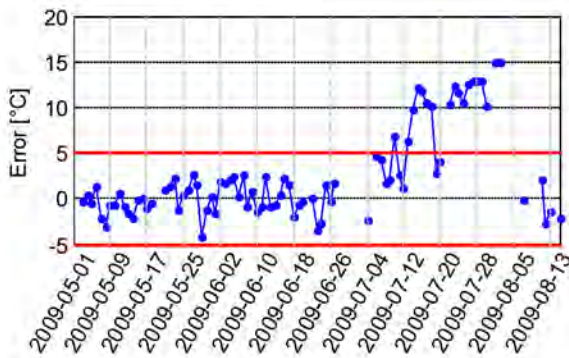


Figure 7: Hydraulic oil temperature prediction error and anomaly limits (WTG 04)

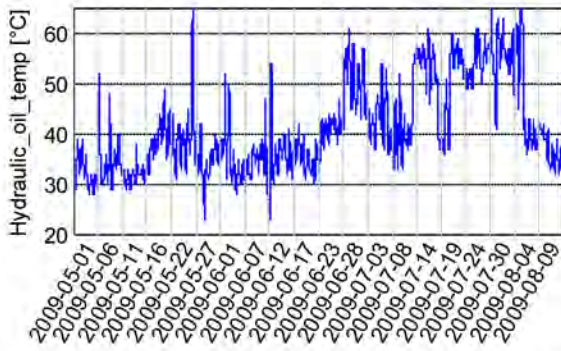


Figure 8: Hydraulic oil temperature raw time series (WTG 04)

If (Anomaly error (5 MF's) Hydraulic oil temp.=high/ very high) & (Anomaly error (3 MF's) HV transformer ph.1 temp. =ok) & (Anomaly error (3 MF's) Top controller temp. =ok) & (Anomaly error (3 MF's) Generator current ph.1 =ok) & (Anomaly error (3 MF's) Nacelle temp.=ok) then (Diagnosis=Hydraulic oil temp. high) (Condition=yellow/red)(Pot. root cause=Ambiguous)

These initial generic rules give no information about the potential root cause. The condition statements "yellow" and "red" are based on the general anomaly definition given in part one [1]. The fault can now be investigated and a rule established based on the diagnosed root cause. In this specific case, the root cause was a leakage in the rotary joint, causing the oil pump to run permanently. The entries in the service report were used to set up the specific rules containing the expert knowledge as follows.

If (Anomaly error (5 MF's) Hydraulic oil temp.=high/ very high) & (Anomaly error (3 MF's) HV transformer ph.1 temp. =ok) & (Anomaly error

(3 MF's) Top controller temp. =ok) & (Anomaly error (3 MF's) Generator current ph.1 =ok) & (Anomaly error (3 MF's) Nacelle temp.=ok) then (Diagnosis=Hydraulic oil temp. high)(Condition=yellow/red)(Pot. root cause=Pump running permanently, leakage possible)

The terms high "yellow" and very high "red" of the hydraulic oil temperature prediction error are defined in the master threshold table with 5°C and 14°C respectively, to reflect the gained knowledge of the component condition. Repair took place on 2009-08-06. Reanalyzing the fault gives the results presented in Table 1.

2009-07-15	2009-07-16	2009-08-01	2009-08-11
Component working as expected	Hydraulic oil temp. too high	Hydraulic oil temp. too high	Component working as expected
None	Pump running permanently, leakage possible	Pump running permanently, leakage possible	None

Table 1: Turbine condition evolution during hydraulic oil leakage occurrence (WTG 04)

The expert knowledge was successfully implemented and the CMS is now able to identify and diagnose similar issues on other turbines of the fleet. It is worth noting, that the generic rules can be set up before fault occurrence based on general engineering knowledge about the system behavior in fault situations.

While each turbine is treated individually with respect to the NBMs, the implemented rules are applicable for all turbines of the same type, which is shown in the next examples.

4.2. Example 2: Gearbox oil temperature increase A

The second example shows an increase in gearbox oil temperature due to dirty filters of the cooling system. Again a 2D waterfall plot is used to emphasize present anomaly patterns (see Figure 9).

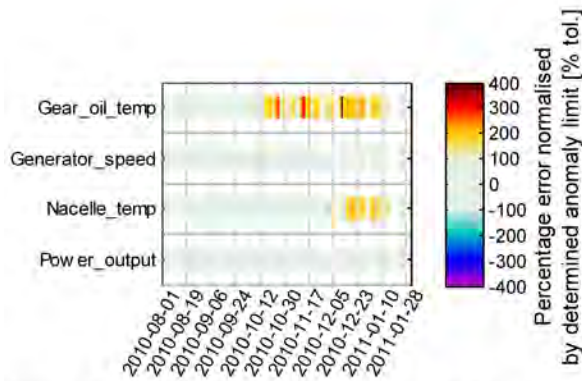


Figure 9: Normalized prediction errors of the gearbox oil temperature and related inputs (WTG 10)

It is visible that the gearbox oil temperature deviates from normal behavior round 2010-10-15, while all inputs are normal. At 2010-11-31 the input nacelle temperature leaves normal behavior and indicates a positive deviation, i.e. higher nacelle temperatures. This effect is a consequence from the increased gearbox temperature, as the gearbox together with the generator are the main heat sources in the nacelle. The prediction error magnitude and the anomaly limits of the gearbox oil temperature are visualized in Figure 10 and the corresponding raw time series in Figure 11.

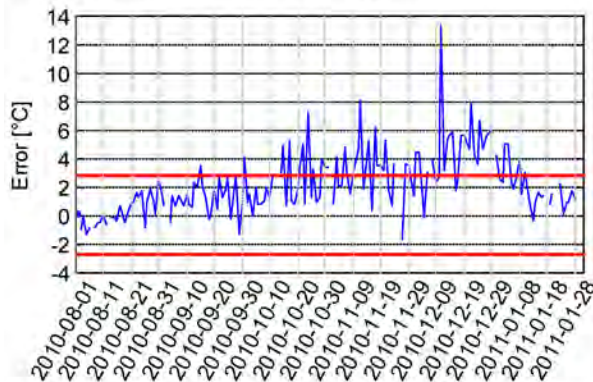


Figure 10: Gearbox oil temperature prediction error and anomaly limits (WTG 10)

The fluctuating pattern of the prediction error is typical for the deviation from normal behavior being depended on the turbine power output, i.e. high prediction error when high efficiency of the cooling system is required (full load) and low when the turbine power output is low. To some extent, this makes the correct classification into the category yellow and red depended on the power output. However, this can be overcome, if only the worst condition is considered until the component condition falls back to normal (green).

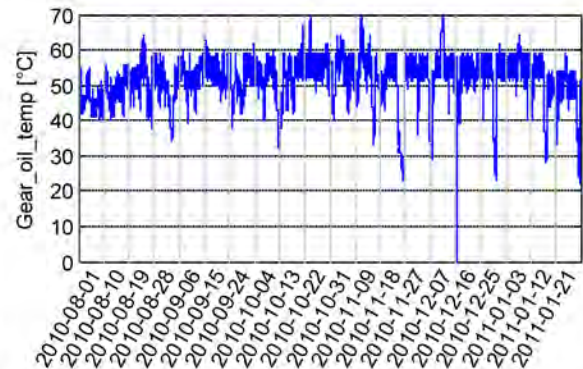


Figure 11: Gearbox oil temperature raw time series (WTG 10)

The raw data do indicate an increase in temperature, but the level of increase appears to be constant beyond mid-august.

The increase in gearbox oil temperature is highlighted by the CMS at 2010-10-15 via the generic rules:

If (Anomaly error (5 MF's) Gearbox oil temperature temp.=high/very high) & (Anomaly error (3 MF's) Generator speed=ok) & (Anomaly error (3 MF's) Nacelle temp.=ok) & (Anomaly error (3 MF's) Power output=ok) **then** (Diagnosis=Gearbox oil temp.high) (Condition=yellow/red) (Pot. root cause= Ambiguous)

The categorization into the condition category "yellow" takes place based on the standard definitions of anomaly. The condition evaluation is summarized in Table 2.

2010-10-14	2010-10-15	2010-12-11	2010-01-12
Component working as expected	Gearbox oil temp. too high	Nor rule applicable	Component working as expected
None	Ambiguous	Ambiguous	None

Table 2: Turbine condition evolution during filter pollution (WTG 04)

At 2010-12-11 the gearbox condition turned to "grey" (no diagnosis possible), as no rule in the rule base reflected the present pattern. This is because the generic rule requires the nacelle temperature to be normal. However, at 2010-12-11 the nacelle temperature model

did also show a high prediction error, which prevented the generic rule from firing. This temperature increase is because of the physical link of the gearbox temperature and nacelle temperature.

Note that during the increase in gearbox oil temperature no automatic diagnosis of the nacelle temperature can take place (indicated by the grey dot).

The service protocol reveals that the six months service was performed on the 2011-01-12 and the maintenance manual indicates that the filter was replaced, causing the normalization of the prediction error. Knowing this, the generic rules can be changed into specific rules with regards to the potential root cause and the general rule setup.

If (Anomaly error (5 MF's) Gearbox oil temperature==high/very high) & (Anomaly error (3 MF's) Generator speed==ok) & (Anomaly error (3 MF's) Nacelle temp.==low) & (Anomaly error (3 MF's) Power output==ok) **then** (Diagnosis=Gearbox oil temp.high) (Condition==yellow/red) (**Pot. root cause**== Cooling insufficient, filter polluted)

Additionally the threshold for the condition "red" was changed to 10°C in the master threshold table, in order to make condition statements equal for all turbines and to better reflect the pollution status. Reanalyzing the prediction error evolution gives the correct results as stated in Table 3.

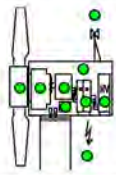
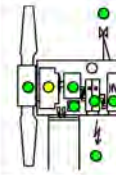
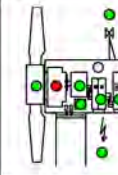
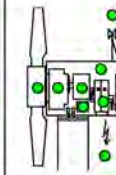
2010-10-14	2010-10-15	2010-12-11	2010-01-12
			
Component working as expected	Gearbox oil temp. too high	Gearbox oil temp. too high	Component working as expected
None	Cooling insufficient filter polluted	Cooling insufficient filter polluted	None

Table 3: Turbine condition evolution during filter pollution (WTG 04)

The filter pollution is a frequent phenomenon at this type of turbine, which allows a check of rule validity for other turbines. This check is done in the next example.

4.3. Example 3: Gearbox oil temperature increase B-G

In this example the specific rules implemented in example 2, are applied to the same issue at other turbines. Figure 12 shows the 2-D waterfall plot during filter pollution at WTG 09.

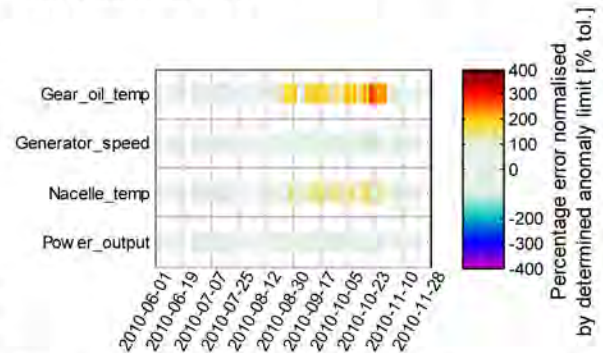


Figure 12: Normalized prediction errors of the gearbox oil temperature and related inputs (WTG 09)

The first anomaly in gear oil temp. occurs at 2010-08-20. Six days later the nacelle temperature leaves the normal operational range. Figure 13 shows the prediction error amplitudes and Figure 14 the raw time series of the gearbox oil temperature during filter pollution.

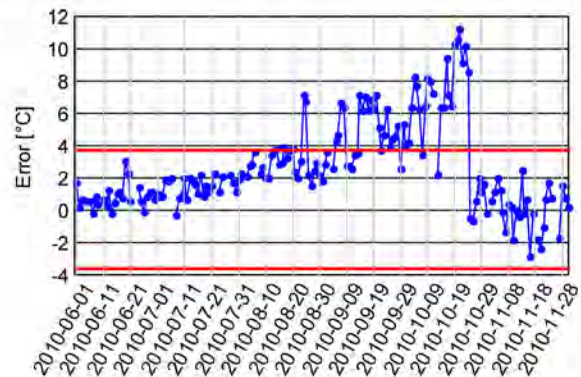


Figure 13: Gear oil temperature prediction error and anomaly limits (WTG 09)

Although the temperature deviation from normal behavior is up to 11°C the increase was not highlighted by the turbine controller, due to usually high thresholds set. After filter replacement the temperature fell back to the normal operational range.

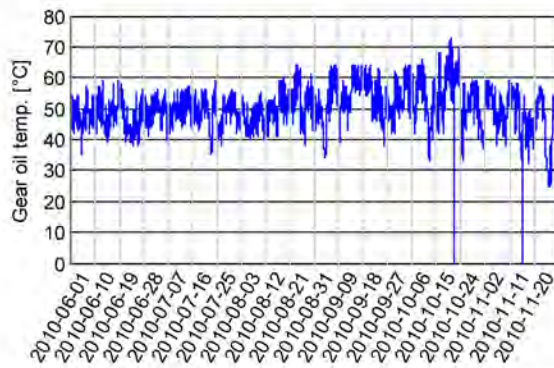


Figure 14: Gear oil temperature raw time series (WTG 09)

The trend is also visible in the raw time series, but the fluctuations due to the different operational modes make it more difficult to identify the anomaly. Note that setting simple thresholds to the raw signal would make the identification of the increase dependent on the operational mode. A partially loaded turbine would thus delay the fault discovery, which shows the effectiveness of the proposed method.

The evolution of the gearbox condition during filter degradation is visualized in Table 4. The filter was replaced on 2010-10-25.

2010-08-19	2010-08-20	2010-10-19	2010-10-26
Component working as expected	Gearbox oil temp. too high	Gearbox oil temp. too high	Component working as expected
None	Cooling insufficient filter polluted	Cooling insufficient filter polluted	None

Table 4: Turbine condition evolution during gearbox oil temperature increase due to dirty filters (WTG 09)

The condition was correctly classified. There were six other turbines detected with the same root cause. The result of the condition evaluation is presented in Table 5 and the corresponding prediction errors of the gearbox oil temperature in Figure 15.

The pattern present between the different turbines is similar. The set threshold for the condition category “red” (10°C) is acceptable as this category should be used only when action is required.

WTG	Yellow	Red	Filter change	Max div.
02	2010-08-24	2010-10-19	2010-11-23	15°C
06	2010-09-14	-	2010-10-15	5.8°C
12	2010-09-08	-	2010-11-17	7.8°C
14	2010-07-15	2010-10-01	2010-11-10	13.1°C
17	2010-10-17	-	2011-01-28	8.1°C
18	2010-08-23	-	2010-10-12	6.6°C

Table 5: Condition evaluation summary for the polluted filter phenomenon

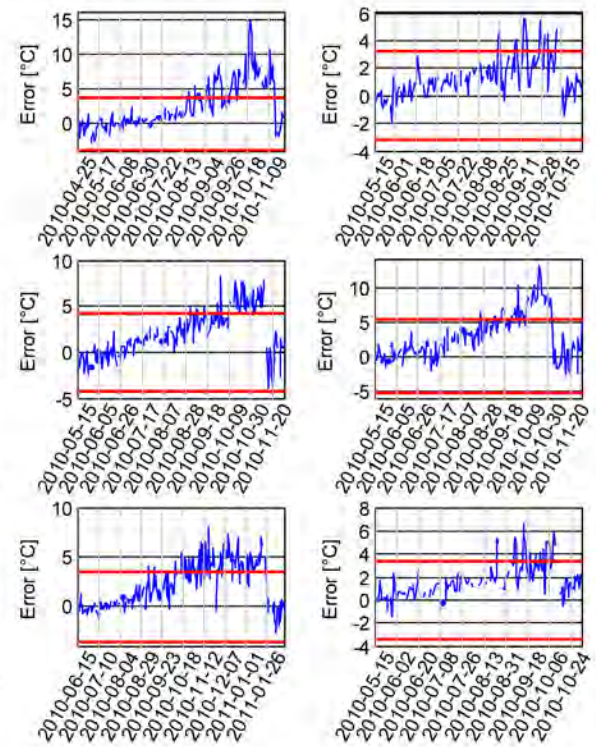


Figure 15: Prediction errors during filter pollution (from top left to bottom right): WTG 02, WTG 06; WTG 12; WTG 14; WTG 17; WTG 18

It should be noted that although each turbine is treated individually with regards to the signal behaviour (each turbine has its own model and model accuracy) the categorization into the condition categories is possible. Due to the different model accuracies the category “yellow” is different for each turbine, since it has not been defined in the master threshold table.

4.4. Example 4: Converter fan malfunction A

Modern type wind turbines operate at variable speed. In order to be able to connect to the grid, a stable frequency is required, equal to the grid frequency.

In wind turbines this is achieved via a frequency converter (AC-DC-AC). It is thus a central component of the turbine, without which no power can be supplied. In following example a relay failure caused the converter temperature to increase, since the fan was not working. Figure 16 shows the 2-D waterfall plot of the converter choke coil temperature and the related inputs and Figure 17 analogously a similar plot of the converter controller temperature.

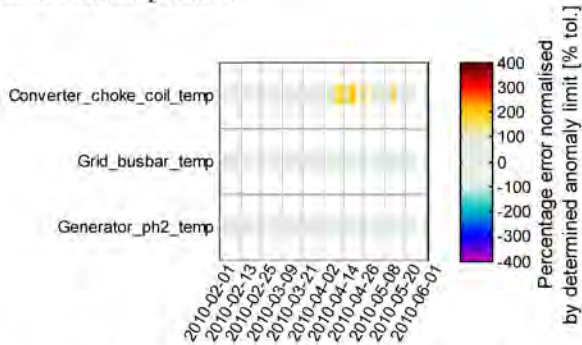


Figure 16: Normalized prediction errors of the Converter choke coil temperature and related inputs (WTG 14)

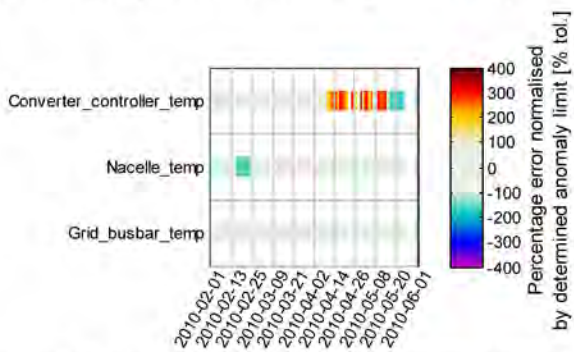


Figure 17: Normalized prediction errors of the Converter controller temperature and related inputs (WTG 14)

Both models show a positive model deviation, while the related inputs to the models behave normal. Note that due to the different thresholds for the models the anomaly is pronounced differently, i.e. the percentage error normalized is higher for the controller temperature because the model is more accurate. Figure 18 and Figure 19 show the prediction error evolution of the two temperatures, and Figure 20 and Figure 21 the raw time series.

In the prediction errors, the anomalies are clearly visible, while they are practically invisible in the raw time series. At 2010-05-07 the turbine had an error due to high temperature of the top controller (master controller in the nacelle from which all the turbine subsystems are controlled and monitored).

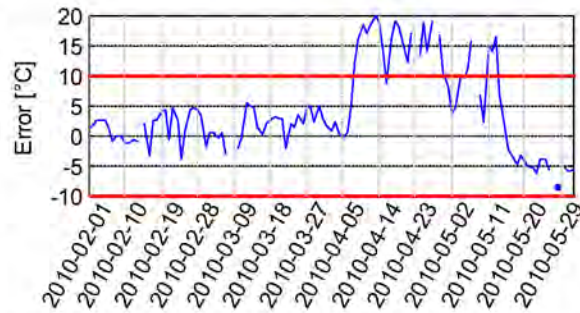


Figure 18: Prediction error converter choke coil temperature WTG 14

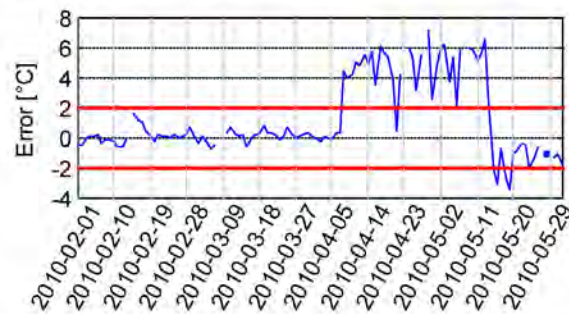


Figure 19: Prediction error converter controller temperature WTG 14

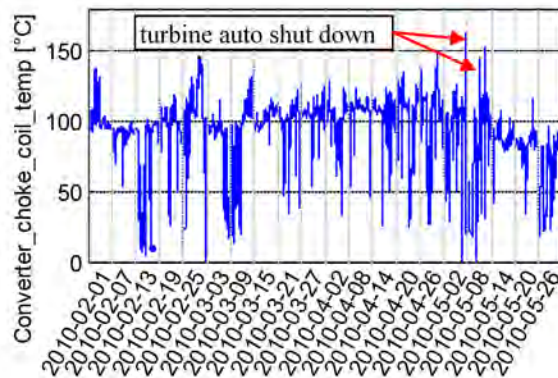


Figure 20: raw time series of the converter choke coil temp. WTG 14

The turbine restarted based on auto reset one hour later. However, due to dying wind speeds the error did not occur again until 2010-05-11 where the turbine was shut down for 6 hours. One day later the error triggered the third time and a service team was sent out and identified the root cause to be a defective relay. Another seven hours of turbine downtime were generated.

The two temperatures considered here by the CMS belong to the same component "Converter".

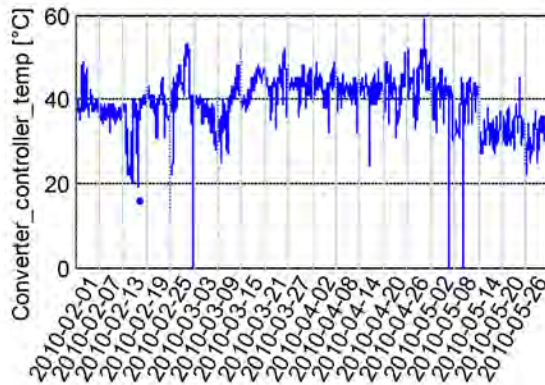


Figure 21: raw time series of the converter choke coil temp. WTG 14

Initially there were four generic rules established to highlight anomalies in the prediction error. They were:

If (Anomaly error (5 MF's) Converter choke coil temp.=high/very high) & (Anomaly error (3 MF's) Grid busbar temp. =ok) & (Anomaly error (3 MF's) Generator ph.2 temp. =ok) **then** (Diagnosis=Converter choke coil temp. high) (Condition=yellow/red) (Pot. root cause=Ambiguous)

If (Anomaly error (5 MF's) Converter controller temp. =high/very high) & (Anomaly error (3 MF's) Nacelle temp. =ok) & (Anomaly error (3 MF's) Grid busbar temp. =ok) **then** (Diagnosis=Converter controller temp. high) (Condition=yellow/red) (Pot. root cause=Ambiguous)

Since both temperatures behave abnormal during this fault and only one condition can be assigned to the component the rules must be weighted in order in order to track the right root cause. Since the converter choke coil temperature level is much higher than the temperature level of the converter controller, heat is transferred from the choke coil to the controller. Hence in case both temperatures show an increase, it is more likely that the choke coil is the cause, as an increase in controller temperature is unlikely to be noticed in the choke coil temperature. For this reason the weights of the rules treating the converter controller temperature were set to 0.9, whereas the weights of the rules treating the converter choke coil temperature were kept by 1. Table 6 summarizes the CMS result.

The condition category "red" was not triggered during fault occurrence, which is caused by the high alarm limits and the comparably inaccurate converter choke coil temperature model (the threshold for "red" lies about 30°C).

2010-04-07	2010-04-08	2010-05-13
Component working as expected	Converter choke coil temp. high	Component working as expected
None	Ambiguous	None

Table 6: Turbine condition evolution during fan malfunction increase due to dirty filters (WTG 14)

The gained knowledge can be implemented in terms of specific rules:

If (Anomaly error (5 MF's) Converter choke coil temp.=high/very high) & (Anomaly error (3 MF's) Grid busbar temp. =ok) & (Anomaly error (3 MF's) Converter controller temp. =high) & (Anomaly error (3 MF's) Generator ph.2 temp. =ok) **then** (Diagnosis=Converter choke coil temp. high) (Condition=yellow/red) (Pot. root cause=Internal VCS fan not working/ cooling insufficient)

Based on the experienced fault pattern and the prediction error evolution, the threshold for the condition category "red" was changed to 15°C in the master threshold table. Reanalyzing the fault gives the result presented in Table 7.

2010-04-07	2010-04-08	2010-04-09	2010-05-13
Component working as expected	Converter choke coil temp. high	Converter choke coil temp. high	Component working as expected
None	Internal VCS fan not working/ cooling insufficient	Internal VCS fan not working/ cooling insufficient	None

Table 7: Turbine condition evolution during fan malfunction (WTG 14)

The period in which the category “yellow” is active is only one day. This is acceptable, since the temperature increase is sharp and a temperature increase of 15°C is considered serious and should initiate investigation.

4.5. Example 5: Converter fan malfunction B, C, D & E

Using the same rules further converter fan malfunctions were diagnosed by the CMS at four other turbines. The condition evaluation is summarized in Table 8. Figure 22 shows the relevant prediction errors during anomaly occurrence.

WTG	Yellow	Red	Issue resolved	Max div.
03	2010-05-22	2010-11-28	2011-01-11	40.2°C
08	2010-12-19	2011-01-15	2011-04-13	25.0°C
12	2009-09-10	-	2009-10-07	13.1°C
18	2010-11-08	2010-12-21	2011-04-13	38.8°C

Table 8: Condition evaluation summary for the polluted filter phenomenon

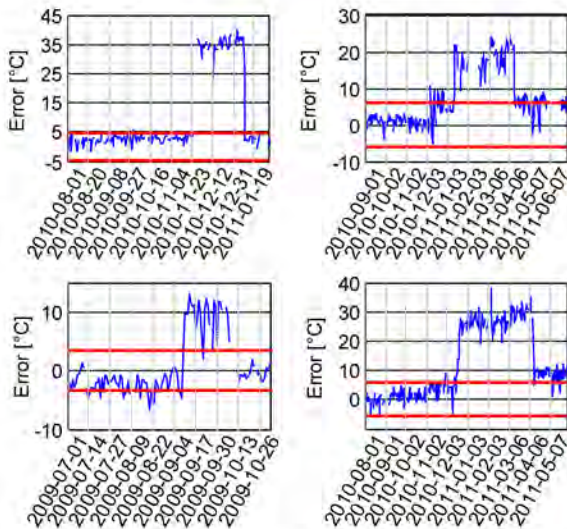


Figure 22: Prediction errors of the converter choke coil temperature during converter fan malfunction (from top left to bottom right): WTG 03, WTG 08; WTG 12; WTG 18;

Even though the turbine shuts down when a high temperature is detected, the turbine restarts automatically, usually two times before the auto restart is suppressed. Latest after the third time of high temperature, a service action is required to resolve the issue. Using the proposed CMS the issue can be detected right away. The time until a service action is required and the turbine no longer auto restarting can range

between one to eight months (for the examples given here).

For WTG 18 the temperature level after the fan was replaced does not fall back to its normal range. This can be either due to a different setting of the thermostat, a lasting damage of the choke coil or a different fan being used. Figure 23 illustrates the raw time series of the converter choke coil temperature, indicating the maximum temperatures during fan malfunction and the temperature level after fan replacement.

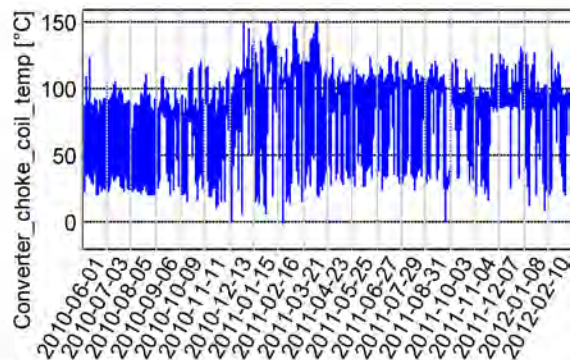


Figure 23: Raw time series of the converter choke coil temp. (WTG 18)

The maximum temperature of 150°C occurred can cause the insulation of the coil to age faster, reducing the overall component live time. In order to not trigger alarms permanently, the condition thresholds for this particular turbine can be adopted to account for the persistent offset. A different approach would be to retrain the model, or to establish a offset in the prediction error. However, the latter bears the disadvantage of tracking the established offsets. Furthermore the prediction error appears to be normal although degradation is present. The same is true for retraining in those situations. Hence the preferred method is the adjustment of thresholds. Note that the permanent offset in choke coil temperature can influence other models which use this signal as input.

Another interesting issue arising from Figure 22 in combination with Table 8 is that for WTG 03 the CMS highlights the category “yellow” at 2010-05-22. However, the hard boundaries (red lines) are not violated before the step increase (2010-11-28). Here the advantage of using FL instead of crisp logic becomes visible. Due to the overlap in membership functions, diagnosis is still possible, although in a crisp sense the temperature still behaves normal (within the boundaries).

Figure 24 illustrates the converter choke coil membership functions for the converter Fuzzy

Interference System (FIS), including the thresholds for condition evaluation.

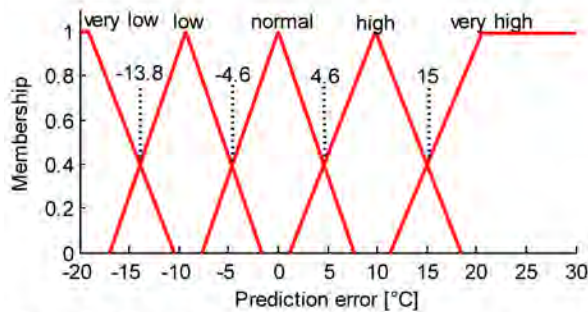


Figure 24: Converter choke coil membership functions for the converter FIS (WTG 03)

The prediction error prior 2010-11-28 is just below 4.6°C. At this prediction error level the membership function “high” already contributes some membership larger than zero to the evaluation of rules. Since at this time already the converter controller temperature is abnormal (see Figure 25), the rule highlighting the issue fires.

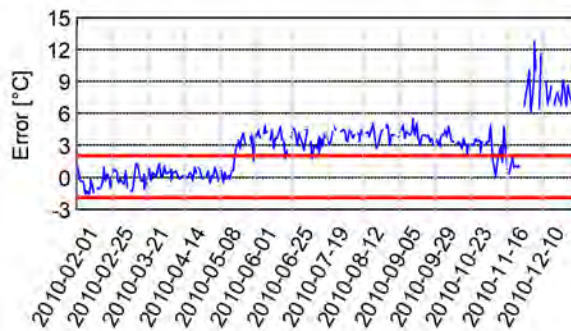


Figure 25: Prediction error of the converter controller temperature during fan malfunction

4.6. Example 6: Anemometer offsets

In this examples anemometer offsets are detected at all turbines. Anemometers are used to determine the cut-in and in some cases also the cut-out wind speed. However, the information coming from the anemometers is useful for the overall turbine performance evaluation in terms of the wind turbine power curve, i.e. the turbine power output over the wind speed.

A typical power curve of a pitch regulated wind turbine is shown in Figure 26.

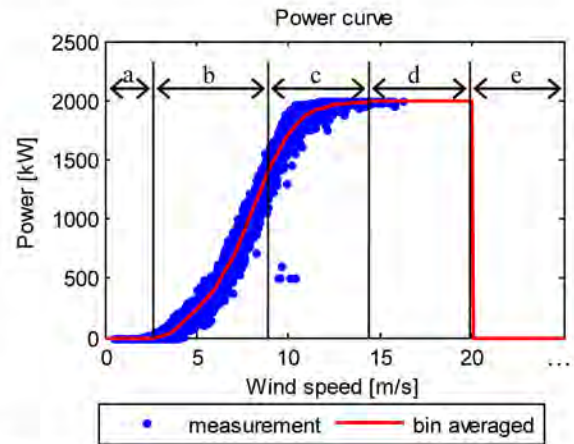


Figure 26: Power curve properties

As illustrated in Figure 26, the power curve can be divided into five regions:

- wind speed < cut in speed – the cut in speed is the wind speed value at which the turbine starts operating; power output zero or negative; pitch angle about 90°
- wind speed > cut in speed < rated speed; power output according to eq.1; Pitch angle between 0...-5°
- wind speed around rated speed; power output according to eq.1; Pitch angle between -5...5°
- wind speed > rated speed; power output according to eq.1; Pitch angle between 5...25°
- wind speed > cut-out speed; power output zero or negative; pitch angle about 90°

If the wind speed is measured incorrect, secondary performance evaluation via the power curve is impossible or leads to wrong results. Hence, it is important to track, if the wind speeds measured fit to the measured power. Figure 27 shows the normalized prediction errors of the wind speed and the related inputs of WTG 01 exemplary and Figure 28 the wind speed prediction error during offset occurrence.

The wind speed prediction error shows a negative deviation, whereas the power output shows a positive deviation at the same time. This correlation is caused by the fact that these two models use each other's signal as input, i.e. the wind speed model uses the power output and the power output the wind speed as input. This leads to the non-ideal situation, that in case of anomaly, both models will show abnormal prediction errors. However, given the set of available possible input variables, this could not be avoided.

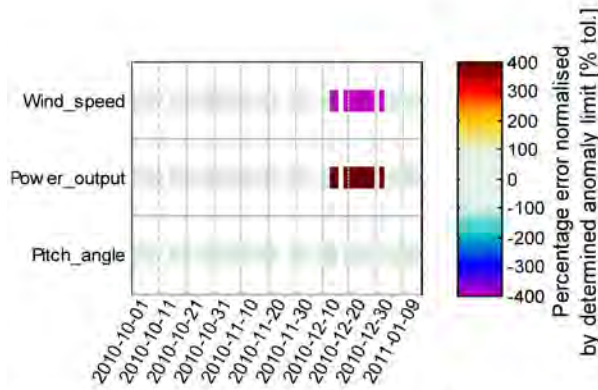


Figure 27: Normalized prediction errors of the wind speed and related inputs (WTG 01)

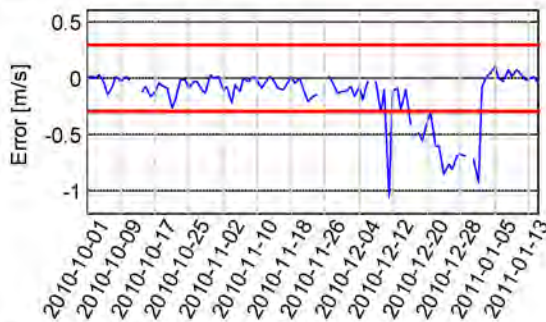


Figure 28: Wind speed prediction error during offset occurrence (WTG 01)

Taking into account that a positive power output prediction error means a better than normal turbine performance the root cause can be identified, as this appears less likely than a performance decrease with lower power output. On the other hand a positive wind speed model deviation would mean a higher wind speed measurement than usual, which in terms of a classical cup anemometer appears unlikely, too. Hence the generic rules to identify the likely cause of anomaly are:

If (Anomaly error (5 MF's) Wind speed=*low/very low*) & (Anomaly error (3 MF's) power output=*high*) & (Anomaly error (3 MF's) pitch angle=*ok*) **then** (Diagnosis=*Wind speed measured too low*) (Condition=*yellow/red*) (Pot. root cause=*Ambiguous*)

If (Anomaly error (5 MF's) power output = *low/very low*) & (Anomaly error (3 MF's) Wind speed=*low*) & (Anomaly error (3 MF's) pitch angle=*ok*) **then** (Diagnosis=*Power output too low*) (Condition=*yellow/red*) (Pot. root cause=*Ambiguous*)

Applying these two rules, the CMS gives the result presented in Table 9.

	2010-12-13	2010-12-14	2010-12-19	2010-12-25
	Component working as expected	Wind speed measured too low	Wind speed measured too low	Component working as expected
	None	Ambiguous	Ambiguous	None

Table 9: Turbine condition evolution during anemometer offset (WTG 14)

The cause of anomaly was an offset set initiated by after a software update mid December 2010. The offset was about -0.7 m/s.

The issue was also highlighted at the other turbines as summarized in Table 10.

WTG	Yellow	Red	Retraining	Max div.
02	-	-	2010-12-25	-1.1m/s
03	-	2010-12-19	2010-12-25	-1.1m/s
04	2010-12-22	2010-12-24	2010-12-25	-0.9m/s
05	-	2010-12-19	2010-12-25	-1.0m/s
06	2010-12-18	2010-12-24	2010-12-25	-0.9m/s
07	2010-12-19	2010-12-24	2010-12-25	-0.7m/s
08	2010-12-20	2010-12-22	2010-12-25	-0.8m/s
09	-	2010-12-18	2010-12-25	-0.8m/s
10	-	2010-12-16	2010-12-25	-0.8m/s
11	2010-12-14	2010-12-15	2010-12-25	-1.0m/s
12	-	2010-12-19	2010-12-25	-1.0m/s
13	-	2010-12-16	2010-12-25	-1.8m/s
14	2010-12-18	2010-12-24	2010-12-25	-0.9m/s
15	-	2010-12-16	2010-12-25	-0.8m/s
16	2010-12-14	2010-12-15	2010-12-25	-1.1m/s
17	-	2010-12-19	2010-12-25	-0.9m/s
18	-	-	2010-12-25	-1.0m/s

Table 10: Condition evaluation summary during anemometer offset

At WTG 02 and WTG 18, no diagnosis was possible, as the wind direction signal was implausible after the software update. This made power output predictions impossible. In turn the rules involving the power output predictions could not be evaluated. In order to account for the offset, all models taking the wind speed as input

as well as the wind speed model itself were retrained, since removal of the offset could not be forced.

4.7. Example 7: Turbine controller malfunction

The last example shows a turbine master controller malfunction- during which a dynamical set point caused the turbine power output to decrease. The dynamical set point made the turbine to reduce its power output even below rated power (region b and c in Figure 26). In the consequence the wind turbine power curve became broader. The reduction was non-constant, which made detection of the malfunction via visual observation of the power curve difficult. In Figure 29 three example power curves during healthy state are depicted and in Figure 30 the power curve during controller malfunction.

The malfunction happened after a controller software update at all WTGs. In the normalized prediction error (Figure 31) as well as the prediction error (Figure 32), the malfunction is clearly identifiable. Note that the inputs: ambient temperature and wind direction are not plotted in Figure 31, as for these signals no NBM was developed.

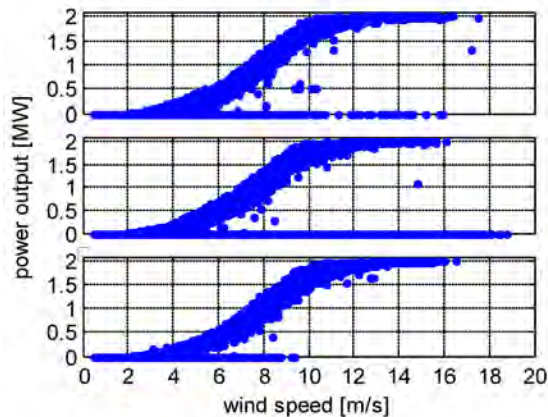


Figure 29: Measured power curve in a healthy state of three example wind turbines

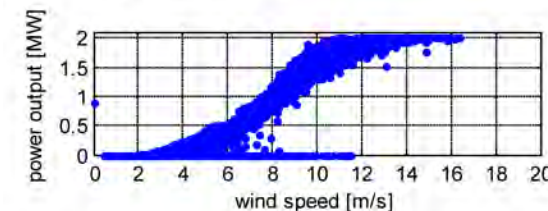


Figure 30: Measured power curve during controller malfunction WTG 01

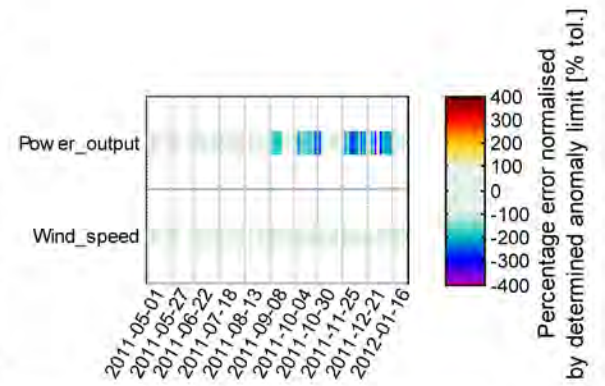


Figure 31: Normalized prediction errors of the power output and related inputs (WTG 01)

The turbine power output is crucial for the cost of energy. It is thus of interest to monitor the turbine power output closely. Therefore the thresholds for the condition category “yellow” were set to $\pm 60\text{kW}$ and “red” to $\pm 120\text{kW}$ deviation. Based on these thresholds and the rules stated in section 4.6, the CMS highlights the issue at the 2011-09-08. An overview about the CMS condition evaluation during the malfunction is given in Table 11.

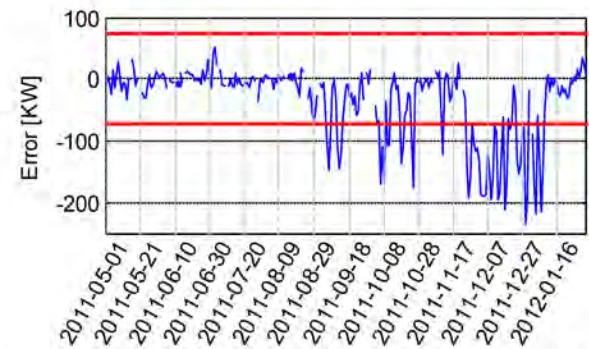


Figure 32: Power output prediction error during controller malfunction (WTG 01)

2011-09-07	2011-09-08	2011-09-13	2011-01-16
Component working as expected	Power output too low	Power output too low	Component working as expected
None	Ambiguous	Ambiguous	None

Table 11: Turbine condition evolution during turbine controller malfunction (WTG 01)

The condition evaluation at the other turbines is summarized in Table 12.

WTG	Yellow	Red	Issue resolved	Max div.
02	-	-	2012-01-16	-234kW
03	2011-10-11	2011-10-25	2012-01-16	-200kW
04	2011-08-31	2011-09-06	2012-01-16	-354kW
05	-	2011-09-06	2012-01-16	-323kW
06	-	2011-09-06	2012-01-16	-410kW
07	-	2011-09-07	2012-01-16	-296kW
08	-	2011-10-10	2012-01-16	-257kW
09	-	2011-09-07	2012-01-16	-262kW
10	2011-09-06	2011-09-07	2012-01-16	-287kW
11	2011-09-12	2011-12-03	2012-01-16	-218kW
12	2011-09-13	2011-12-03	2012-01-16	-214kW
13	-	2011-09-07	2012-01-16	-204kW
14	-	2011-09-07	2012-01-16	-198kW
15	2011-09-03	2011-09-06	2012-01-16	-243kW
16	2011-08-28	2011-09-07	2012-01-16	-241kW
17	2011-09-12	2011-10-11	2012-01-16	-318kW
18	-	-	2012-01-16	-244kW

Table 12: Condition evaluation summary during controller malfunction

Once more for WTG 02 and WTG 18 no diagnosis was possible due to implausible wind direction signals after a controller software update. Dependent on the wind resource at the particular turbine location the issue is highlighted at different dates and with different condition categories. Figure 30 shows that the power curve deviation is particular large for wind speeds between 9 and 14m/s. At the site and time of year the turbines do not operate in this wind speed range very often, which can cause some delay in diagnosis. However, the performance of the CMS is satisfactory, since the deviation from normal behavior in other wind speed ranges was small as indicated by the fluctuating prediction error pattern. Figure 33 shows the prediction error for WTG 03, where the issue was detected comparably late.

It can be seen that the prediction error violates the anomaly limit two times between 2011-09-01 and 2011-09-15 before it returns to its normal range again. However, currently the CMS is set up to highlight the anomaly only when the anomaly limit is violated minimum three times within seven days. Beginning of October the prediction error magnitude increases for a longer period and the issue is highlighted. This case points to the difficulty of the tradeoff between preventing false alarms and early anomaly detection.

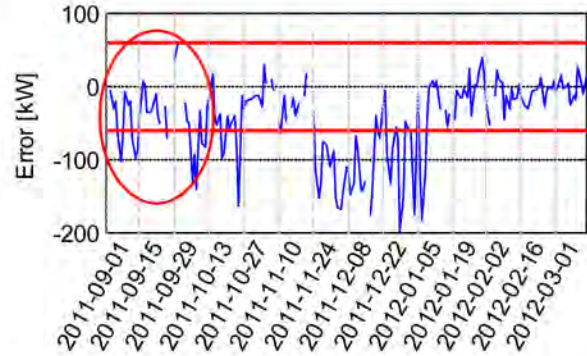


Figure 33: Power output prediction error (WTG 03)

5. Results and discussion

The results presented show that a number of different failures and malfunctions can be identified using the proposed method, combining ANFIS NBM and FL. In contrast to Sanz-Bobi, Garcia and del Pico [4] a Full Signal ReConstruction (FSRC) model approach is used to set up the normal behavior models. In earlier research of the authors [5] this type of models were found to detect faults earlier as well as allowing a broader anomaly spectrum to be detected. For instance with autoregressive models as proposed by Sanz-Bobi, Garcia and del Pico [4] as well as Zaher, McArthur and Infield [6] mostly relative changes in signal behavior can be detected, whereas with FSRC models as proposed by Schlechtingen and Santos [5] also absolute changes in level are detectable. In this research the FSRC model approach has been successfully applied to a broad range of different SCADA signals, proving their ability to detect various types of relevant anomalies.

The research and advances made in the past decade (e.g. by [4] [6] [5] [1] [7]) clearly show the value of also using SCADA signals to monitor wind turbine components conditions. However, so far the focus has been on detecting faults in the gearbox and the generator. Here a CMS system has been proposed that can be successfully applied not only to detect deviation in turbine performance, but also for detection of other malfunctions at smaller components. Using FL, rules can be established which allow implementation of expert knowledge and furthermore highlight the present anomalies in the prediction error. The great advantage is that the generic rules can be established during system setup, which allows detection of present anomalies in the data, right from the beginning (after model training). The overview plot as shown in Figure 3 allows visualization of the fuzzy expert application module results with regards to the most recent condition status of each component/ subsystem.

Sanz-Bobi, Garcia and del Pico [4] show in their research that the implemented rules for gearbox failure diagnosis are also valid to detect similar faults at other turbines. Using different types of faults and NBMs in combination with FL it has been shown here that automatic analysis and diagnosis is possible also for other components and thus generalizing their result. The research further shows that in contrast to the fuzzy expert system proposed by Sanz-Bobi, Garcia and del Pico [4], which is set up to detect only positive deviations from normal behavior, it is important to extend the fuzzy expert system to also consider negative prediction error deviations.

Having set up NBMs for a broad range of different signals (especially of model inputs) gives the expert full flexibility of establishing rules reducing false alarms (via a check if input signals behave normally) and identify the correct root cause. Taking into account earlier research activities of the authors [5], where based on a different data set generator stator anomalies and gearbox bearing damages were identified via FSRC NBMs, the authors are confident that also large component damages can be identified and correctly diagnosed using the proposed method.

Limitations of the method arise from limited access to sensor data. Operators currently do not have full admission to data from all sensors at their wind turbines. The more different sensor data is available, the more diverse and sophisticated models can be built and the more failures can be detected. Furthermore the proposed system setup does not allow detection of very fast progressing faults. The chosen averaging period of one day is a trade of between model accuracy and diagnosis delay time. The system is meant for early fault detection, which is why fast progressing faults are not in focus. A CMS should allow the operator to schedule repair and take informed decisions.

6. Future aspects

With a rising number of rules it becomes more and more important to check, if the rules established earlier still fire, when a new rule is implemented potentially interfering with the old rule. For this purpose, development of a fault reference data base will be required in future in which fault patterns are stored together with the desired diagnosis. After implementation of new rules it should be checked, if the fuzzy expert system still gives the correct results.

With a rising number of SCADA signals, it is also required to further diversify the fuzzy expert system with regards to number of components/ subsystems. In the current development there are nine component/ subsystems defined. With many SCADA signals clustered in a component/ subsystem problems arise, if

two different anomalies or failures are present at the same time in this component. This is because the fuzzy expert application module can only give one diagnosis at a time for each component/ subsystem defined. Future work will therefore require a finer diversification of different component/ subsystems.

In the current development, no difference in diagnosis can be made when the same fault pattern has different root causes. To reduce the potential number of conflicts, future research will therefore focus on the implementation of other measurements, such as high resolution vibrational data or oil sample analysis in the CMS. Furthermore turbines from a different type and brand will be included in the research.

7. Conclusions

In this paper, application examples of the CMS proposed in part one have been given. The examples show the crucial information contained in SCADA signals. Not only with respect to the component conditions, but also to the overall turbine performance. Using FSRC ANFIS NBM in combination with FL this information can be efficiently extracted. The CMS proposed supplies the basis for automated fault detection and diagnosis and aids experts in diagnosis of fault patterns. Taking into account a large number of different SCADA signals, the architecture gives experts a large degree of freedom with respect to knowledge implementation.

Moreover it can be concluded that rules establish with a fault experienced at one turbine can be applied to other turbines of the same type and brand to identify and diagnose similar patterns correctly. This reduces diagnosis effort required drastically and aids operators in achieving their cost reduction targets.

8. References

- [1] Schlechtingen, M., Santos, I.F. and Achiche, S. (2012). "Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description." *Journal of Applied Soft Computing*.
- [2] Randall, Robert B. (2011). "Vibration-based Condition Monitoring." *Wiley*.
- [3] Yang, W. and Jiang, J. (2011). "Wind turbine condition monitoring and reliability analysis by SCADA information." *IEEE, pp.1872-1875*.
- [4] 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 Vol.57, pp.552-568*.
- [5] Schlechtingen, M. and Santos, I.F. (2011). "Comparative analysis of neural network and regression

based condition monitoring approaches for wind turbine fault detection. "*Mechanical Systems and Signal Processing* Vol.25 No.5, pp.1849-1875.

[6] **Zaher, A., McArthur, S.D.J. and Infield, D.G.** (2009). "Online Wind Turbine Fault Detection through Automated SCADA Data Analysis. "*Wind Energy* Vol.12, pp.574-593.

[7] **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.

3.3.1. Extended research and development - Additional information

After publication of [P3] models for power plant 2 were established. While for power plant 1 45 models were developed, 80 were set up for power plant 2 analogously to the described procedure. In Table 3.1, the model performance is summarized together with the signal normal ranges. The normalized percentage performance (P_{norm}) is calculated as:

$$P_{norm} = 100\% \left(1 - \left| \frac{\sigma(E) - r_{max}}{r_{max}} \right| \right) \quad (4)$$

With E being the prediction error (see publication [P3]), σ the SD and r_{max} the maximum value during normal operation.

Table 3.1: Example model performance after training for a random turbine of power plant 2

Model	P	P _{avg}	Normal range	Unit	P _{norm}
Wind speed 1	0,18	0,06	0 to 40	m/s	0,15 %
Wind speed 2	0,17	0,07	0 to 40	m/s	0,18 %
Wind speed cross.	0,14	0,08	- ^{*)}	m/s	-
Cabinet A9 int. temp.	0,59	0,30	0 to 60	°C	0,50 %
Pitch angle A	0,46	0,19	-10 to 90	°	0,21 %
Pitch angle B	0,46	0,19	-10 to 90	°	0,22 %
Pitch angle C	0,46	0,19	-10 to 90	°	0,21 %
Pitch angle A cross.	0,01	0,00	- ^{*)}	°	-
Pitch angle B cross.	0,01	0,01	- ^{*)}	°	-
Pitch angle C cross.	0,01	0,01	- ^{*)}	°	-
Pitch oil pres. A	0,26	0,16	-1 to 270	bar	0,06 %
Pitch oil pres. B	0,35	0,23	-1 to 270	bar	0,08 %
Pitch oil pres. C	0,33	0,23	-1 to 270	bar	0,08 %
Pitch oil pres. A cross.	0,20	0,15	- ^{*)}	bar	-
Pitch oil pres. B cross.	0,23	0,19	- ^{*)}	bar	-
Pitch oil pres. C cross.	0,21	0,15	- ^{*)}	bar	-
Hub temp.	1,06	0,58	0 to 45	°C	1,28 %
Hydraulic oil temp.	0,70	0,13	0 to 60	°C	0,22 %
Hydraulic oil pres. pump	0,44	0,27	-1 to 250	°C	0,11 %
Rotor speed	0,30	0,09	0 to 17	rpm	0,53 %
Main bearing temp.	1,84	1,32	0 to 60	°C	2,19 %
Gearbox oil temp.	1,42	0,29	0 to 65	°C	0,44 %
Gearbox bearing temp. HSS gen.	1,32	0,44	0 to 75	°C	0,59 %
Gearbox bearing temp. HSS rot.	1,49	0,36	0 to 75	°C	0,48 %
Gearbox bearing temp. IMS gen.	1,66	0,35	0 to 65	°C	0,54 %
Gearbox bearing temp. IMS rot.	1,34	0,25	0 to 60	°C	0,41 %
Gearbox oil pres.	0,23	0,13	-1 to 6	bar	2,08 %
Gearbox oil pres. before inline filter	0,65	0,35	-1 to 13	bar	2,71 %
Gearbox oil pres. after inline filter	0,66	0,36	-1 to 13	bar	2,73 %
Gearbox oil pres. diff. inline filter	0,03	0,02	- ^{*)}	bar	-
Gearbox oil pres. diff. offline filter	0,03	0,01	-1 to 4	bar	0,33 %
Brake temp. gear	0,51	0,16	0 to 65	bar	0,24 %
Generator stator temp. R	2,89	1,00	0 to 135	°C	0,74 %

Model	P	P _{avg}	Normal range	Unit	P _{norm}
Generator stator temp. S	2,91	0,97	0 to 135	°C	0,72 %
Generator stator temp. T	2,85	1,00	0 to 135	°C	0,74 %
Generator stator temp. R cross.	0,16	0,12	- ^{*)}	°C	-
Generator stator temp. S cross.	0,14	0,10	- ^{*)}	°C	-
Generator stator temp. T cross.	0,10	0,06	- ^{*)}	°C	-
Generator bearing temp. NDE	2,89	1,57	0 to 95	°C	1,66 %
Generator bearing temp. DE	0,93	0,40	0 to 80	°C	0,49 %
Generator speed	27,49	8,17	0 to 1500	rpm	0,54 %
Cabinet A3 int. temp. left	1,51	0,97	0 to 45	°C	2,15 %
Cabinet A3 int. temp. right	1,05	0,84	10 to 55	°C	1,52 %
Nacelle temp.	1,27	0,66	0 to 50	°C	1,31 %
Humidity tower	2,58	1,72	0 to 100	%	1,72 %
Cabinet A1 int. temp.	1,03	0,23	0 to 45	°C	0,50 %
Cabinet A1 ext. temp.	1,31	0,56	0 to 45	°C	1,23 %
Cabinet A2 int. temp.	1,76	0,96	0 to 45	°C	2,13 %
Converter cool water temp.	1,48	0,72	0 to 65	°C	1,11 %
IO module 1 temp.	0,55	0,44	0 to 55	°C	0,79 %
IO module 2 temp.	0,73	0,54	10 to 60	°C	0,90 %
IO module 1 temp. cross.	0,51	0,45	- ^{*)}	°C	-
IO module 3 temp.	1,59	1,04	10 to 55	°C	1,88 %
Main reactor temp. R	3,50	2,36	0 to 80	°C	2,95 %
Main reactor temp. S	3,25	2,42	0 to 80	°C	3,02 %
Main reactor temp. T	2,71	1,97	0 to 80	°C	2,46 %
Main reactor temp. R cross.	0,55	0,42	- ^{*)}	°C	-
Main reactor temp. S cross.	0,72	0,58	- ^{*)}	°C	-
Main reactor temp. T cross.	0,39	0,26	- ^{*)}	°C	-
Grid filter temp. B1	0,57	0,37	0 to 50	°C	0,73 %
Grid filter temp. B2	0,61	0,39	0 to 50	°C	0,78 %
Grid filter temp. B3	0,61	0,40	0 to 50	°C	0,80 %
Grid filter temp. B1 cross.	0,32	0,15	- ^{*)}	°C	-
Grid filter temp. B2 cross.	0,30	0,06	- ^{*)}	°C	-
Grid filter temp. B3 cross.	0,32	0,14	- ^{*)}	°C	-
Voltage R	0,82	0,55	0 to 450	V	0,12 %
Voltage S	0,87	0,55	0 to 450	V	0,12 %
Voltage T	0,87	0,61	0 to 450	V	0,13 %
Voltage R cross.	0,48	0,36	- ^{*)}	V	-
Voltage S cross.	0,43	0,31	- ^{*)}	V	-
Voltage T cross.	0,38	0,30	- ^{*)}	V	-
Current R	50,00	23,53	0 to 2100	A	1,12 %
Current S	49,86	23,41	0 to 2100	A	1,11 %
Current T	50,09	23,60	0 to 2100	A	1,12 %
Current R cross.	0,86	0,60	- ^{*)}	A	-
Current S cross.	1,28	1,17	- ^{*)}	A	-
Current T cross.	1,03	0,89	- ^{*)}	A	-
Power output	50,55	14,53	-100 to 2350	kW	0,62 %
Transformer oil temp. filtered	4,69	2,11	0 to 110	°C	1,92 %
Transformer room temp.	2,89	1,53	0 to 65	°C	2,35 %

^{*)} for cross predictions the same normal ranges apply as for the original signal

Apart from one exception (Main reactor temp. S) the model performance P_{avg} is satisfactory with P_{norm} less than 3% of the maximum occurring normal value.

Figure 3.1 illustrates the identified input output sets.

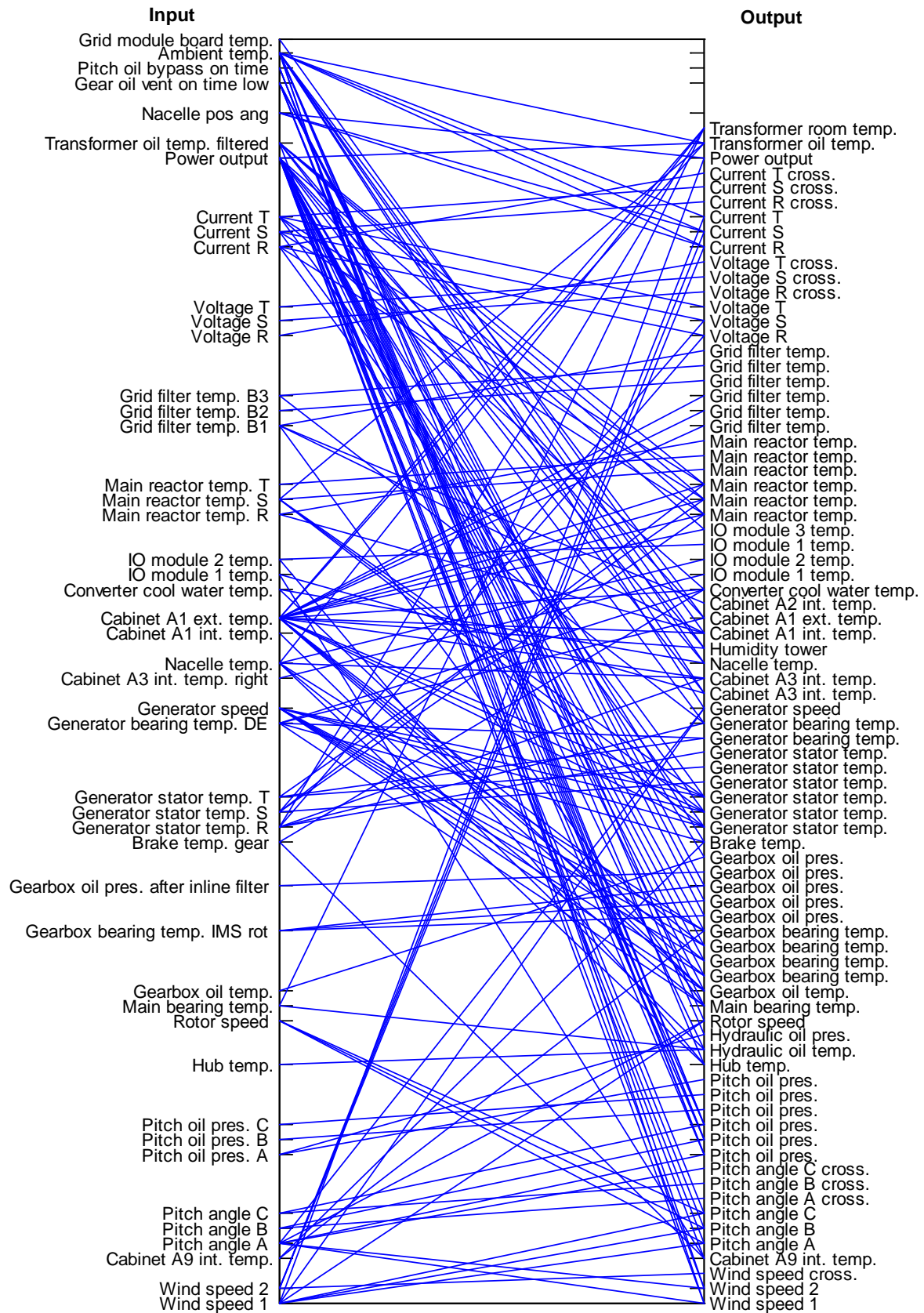


Figure 3.1: Input-output sets of the developed models for power plant 2

Research of this project shows, that identification of good input output sets leading to an acceptable and stable model performance is not trivial. Among other factors, it depends on:

- The available data set (correlated signals available)
- The physical process causing the signal behavior
- Reliability of the used input

The first factor can be aided by linear analysis (use of the cross correlation function) or nonlinear analysis like the described Genetic Algorithm combined with Partial Least Squares (GAPLS). The second factor requires understanding of the physical process as it is described in [P2] for the power output.

In practical applications, understanding of the physical process is very important, but usually not mentioned in literature, due to the focus often being on the prediction error only. With condition monitoring the fault visibility gains importance and should not be neglected.

The third factor is probably the most important. Inputs signals may be well describing the output signal behavior, but due to the operational practice or performed service/maintenance, this input may be considered inappropriate and consequentially discarded. An example for such a signal is the cabinet temperature A9 of power plant 2, which is strongly dependent on the cabinet filter condition. The cabinet is mounted close to the mechanical brake. In case of emergency brakes or other unusual brake events, the filter pollutes and the cabinet temperature increases. Using this signal as input leads situations, in which the output signal cannot be properly monitored. This knowledge can often only be gained over time and in combination with large data set. Hence, after initial model setup a revision may be required after operational experience is gained to optimize the model and CMS performance.

4 Vibration data analysis

In this chapter, state of the art methods to analyze the available wind turbine vibration measurements of power plant 2 are identified and their performance evaluated. The most promising methods are incorporated in the CMS to further evaluate their performance with real world data. Depending on the component to be monitored, different algorithms are required. The components are separated according into the machine elements bearings and gears. Additionally a simple algorithm to monitor the rotor in terms of unbalance (aerodynamic or mass caused) is set up. The algorithm is based on detection of the 1p and 3p, i.e. one time and three times per revolution frequency component in the frequency spectrum. The 1p modulations may indicate general rotor unbalance, whereas changes in the 3p frequency components (at turbines with three blades) may be of interest to monitor due to the impact of this component on the overall fatigue life for the turbine. The derived parameter is the amplitude height. Due to its simplicity and missing validation cases, the algorithm is not further described here.

4.1. Publication [P5] (submitted)

Publication [P5] emphasizes bearing defect detection algorithms. Over many years, it has been established that the benchmark method for bearing diagnosis is envelope analysis, where a signal is bandpass filtered in a high-frequency band in which the fault impulses are amplified by structural resonances [31]. In this research envelope analysis was found to be a strong tool for identification of the faulty components. However, the applied preprocesses (selected filter band and signal separation method) strongly influence the diagnostic information obtained by calculating the signal envelope. Hence the main objection here was to identify suitable preprocess algorithms giving good results for the available wind turbine data and contribute to the validation of the state of the art methods. Furthermore, research was performed to automate diagnosis.

Slow rotating bearings, e.g. as those in the planetary stage of the gearbox or the main bearing have been in the focus of this research, since their diagnosis requires different analysis methods. Research to monitor those components reliable in practical applications is left for the future. Publication [P5] summarizes the findings.

Note: The analysis in publication [P5] is based on the measurements coming from power plant 2, but for reasons of confidentiality, it has been preferred to disguise the real number of turbines from 21 to 17 in the journal paper.

Automated wind turbine gearbox bearing diagnosis algorithm based on vibration data analysis and signal pre-whitening

Meik Schlechtingen*

Department of Technical Operation Wind Offshore, EnBW Erneuerbare Energien GmbH, Germany
Admiralitätstr.4; 20459 Hamburg; Email: m.schlechtingen@enbw.com; Tel.: +49 40533268134

Ilmar Ferreira Santos

Department of Mechanical Engineering, Section of Solid Mechanics, Technical University of Denmark, Denmark

ARTICLE INFO

Article history:

Keywords:

Signal pre-whitening
Bearing diagnosis
Wind turbine gearbox
Data mining
Condition monitoring

ABSTRACT

In this paper an automated wind turbine gearbox bearing diagnosis algorithm is presented. The algorithm is based on most recent research results for separating discrete (gear) from random (bearing) frequency components using Cepstral Editing Procedure (CEP) based signal Pre-Whitening (PW). The proposed automated procedure builds up on the semi-automated procedure described by Sawalhi et al. in 2007. The procedure is updated with regards to the most recent achievements made concerning signal separation and extended with a frequency content identifier and rule based diagnosis to fully automate the diagnosis process. Furthermore, this paper gives a selection of important statements made in literature throughout the last decade to summarize the algorithms used and focus on the most important issues. Each of the involved processing steps is discussed with respect to its effectiveness based on real data.

The proposed procedure is applied to wind turbine data coming from seventeen wind turbines of the 2 MW class, for which vibration data are available containing both healthy and damaged states. Three application examples are given, where the automated procedure successfully diagnosed High Speed Shaft (HSS) bearing damages in the data sets.

1. Introduction

Wind turbine condition monitoring is of increasing importance as the wind turbines are built more and more in remote locations. A special challenge is the placement of wind turbines at sea (offshore). Here the energy yield is usually much higher than on land (onshore), but special (expensive) logistics is required for construction, service & maintenance and decommissioning. Although the higher energy yield is desirable, the cost of energy is still very high in comparison to onshore turbines. From a society perspective a sharp drop in cost of energy must be achieved in the coming years. Next to the usually high construction costs of wind power plants offshore, unexpected downtime due to component failure bears large potential for cost reduction, if faults can be foreseen.

Difficulties with effective monitoring of wind turbines arise from the number installed units, which is increasing rapidly. A 2 GW conventional power station may consist of four 500 MW units, which can be monitored constantly in a cost effective manner [1], but several hundred 3 MW offshore wind turbines would be required to achieve this installed capacity, each in need of monitoring. Hence, a high degree of automation is required to reduce monitoring costs.

Due to their crucial role the components main bearing, gearbox and generator are usually monitored via vibration measurements at offshore turbines. Vibration analysis is by

far the most prevalent method for machine condition monitoring [2]. However, also other information sources can be considered such as Supervisory Control And Data Acquisition (SCADA) data, e.g. temperatures, currents, voltages, power output, etc. in case no vibration measurements are available or to enhance diagnostics [3]. In general, operators should consider all information at hand in order to schedule their service & maintenance and achieve their cost reduction targets. In this research the focus lies on automated wind turbine gearbox bearing diagnosis based on vibration data analysis using signal Pre-Whitening (PW).

A large part of condition monitoring consists of separating the mixed signals at each measurement point into components coming from individual sources [2]. Here the task is to separate the deterministic (e.g. originating from gears) from the random and cyclostationary (e.g. originating from rolling element bearings) components in the signal. Signal PW is a comparably new method and can be used for this separation.

In 2007 Sawalhi and Randall [4] proposed a semi-automated method for diagnosing bearing faults. Using three case studies Sawalhi and Randall show the applicability of their method. It consist of (I) removal of speed fluctuations (II) removal of discrete frequencies (III) removal of the smearing effect of the signal transfer path (IV) determination of the optimum band for filtering &

demodulation and (V) determination of the fault characteristic frequencies.

In the same year Sawalhi et al. [5] studied the enhancement of fault detection and diagnosis in rolling element bearings using Minimum Entropy Deconvolution (MED) combined with Spectral Kurtosis (SK). This corresponds to step III and IV of the semi-automated procedure. The results show that using the MED technique dramatically sharpens the pulses originating from the impacts of the balls with the spall and increases the kurtosis value to a level that reflects the severity of the fault [5]. In this work a good match between the fault size and the Spectral Kurtosis (SK) value was found, when MED is applied to enhance the signals impulsiveness. The effect of MED as pre-process is further studied and a significant increase in diagnostic information found by Barszcz and Sawalhi [6] in 2012. These findings are further discussed in this research for the case of wind turbine gearbox bearings.

Using the same three test cases as described by Sawalhi and Randall [4], the proposed semi-automated method is summarized and further developed in 2011 by Randall and Antoni [7] who presented a tutorial for rolling element bearing diagnostics. In this tutorial different processing techniques for discrete and random separation are summarized (corresponding to step II of the semi-automated procedure), which are: Linear Prediction (LP), Adaptive Noise Cancellation (ANC), Self Adaptive Noise Cancellation (SANC), Discrete/ Random Separation (DRS) and Time Synchronous Averaging (TSA).

In the LP approach an autoregressive linear model built with past values is used to model the current value. Such a model can only represent the deterministic part of the signal. The random part is given by the residual. However, the model order (number of previous values considered) is generally unknown, which requires some search for an optimum order.

The ANC algorithm requires a simultaneously measured coherent signal from a different location, for adjustment of the adaptive filter. Such synchronous data recordings are not always available for operators. In this case the SANC algorithm can be used, in which a delay of the input signal is taken for adoption of the adaptive filters. Antoni and Randall [8] show however, that the same effect could be achieved, but much more efficiently using DRS.

In DRS the efficient Fast Fourier Transformation (FFT) is used. The transfer function between the primary signal and its delayed version (representing only the deterministic part) is calculated in the same way as the H_1 Frequency Response Function (FRF) [9]. This has a value near one at discrete frequencies and near zero at other frequencies [9]. This property is used to filter the signal by fast convolution in the frequency domain. However, it does give a notch filter of fixed width that can have detrimental effects at high and low frequencies [9]. At low frequencies (low harmonics of the bearing frequencies if they exist), the bearing harmonics may still be within the bandwidth of the comb filter, and they may be treated as discrete frequencies [9]. At high frequencies, some random modulation of the

discrete frequency components may be left in the form of a narrow-band noise and not completely removed [9].

In TSA specific discrete frequency components can be removed by averaging together a series of signal segments each corresponding to one period of a synchronising signal [7]. The drawback is that it requires separate operation for every different harmonic in the signal including a separate resampling in each case so as to give a specified integer number of samples per period [7].

Further research was carried out by Sawalhi and Randall with regards to the separation of the discrete and random components (corresponding to step II of the semi-automated procedure). In 2011 a method based on cepstrum editing is proposed by Sawalhi and Randall [10] to remove selected components from a time signal. The cepstrum collects uniformly spaced harmonics and sideband families such as arise from gear mesh harmonics and modulations of these by the gear rotational speeds [10]. For removal of undesired components, the real cepstrum is computed, selected components (quefrequencies) are set to zero and the result inverse transformed back to the initial time domain.

Also in 2011 this procedure was given the name Cepstral Editing Procedure (CEP) by Randall and Sawalhi [9]. Further application examples are given and a performance comparison to the traditionally used TSA and DRS method is made. It is pointed out that this method has a number of advantages compared with alternative methods. Unlike TSA and DRS, it can remove sidebands as well as harmonics and also narrow-band noise peaks rather than just discrete frequency harmonics [9].

Sawalhi and Randall [11] presented an even more drastic editing in the cepstrum, by setting all components to zero except that at zero quefrequency.

In 2012 Niaoqing et al. [12] tested the method using test rig data and inner ring faults with different fault sizes. They used a Stochastic Resonance Model (SRM) to enhance the Envelope spectrum. The method proved to be robust and suitable for the application with more vibration interference. However, the stochastic resonance method requires tuning of the system parameters, which can be difficult in practical applications, involving a large number of units to be monitored. In this work, also the SK value as fault size indicator (first used by Dyer and Stewart [13]) was examined. Results show that the SK after signal processing, contains some fluctuations, but in general is sensible to the fault size. The findings confirm the findings from Sawalhi et al. [5], who conclude that the SK value can be used as fault size indicator after impulsiveness enhancement. In the case of Sawalhi et al. this was achieved by MED in contrast to the SRM algorithm used by Niaoqing et al. [12].

The main original contribution of this paper lies in the further evaluation of signal PW using the CEP in combination with the SK value as a fault size indicator for wind turbine gearbox bearings. Furthermore an automated method for bearing diagnosis is presented using the achievements made by Sawalhi and Randall, as well as own research results. In this respect the semi-automated method

as proposed by Sawalhi and Randall is updated with regards to the most recent developments and further developed with regards to automated envelope spectra diagnosis as it is required for wind turbine gearbox bearing diagnosis.

For this purpose data from seventeen operating wind turbines are available, both in fault free states as well as in states where the fault signatures indicate bearing damage. Using these data the effectiveness of the automated procedure is shown. Additionally the use of the SK value as fault size indicator is discussed using the available data and faults.

In this paper many statements are collected from related papers in this research area in order to summarize the most important conclusions and findings necessary for the understanding of the automated procedure. These statements are consistently cited, allowing finding more information about the specific topics of interest in the original source.

In section 2 of this paper, the available data sets are introduced. In section 3 the automated procedure for bearing fault diagnosis is described in detail with a focus on the individual processing steps involved. Furthermore the efficiency and the effect the involved pre-processes are evaluated. Three application examples are given in section 4. Results and Discussion are made in section 5 and necessary future research discussed in section 6. Finally conclusions are drawn in section 7.

2. Data set description

The data sets available come from seventeen Wind Turbine Generators (WTGs) of the 2 MW class, where vibration measurements are recorded cyclically every second week. The gearbox is equipped with three accelerometers (type: Gram&Juhl GJ1200M8), with one sensor placed at the planetary stage, one in the Inter Mediate Shaft (IMS) region and one in the High Speed Shaft (HSS) region (see Figure 1).

The cyclic measurements available for research are sampled with the settings as summarized in Table 1.

Name	Sampling freq.	Sample lengths	Number of samples
#1	2.59 kHz	12.51 s	32000
#2	15.36 kHz	4.17 s	64100
#3	41.67 kHz	1.54s	64100

Table 1: Summery of recordings

Around 32 measurements per turbine, sensor and sampling frequency are available from the past 1.5 years of operation. Thus in total around 1632 recordings are available. The recording system is set up to only record, when the turbine is operating under full load to allow comparison of measurements at different times (causing the slight differences in number of measurements available for each turbine).

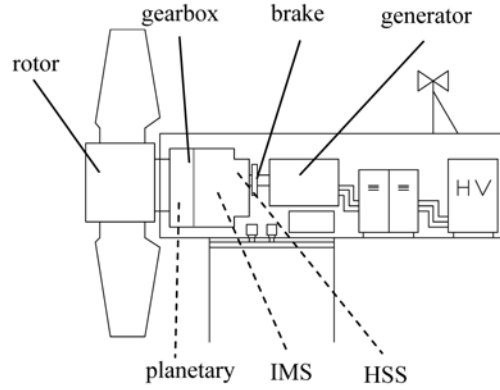


Figure 1: Wind turbine schematic

Due to their longer recording times of measurement #1 and #2, these time series are given by the system as tracked time series. By tracking, speed variations are encountered by resampling the signal synchronous to the speed of the shaft. Here these data are resampled to the nominal speed. For #3 resampling is not necessary as the recording length is short.

3. Automated procedure for bearing fault detection

The semi-automated procedure presented by Sawalhi and Randall [4] in 2007 consists of a five step strategy to arrive at the envelope spectrum to be interpreted for fault signatures. Due to recent advances made by the different researchers, this procedure requires an update with regards to the algorithms used and the signal processing steps involved.

The following update is proposed to fully automate diagnosis:

1. Order tracking – Remove speed fluctuations (optional)
2. Pre-Whitening (PW) based on CEP - Signal decomposition
3. Minimum Entropy Deconvolution (MED) – Remove smearing effect of the signal transfer path (optinal)
4. Spectral Kurtosis (SK) – Determine optimum band for filtering and demodulation (optional)
5. Envelope analysis – Determine fault characteristic frequencies
6. Frequency content identification – Identify frequency components and modulations
7. Diagnosis – Application of rules interpreting the pattern present.

3.1. Order tracking

Order Tracking (OT) means resampling the time series according to the angular position of the shaft. It is useful for variable speed machines in order to remove the rotational speed related frequency fluctuations. For OT generally a

tacho signal or a once per revolution trigger is required that indicates the shafts angular position. In case no tacho signal is available Bonnardot and Randall [14] found it to be possible to extract the instantaneous speed information by phase demodulation of a number of gear mesh frequencies.

When applying digital or analog OT, frequency components which are independent of the shaft speed such as natural frequencies will shift and consequentially their amplitude in the spectra will decrease.

For the automated bearing diagnosis algorithm OT is recommended, but not necessarily mandatory, if the speed variations are limited and the CEP is used [9].

3.2. Pre-whitening

Decomposition of vibration signals into deterministic and nondeterministic parts is recommended for machinery diagnostics as it can be used for the separation of signals emitted from different machinery elements [15]. One of the major sources of masking the relatively weak bearing signals is discrete frequency “noise” from gears, since such signals are usually quite strong, even in the absence of gear faults [7]. Even in machines other than gearboxes, there will usually be strong discrete frequency components that may contaminate frequency bands where the bearing signal is otherwise dominant [7]. It is usually advantageous therefore to remove such frequency noise before proceeding with bearing diagnostic analysis [7].

The fundamental idea of separation is based on the assumption that gear related vibrations are phase locked to the rotational speed, i.e. after order tracking they will appear as discrete frequencies in the spectrum. However, bearings virtually always have some slip for the following reason: the load angle ϕ from the radial plane varies with the position of each rolling element in the bearing as the ratio of local radial to axial load changes, and thus each rolling element has a different effective rolling diameter and is trying to roll at a different speed [2]. The cage ensures that the mean speed of all rolling elements is the same, by causing some random slip [2]. This slip is typically of the order of 1-2%, both as deviation from the calculated value and as a random variation around the mean frequency [2]. The slip, while small, does give a fundamental change in character of the signal, and is the reason why envelope analysis extracts diagnostic information not available from frequency analysis of the raw signal [2].

The separation method here is signal PW for the reasons given in the introduction.

Basic element of the signal PW procedure is the real cepstrum. The cepstrum has a number of versions and definitions, but all can be interpreted as a “spectrum of a log spectrum” [9]. The complex cepstrum reads:

$$C(\tau) = \mathfrak{F}^{-1} \log[\mathfrak{F}[x(t)]] \quad (1)$$

$C(\tau)$ is the complex cepstrum; \mathfrak{F}^{-1} is the inverse Fourier transform; \mathfrak{F} is the Fourier transform; $x(t)$ is the time signal.

In order to arrive at the real cepstrum, only the real part of the complex cepstrum is considered. Information about the history of the cepstrum is summarized by Oppenheim and Schaffer [16]. Additional general information in the context of vibration analysis is given by Randall and Hee [17] and in the context of cepstrum editing by Randall and Sawalhi [9].

The procedure to achieve signal PW is depicted in Figure 2. In order to proceed from the real cepstrum to the edited cepstrum, all quefrequencies are set to zero, apart from the first, which is determining the mean value of the log spectrum and thus giving appropriate scaling of the resulting signals [9].

After editing, the edited log amplitude spectrum is exponentiated and recombined with the original phase spectrum to form the complex log spectrum [9]. When processing the signals with Matlab[®] the required polar to cartesian coordinate transformation can be done with the `pol2cart` function using the phase and the exponentiated edited log amplitude spectrum as input. The output can be interpreted as the real and imaginary part of the complex spectrum.

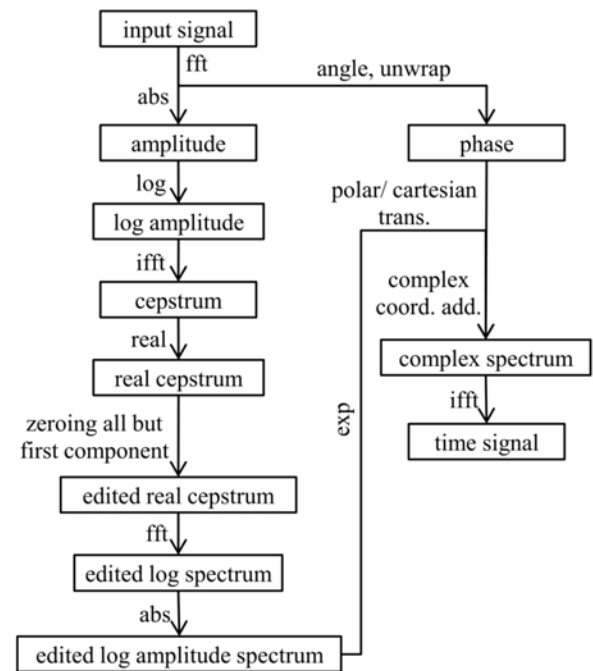


Figure 2: PW based on CEP schematic

Since the procedure is based on the efficient Fast Fourier Transformation (FFT), the whole editing procedure is very fast to calculate and can be fully automated without the need of parameter optimization. The real cepstrum before

editing of #3, measurement 29 at WTG 02 is shown in Figure 3.

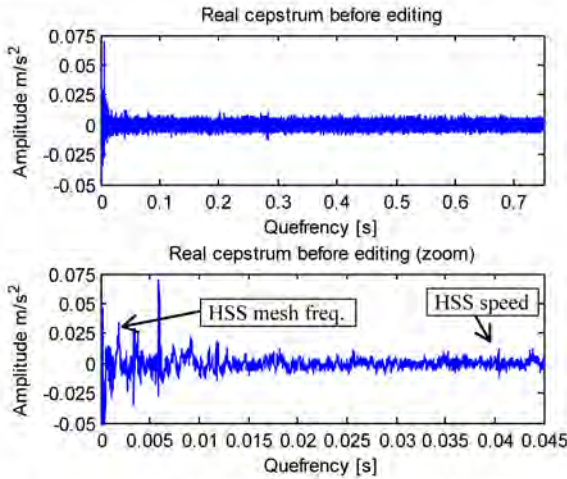


Figure 3: Real cepstrum before editing (#3, HSS, meas. 31, WTG 02)

The HSS mesh frequency as well as the HSS speed can be clearly identified, together with other deterministic components of unknown source. The effect of PW on the corresponding time series as well as the Root Mean Squared (RMS) spectrum is illustrated in Figure 4.

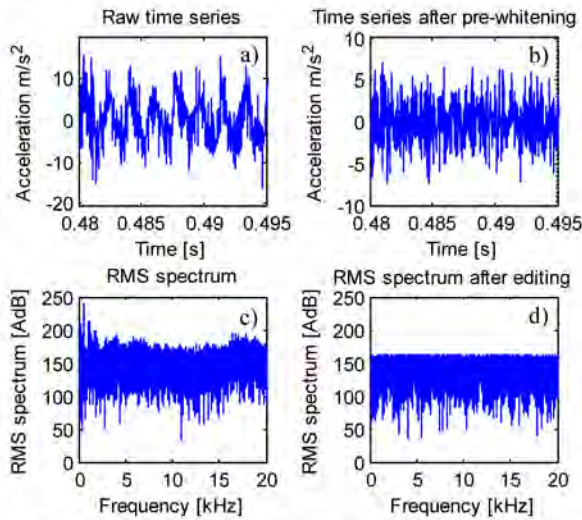


Figure 4: The effect of signal PW (#3, HSS, meas. 31, WTG 02)

Prior PW, the raw time series (a) was contaminated with deterministic components causing cyclical fluctuations. The PW procedure removed the dominance of these components and reduced them to the noise level (b). The RMS spectrum before PW contained dominating frequency components and frequency bands containing more energy (c). After PW

the RMS spectrum has the same level for all frequencies (d).

3.3. Minimum Entropy deconvolution

The Minimum Entropy Deconvolution (MED) method is designed to reduce the spread of impulse response functions, to obtain signals closer to the original impulses that gave rise to them [7]. It was first proposed by Wiggins [18] to sharpen the reflections from different subterranean layers in seismic analysis [7]. The basic idea is to find an inverse filter that counteracts the effect of the transmission path, by assuming that the original excitation was impulsive, and thus having high kurtosis [7].

$$K = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^4}{(N-1)s^4} \quad (2)$$

K is the Kurtosis; s is the standard deviation; N is the number of data points; \bar{Y} is the mean.

The name MED derives from the fact that increasing entropy corresponds to increasing disorder, whereas impulsive signals are very structured, requiring all significant frequency components to have zero phase simultaneously at the time of each impulse [7]. Thus, minimizing the entropy maximizes the structure of the signal, and this corresponds to maximizing the kurtosis of the inverse filter output (corresponding to the original input to the system) [7]. The method might just as well be called maximum kurtosis deconvolution because the criterion used to optimize the coefficients of the inverse filter is maximization of the kurtosis (impulsiveness) of the inverse filter output [7]. The MED method was applied to bearing diagnostics by Sawalhi, Randall and H. Endo [5].

Using the MED method, the kurtosis can be increased, if impulses are present. Note that in case the discrete frequencies are not removed prior the MED, they will also be amplified.

The Matlab® code used for performing the MED was developed by G. McDonald [19] based on the algorithm proposed by Wiggins [18].

A critical choice using MED is the filter size. If the filter size is chosen too large, start-up and ending transients may occur, dominating the minimization method and leading to a loss of diagnostic information. The effect of the filter size on the time series is illustrated in Figure 5, for a time series with start-up and ending transients for large filter size (a) and medium filter size (b). For a low filter size (c) reasonable impulse amplification is achieved without transients. (d) shows the Time Series (TS) after PW without MED.

This is in contrast to the findings of Barszcz and Sawalhi [6] who found very large filters (filter size > 2000) to perform better in case of their surveyed wind turbine gearbox bearing inner ring damage.

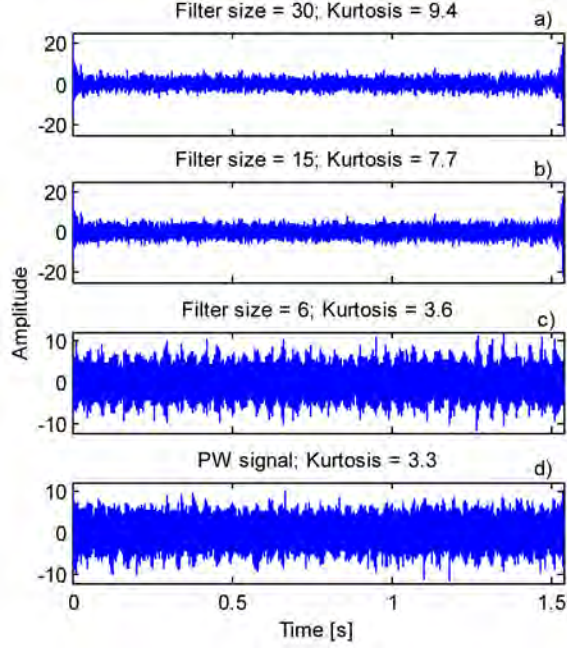


Figure 5: Time series after PW & MED a) filter size = 30; b) filter size = 15; c) filter size = 6; d) signal after PW without MED (#3, HSS, meas. 19, WTG 05)

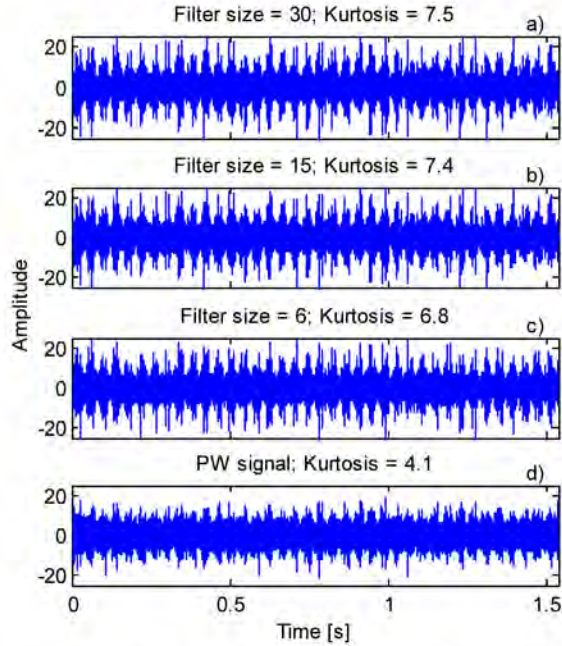


Figure 6: Time series after PW & MED a) filter size = 30; b) filter size = 15; c) filter size = 6; d) signal after PW without MED (#3, HSS, meas. 32, WTG 05)

Their finding is based on the kurtosis values achieved, apparently using a TS where MED did not lead to

start-up and end transients. For blind analysis the filter size used should give reasonable results for all TS processed.

Figure 6 shows the effect of the filter size for a time series without transients. It is seen that choosing a filter size of six already gives a reasonable impulse amplification and increase in kurtosis, which is why this filter size is used throughout the research as a stable compromise.

Figure 7 illustrates the kurtosis as a function of the filter size for the TS with and without transients.

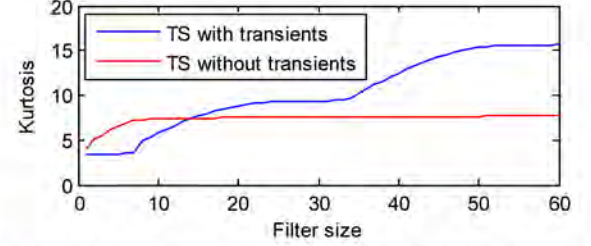


Figure 7: Kurtosis value as function of the filter size; blue (#3, HSS, meas. 19, WTG 05) and red (#3, HSS, meas. 32, WTG 05)

The kurtosis for the TS without transients occurring converges with increasing filter size, whereas the kurtosis values with transients stepwise increases. The transients are believed to be caused by the first and the last value in the time series. In case of their occurrence, these values may be modified or deleted to avoid the transients and obtain proper results. However, further root cause investigation is required, which is outside the scope of this paper.

For the proposed automated procedure, applying MED is not necessarily required, since after successful PW the impulsiveness of the TS is already enhanced, building a good basis for a consecutive envelope analysis. Figure 8 shows the effect of MED of the PW signal in the time domain.

In the raw time series (a) the individual impulses are visible, but they are embedded in the cyclic variations of the deterministic components. Signal PW removes the deterministic components, but leaves a high noise level embedding the impulses (b). Applying MED enhances the impulses and gives the clearest signal (c). The kurtosis increased from 4.5 after PW to 8.0 after PW & MED. After introduction of the envelope spectrum the influence of MED in the frequency domain will be discussed.

3.4. Spectral kurtosis

Spectral kurtosis (SK) provides a means of determining which frequency bands contain a signal of maximum impulsivity [7]. The definition is given by equation 3.

$$SK = \frac{\langle H^4(t, f) \rangle}{\langle H^2(t, f) \rangle^2} - 2 \quad (3)$$

SK is the spectral kurtosis, $H(t, f)$ is the Amplitude envelope function.

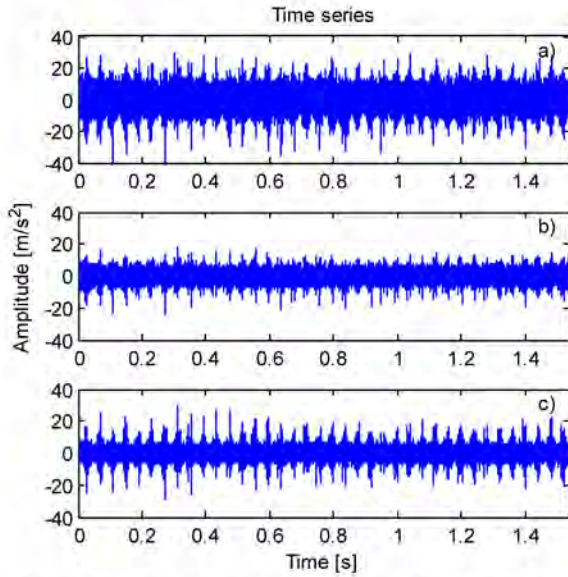


Figure 8: The time series a) no processing, b) PW, c) PW & MED (#3, HSS, meas. 31, WTG 02)

It was first used in the 1980s for detecting impulsive events in sonar signals [20]. It was based on the Short Time Fourier Transform (STFT) and gave a measure of the impulsiveness of a signal as a function of frequency [7]. The spectral kurtosis extends the concept of the kurtosis, which is a global value, to that of a function of frequency that indicates how the impulsiveness of a signal, if any, is distributed in the frequency domain [7]. The principle is analogous in all respects to the Power Spectral Density (PSD) which decomposes the power of a signal vs frequency, except that fourth-order statistics are used instead of second order [7]. This makes the spectral kurtosis a powerful tool for detecting the presence of transients in a signal, even when they are buried in strong additive noise, by indicating in which frequency bands these take place [7]. However, the optimal window length used for the STFT is generally unknown. To obtain a maximum value of kurtosis, the window must be shorter than the spacing between the pulses but longer than the individual pulses [2]. The center frequency and bandwidth (window length and center) is thus varied continuously from low to high values, while recording the SK of each window and thus arriving at a full kurtogram, which can be used to identify the optimum filter settings (center frequency and bandwidth). This procedure is used for bearing diagnosis by Antoni and Randall [21].

The signal is filtered according to the optimum filter settings determined, arriving at a signal that has maximum kurtosis for envelope analysis.

However, computation of the kurtogram for all possible combinations of center frequencies and bandwidths is

obviously costly and not convenient for practical purposes [7]. Therefore the fast kurtogram was proposed by Antoni [22] in 2007. A „1/3-binary tree“ was used, where each halved-band is further split into three other bands, thus producing a frequency resolution in the sequence $1/2, 1/3, 1/4, 1/6, 1/8, 1/12, \dots, 2^{-k-1}$ with corresponding “scale levels” $k=0, 1, 1.6, 2, 2.6, \dots$ to compute the SK in this bands.

With this limited amount of combinations the computational time could be drastically decreased, but virtually the same result can be achieved than with using the full kurtogram. The fast kurtogram algorithm for detection of transient faults proposed by Antoni [22], was made available for public in terms of a Matlab[®] function [23], which is used here.

Figure 9 shows the kurtogram for the example wind turbine and time series after PW and PW & MED.

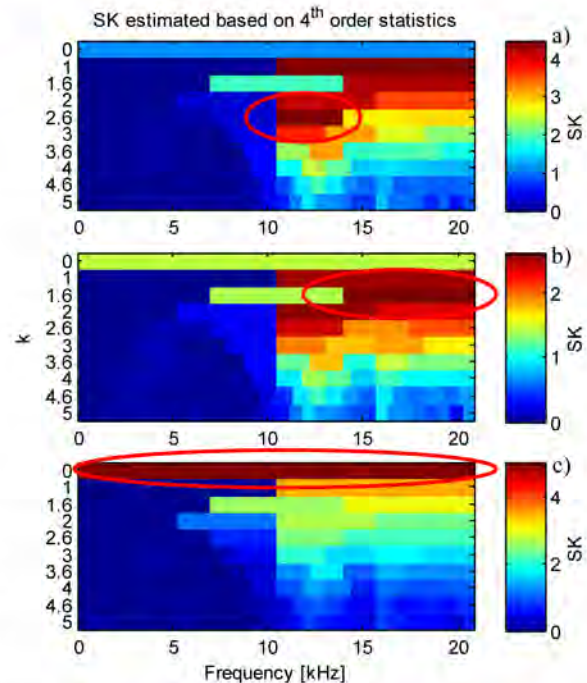


Figure 9: Kurtogram based on 4th order statistics a) raw time series, b) PW, c) PW & MED (#3, HSS, meas. 31, WTG 02)

The maximum SK is achieved at different frequency bands, depending on the processing involved. The kurtogram for the raw time series (a) has its maximum at $k = 2.6$, corresponding to a center frequency of 12.1 kHz and a bandwidth of 3.7 kHz. After PW (b) maximum SK is in the band with a center frequency of 17.6 kHz and a bandwidth of 6.9 kHz ($k=1.6$). After PW & MED (c), maximum kurtosis is in the full frequency band.

Sawalhi and Randall [11] conclude in their research, based on a fault in a planetary bearing at a helicopter gearbox, that even without filtration in the optimum frequency band, complete whitening of the spectrum means that faults have a good chance of manifesting themselves in

an envelope analysis covering the full frequency range, because the most impulsive bands will tend to dominate peaks in the time signal. It is further concluded, that if desired, spectral kurtosis can in any case be applied after cepstral PW, to give an optimal result. In general, a shift of the maximum kurtosis to a lower scale level (broader band) is desirable, as this allows a broader frequency spectrum to be analyzed in the envelope spectrum. A small bandwidth can cause problems in spectra interpretation as the higher harmonics may not be visible, due to the limited maximum frequency.

3.5. Envelope analysis

Envelope analysis has become one of the prominent vibration signal processing techniques for detection and diagnosis of rolling element bearing incipient failure [24].

The envelope can be calculated by performing a Hilbert transform of the squared signal. The signal is band pass filtered according to the band width and center frequency identified through the SK and the kurtogram beforehand. The envelope is then further transformed via a FFT to the frequency domain to arrive at the squared envelope spectrum, which can be used to identify the damage frequencies. Next, the effect of the different pre-processing algorithms (PW, MED and SK) on the squared envelope spectrum is discussed.

3.5.1. The effect of PW

The focus lies on the improvement of the envelope spectra when the time series is pre-whitened. A comparison is made between the raw TS, the PW TS and the LP residual.

The LP (AutoRegressive (AR)) model used here for comparison takes the form

$$\hat{x}(n) = \sum_{k=1}^p a(k)x(k-n) \quad (3) [2]$$

$\hat{x}(n)$ is the predicted current value, p is the number of previous values (model order), $a(k)$ is the weighting coefficient and $x(k-n)$ is the previous value.

The weighting coefficients are obtained using the Yule-Walker equations model and Akaike's entropy-based Information Criterion (AIC) [25] to estimate the model order p .

Although the model order can be estimated using AIC it must be determined for each time series individually as the rotational speed can vary slightly, influencing the discrete frequency components and hence requires different model parameters. The model order can be quite high for time series with a high sampling frequency. The AIC lead to an optimum model order of 598 for the example TS used. Here PW gives some advantage as no parameter must be optimized, making PW a computationally very efficient method.

Figure 10 emphasizes that the PW procedure gives virtually the same result as LP (b and c for all WTGs are very similar). The fundamental IMS frequency (6.4 Hz), the IMS meshing frequency (115.1 Hz), the HSS meshing frequency (550 Hz) component as well as the modulations of the meshing frequency have been successfully removed by both methods at all WTGs (b,c), leaving the bearing damage frequency (297 Hz), its harmonics and shaft speed modulations (± 24.8 Hz) together with the fundamental shaft speed (24.8 Hz). The PW based on CEP thus gives good results in a very efficient manner.

Without separation, the envelope spectra can be dominated by the deterministic frequencies (compare WTG 04 (a) and WTG 10 (a)), making it impossible to diagnose the bearing defect reliably.

3.5.2. The effect of MED

It was already shown that the MED method can increase the kurtosis value in the time domain and enhance the impulsiveness of the signal. In this section it will now be investigated, whether or not this gives an advantage in respect of fault visibility in the envelope spectra using the available data sets. The fault visibility is increased, when typical fault patterns (e.g. damage frequencies) are enhanced, allowing a certain diagnosis. Figure 11 shows the squared envelope spectra of the TS with PW and PW & MED. The strongest improvement archived is at WTG 04. Here the shaft speed modulations around the damage frequency ($297 \text{ Hz} \pm 24.8 \text{ Hz}$) are stronger pronounced with MED. The same is true for WTG 02. At the other turbines the improvement is marginal as (a) and (b) are very similar. However, the amplitudes of PW & MED are higher, which becomes evident by the lower noise level.

It can be concluded, that MED can improve the results and is thus recommended, but diagnosis should also be possible without MED.

3.5.3. The effect of band pass filtering using SK

The time series of the four example WTGs of Figure 10 and Figure 11, have the highest SK value in the full band after applying MED. Hence SK filtering does not give any improvement. Figure 12 compares the envelope spectra with and without optimal band pass filtering prior MED, but after PW.

Both spectra clearly show the present inner race damage signature, which consists of the Ball Pass Frequency Inner race (BPFI) (297 Hz) and present modulations at the shaft speed (24.8 Hz). Differences exist in the amplitudes and the higher modulation of the shaft speeds around the second harmonic of the BPFI. The results appear to be slightly better with band pass filtering. The conclusion drawn by Sawalhi and Randall [11] can thus be confirmed with the data at hand at least prior MED.

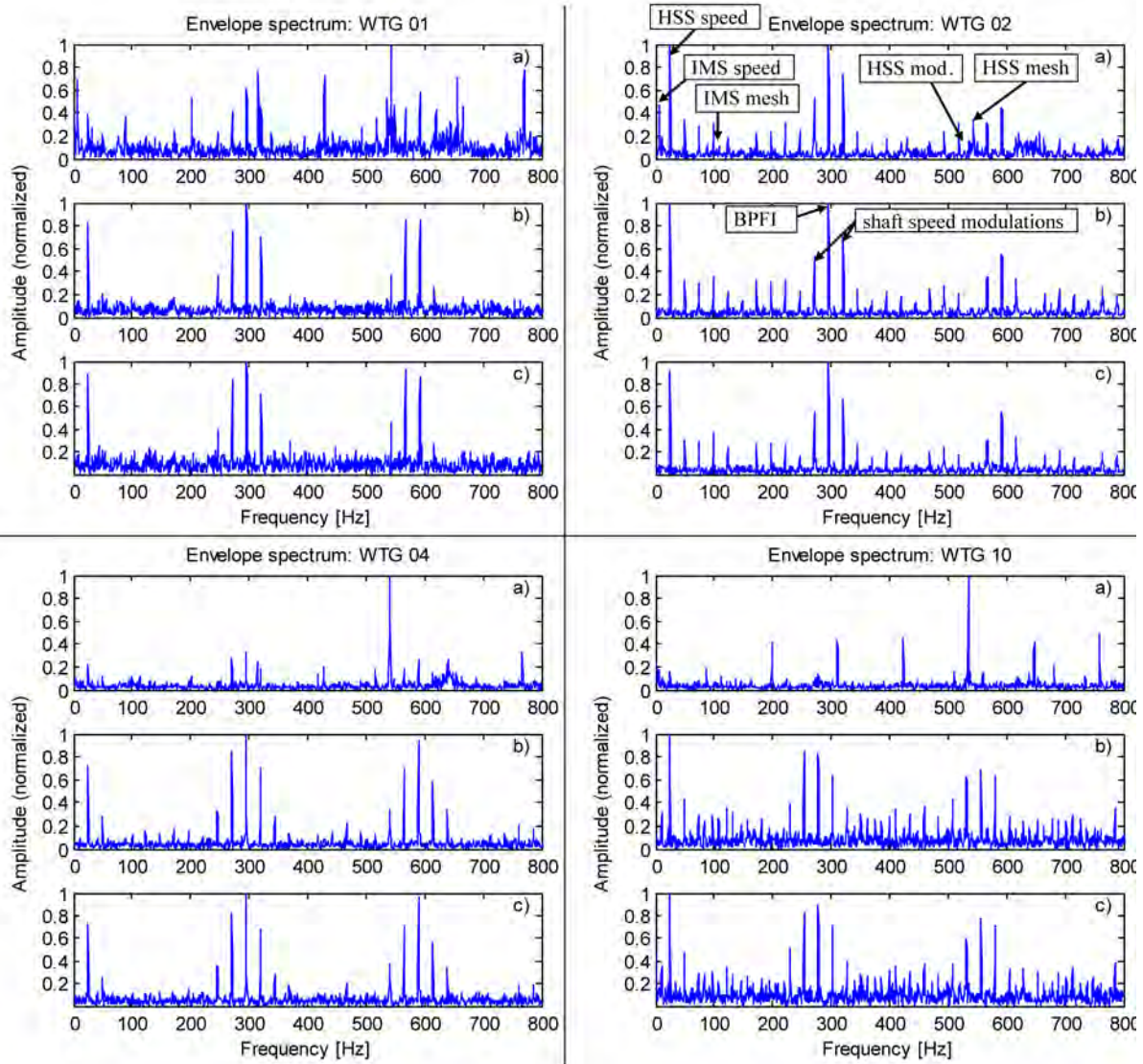


Figure 10: The effect of signal PW; a) no processing, b) linear prediction, c) PW (#3, HSS, meas. 31, WTG 01, 02, 04, 10)

The SK algorithm is fast to compute, (the kurtogram shown in Figure 9 was calculated in 0.5s using a regular 2.6 GHz processor). It is therefore recommended to apply optimal band filtering based on the SK, bearing in mind that in most cases, no improvement to the envelope spectra is made. However, a few cases existed in the data, where the maximum SK was not in full band. In these cases, it is expected that SK filtering improves the envelope spectrum in a fault situation.

3.6. Frequency content identification

For bearing diagnosis, the bearing damage frequencies as well as the individual shaft speeds are required. This information is usually available, either through theoretical calculations based on the geometry and design or in terms of damage frequencies coming from the bearing manufacturer (potentially originating from experimental studies). For further analysis it is assumed that these frequencies are available. In case those frequencies are not at hand, estimation methods can be applied.

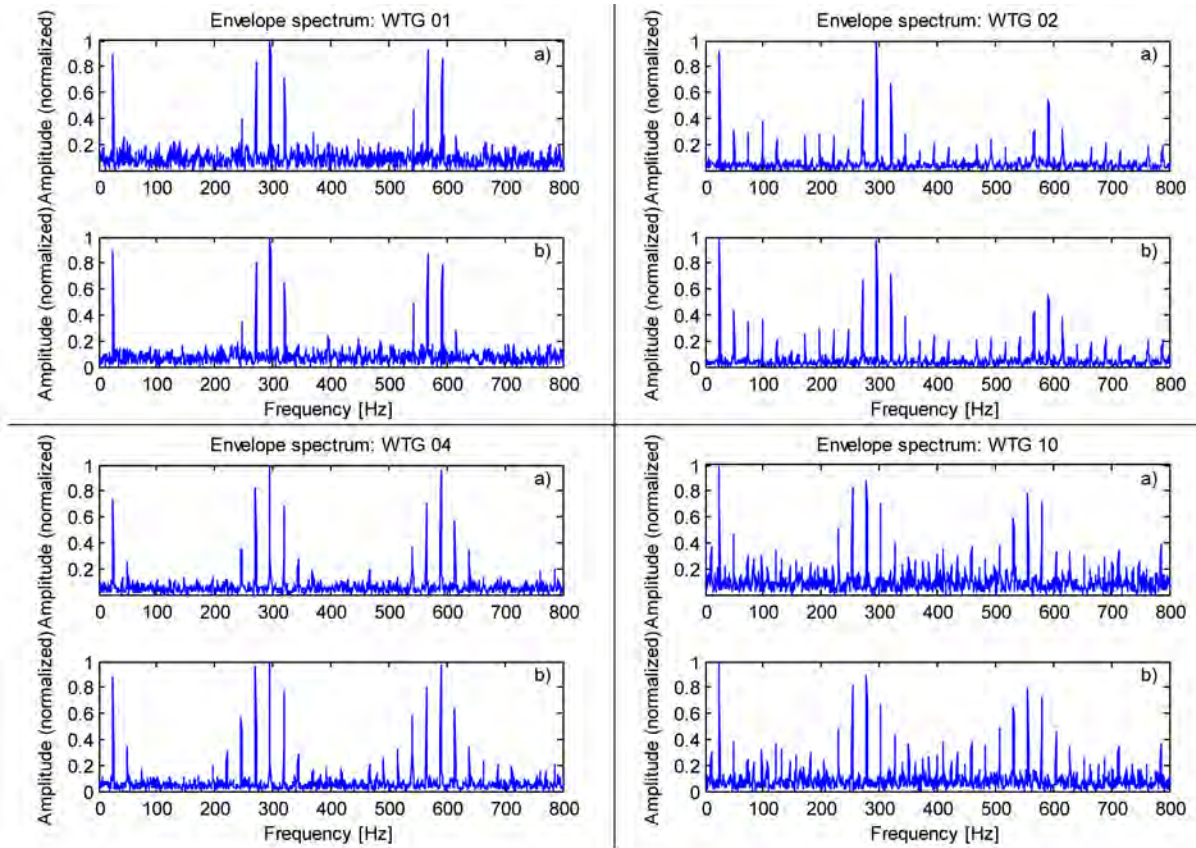


Figure 11: The effect of MED; a) PW without MED, b) PW & MED (#3, HSS, meas. 31, WTG 01, 02, 04, 10)

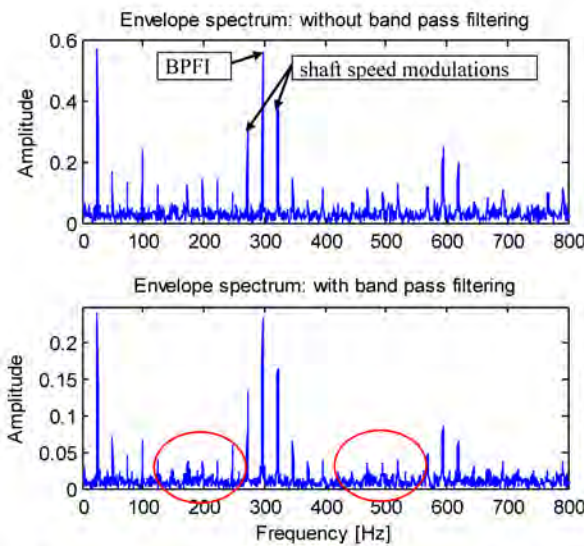


Figure 12: Envelope spectra with and without band pass filtration (PW without MED) (#3, HSS, meas. 31, WTG 02)

As pointed out by Randall and Antoni [7], the deviations from the theoretical values of the damage frequencies may be in the range of 1-2%. Any algorithms developed must therefore be capable of dealing with variations of this order, without affecting the diagnosis.

The Frequency Content Identifier (FCI) developed here deals with this by applying a tolerance band around the known damage frequency with a variation tolerance of 2%. Next to the pure damage frequencies, the presence of modulation sidebands is of interest, too. The developed concepts of damage and modulation frequency identification are visualized in Figure 13.

Before applying the FCI the envelope spectrum must first be checked for present peaks.

A data point is a local peak if (a) it is large and locally a maximum value within a window; the value need not necessarily be large nor a global maximum; and (b) it is isolated i.e. not too many points in the window have similar values [26].

Using this definition two types of peak detectors are set up, which are visualized in Figure 14. Both detectors consist of a moving window shifting along the whole envelope spectrum.

Fundamental frequency	Modulation frequency				
	RPM	BPFI	BPFO	FTF	BSF
RPM	-	0, 1 ^{*)}	0, 1 ^{*)}	0, 1 ^{*)}	0, 1 ^{*)}
BPFI	0, 1 ^{*)}	-	0, 1 ^{*)}	0, 1 ^{*)}	0, 1 ^{*)}
BPFO	0, 1 ^{*)}	0, 1 ^{*)}	-	0, 1 ^{*)}	0, 1 ^{*)}
FTF	0, 1 ^{*)}	0, 1 ^{*)}	0, 1 ^{*)}	-	0, 1 ^{*)}
BSF	0, 1 ^{*)}	0, 1 ^{*)}	0, 1 ^{*)}	0, 1 ^{*)}	-

^{*)} 0: fundamental frequency; 1: first harmonic

Table 3: Wanted modulation contents

In these tables RPM is the shaft speed; BPFI is the ball pass frequency inner race; BPFO is the ball pass frequency

outer race; FTF is the fundamental train frequency and BSF is the ball spin frequency.

The result of the FCI for each component is stored in a matrix (see Table 4). If the frequency content is present (Pres.) the matrix entry will contain a one or zero otherwise. This matrix can afterwards be used to set up rules for automatic diagnosis. Note that the frequency contents labeled Mod_ are identified using the modulation frequency identifier and the other frequency contents are identified using the damage frequency identifier (compare Figure 13).

No.	Freq. Cont.	Pres.
1	RPM h0	0/1 ^{*)}
2	RPM h1	0/1 ^{*)}
3	RPM h2	0/1 ^{*)}
4	RPM h3	0/1 ^{*)}
5	RPM h4	0/1 ^{*)}
6	BPFI h0	0/1 ^{*)}
7	BPFI h1	0/1 ^{*)}
8	BPFI h2	0/1 ^{*)}
9	BPFI h3	0/1 ^{*)}
10	BPFI h4	0/1 ^{*)}
11	BPFO h0	0/1 ^{*)}
12	BPFO h1	0/1 ^{*)}
13	BPFO h2	0/1 ^{*)}
14	BPFO h3	0/1 ^{*)}
15	BPFO h4	0/1 ^{*)}
16	FTF h0	0/1 ^{*)}
17	FTF h1	0/1 ^{*)}
18	FTF h2	0/1 ^{*)}
19	FTF h3	0/1 ^{*)}
20	FTF h4	0/1 ^{*)}
21	BSF h0	0/1 ^{*)}
22	BSF h1	0/1 ^{*)}

No.	Freq. Cont.	Pres.
23	BSF h2	0/1 ^{*)}
24	BSF h3	0/1 ^{*)}
25	BSF h4	0/1 ^{*)}
26	Mod RPM h0 BPFI h0	0/1 ^{*)}
27	Mod RPM h0 BPFI h1	0/1 ^{*)}
28	Mod RPM h0 BPFO h0	0/1 ^{*)}
29	Mod RPM h0 BPFO h1	0/1 ^{*)}
30	Mod RPM h0 FTF h0	0/1 ^{*)}
31	Mod RPM h0 FTF h1	0/1 ^{*)}
32	Mod RPM h0 BSF h0	0/1 ^{*)}
33	Mod RPM h0 BSF h1	0/1 ^{*)}
34	Mod BPFI h0 RPM h0	0/1 ^{*)}
35	Mod BPFI h0 RPM h1	0/1 ^{*)}
36	Mod BPFI h0 BPFO h0	0/1 ^{*)}
37	Mod BPFI h0 BPFO h1	0/1 ^{*)}
38	Mod BPFI h0 FTF h0	0/1 ^{*)}
39	Mod BPFI h0 FTF h1	0/1 ^{*)}
40	Mod BPFI h0 BSF h0	0/1 ^{*)}
41	Mod BPFI h0 BSF h1	0/1 ^{*)}
42	Mod BPFO h0 RPM h0	0/1 ^{*)}
43	Mod BPFO h0 RPM h1	0/1 ^{*)}
44	Mod BPFO h0 BPFI h0	0/1 ^{*)}

No.	Freq. Cont.	Pres.
45	Mod BPFO h0 BPFI h1	0/1 ^{*)}
46	Mod BPFO h0 FTF h0	0/1 ^{*)}
47	Mod BPFO h0 FTF h1	0/1 ^{*)}
48	Mod BPFO h0 BSF h0	0/1 ^{*)}
49	Mod BPFO h0 BSF h1	0/1 ^{*)}
50	Mod FTF h0 RPM h0	0/1 ^{*)}
51	Mod FTF h0 RPM h1	0/1 ^{*)}
52	Mod FTF h0 BPFI h0	0/1 ^{*)}
53	Mod FTF h0 BPFI h1	0/1 ^{*)}
54	Mod FTF h0 BPFO h0	0/1 ^{*)}
55	Mod FTF h0 BPFO h1	0/1 ^{*)}
56	Mod FTF h0 BSF h0	0/1 ^{*)}
57	Mod FTF h0 BSF h1	0/1 ^{*)}
58	Mod BSF h0 RPM h0	0/1 ^{*)}
59	Mod BSF h0 RPM h1	0/1 ^{*)}
60	Mod BSF h0 BPFI h0	0/1 ^{*)}
61	Mod BSF h0 BPFI h1	0/1 ^{*)}
62	Mod BSF h0 BPFO h0	0/1 ^{*)}
63	Mod BSF h0 BPFO h1	0/1 ^{*)}
64	Mod BSF h0 FTF h0	0/1 ^{*)}
65	Mod BSF h0 FTF h1	0/1 ^{*)}

^{*)} 0: Frequency component not present; 1: frequency component present

Table 4: Content matrix

This setup allows for future addition of further contents in envelope spectra that may be checked for. This could be for instance higher harmonics of interest. In this case the content matrix can be extended.

For the application example used here the frequency content around the BPFI frequency together with the tolerance limits is shown in Figure 17.

In line with Figure 17, the entries: BPFI_h0 and Mod_RPM_h0 BPFI_h0 will contain a "1", since the fundamental BPFI frequency as well as shaft speed modulations around this frequency are present.

3.7. Diagnosis

The FCI outputs the content matrix containing the indicators of whether a particular frequency or frequency pattern (modulation) is present in the envelope spectrum.

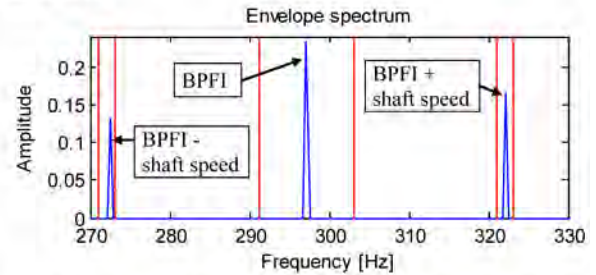


Figure 17: Envelope spectrum after peak filtering and the tolerance limits (#3, HSS, meas. 29, WTG 02)

Based on this matrix rules can be set up which allow the expert to implement his knowledge in fault diagnosis. Rules based on common knowledge about typical bearing damages are for instance:

- When BPFI frequency present & shaft speed modulations around BPFI present *then* Inner race damage
- When FTF frequency present *then* Rolling element damage
- When BPFO frequency present *then* Outer race damage
- When BPFO frequency present & shaft speed modulations around BPFO present *then* Outer race damage + unbalance or misalignment

Instead of formulating rules in a classical way, Fuzzy Logic (FL) is applied, since FL rules can be expressed in a very compact manner. Note that the advantages of fuzzification are not used here. Each line in the content matrix has its own input variable to the FL system. Because the content matrix does only contain two values - either "0" (frequency content not present) or "1" (frequency content present) - the shape of the Membership Function (MF) does not influence the result. The applicable MF has always full membership, due to the functions being centered around 0 and 1. The output is the diagnosis based on the rule evaluation. The initialized MFs are visualized in Figure 18 and Figure 19.

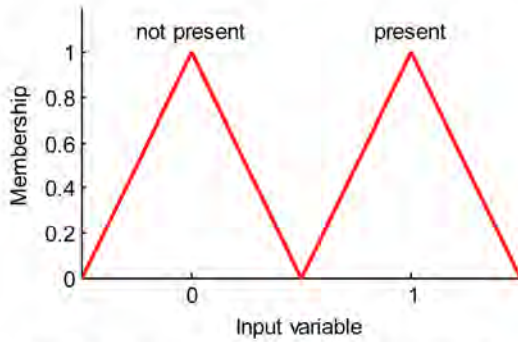


Figure 18: Membership functions initialized for each input variable

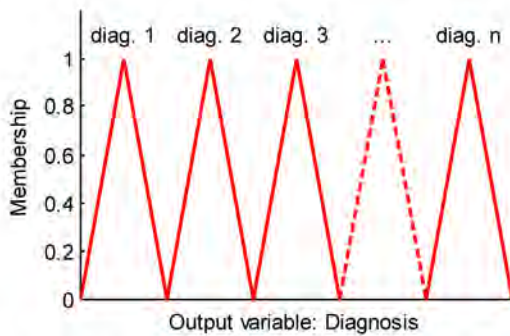


Figure 19: Membership functions initialized for each diagnosis

The rules are specified in a table whose structure is shown in Figure 20.

	RPM _{h0}	RPM _{h1}	RPM _{h2}	...	Mod_BSF _{h0}	Mod_FTF _{h1}	Diagnosis	Diagnosis code	weight	and/or
Rule1	-	-	-	...	-	-	Diagnosis 1	1	1	1
Rule2	1	-	0	...	-	-	Diagnosis 2	2	0.9	1
Rule3	-	-	-	...	-	-	Diagnosis 3	3	1	1
Rule4	1	1	-	...	-	-	Diagnosis 4	4	0.9	1
...

-:- input not considered in rule
 0: frequency content not present
 1: frequency content present
 Continuous number
 Arbitrary text
 Rule weight 0...1
 0: logical or
 1: logical and

Figure 20: Rule specification table structure

4. Application examples

From the seventeen wind turbines in the power plant, six have inner race damage at the high speed shaft bearing. Unfortunately this bearing cannot be inspected by endoscopic methods due its closed design. Hence tracking of the fault size was not possible. However, after bearing replacement factory investigations confirmed straight axial cracks.

In this section three application examples for the automated fault detection algorithm are given, using the proposed procedure.

4.1. Example 1

The available discontinuous measurements are evaluated for faults after PW & MED and SK processing. The envelope spectra after filtering are analyzed via the FCI and the rules specified. Figure 21 shows a waterfall plot of the envelope spectra together with the spectral kurtosis values of the full frequency band of the most recent ten measurements.

Rising amplitudes in the envelope spectra are visible. This corresponds nicely with the increase in SK of the full frequency band. Since March 2012 a rise in SK value can be observed. Although it is not proven here, the findings of Sawalhi et al. [5], Niaoqing et al. [12] and Sawalhi and Randall [27] make it very likely that the increase in SK can be linked to the crack growths at the inner ring.

Note that here the SK of the full band is plotted, instead of the SK of the band with maximum impulsiveness. This choice was made to avoid fluctuations in the SK trend, simply due to shifts in the frequency band with maximum SK. In this sense the same reference band is used.

Moreover it was realized that the MED algorithm makes the SK values in a healthy state more fluctuating. Hence the SK trend in Figure 21 is based on SK values prior application of MED. It is visible that the SK trend

corresponds with the envelope spectra. Repair took place in November 2012.

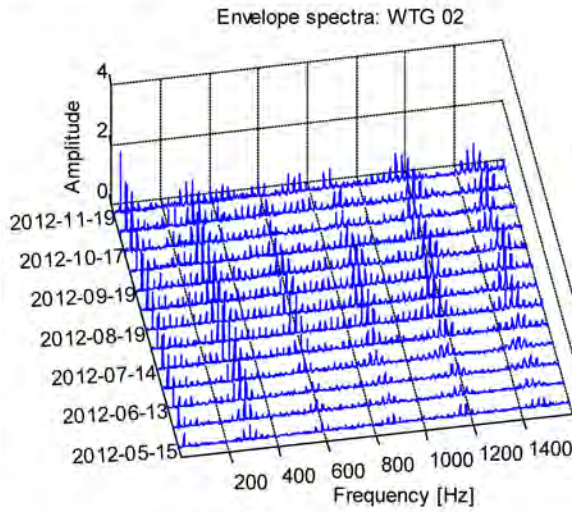


Figure 21: Envelope spectra and spectral kurtosis (WTG 02)

Figure 22 illustrates the SK trend after applying MED as a comparison. This figure presents a novelty as for the first time the SK evolution is shown over time covering both, the long healthy and the damaged period.

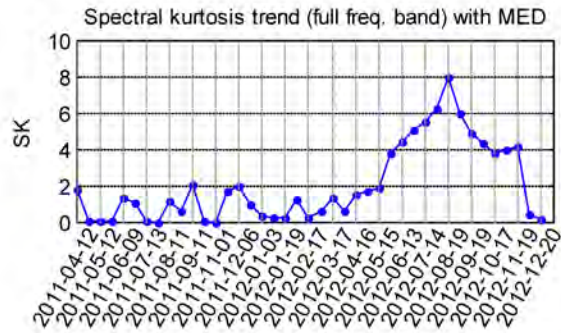


Figure 22: Spectral kurtosis trend after MED (WTG 02)

The fluctuations in the healthy state are characterized by a large variation. Once the crack develops, the SK values stabilize and indicate a trend, most likely linked to the crack size. Beyond 2012-08-19 the SK values decrease. A similar

behavior was observed by Sawalhi and Randall [28]. They suspected the decrease to be due to the fact that the entry and exit edges of the spall would tend to wear [2]. The actual fault here aids this presumption, as the characteristic fault frequency amplitude decreases in the later time series, but rich sideband modulations occur.

Differences in the SK trend with and without MED exist in the healthy period as well as in the overall amplitude level. The fluctuating pattern of the SK trend prior fault development makes the SK trend after MED a more difficult parameter to monitor, which is why the SK trend without MED is preferred.

A further difference can be observed at 2011-12-06, where the SK trend without MED shows a noticeable increase, whereas it is not as obvious in the SK trend after MED. The increase at 2011-12-06 has its reason, as in this period the envelope spectra gives indication of fault (see Figure 23), which disappears again after 2012-01-03. This phenomenon is potentially caused by flattening of the developing crack due to overrollment and a continuous crack growth beyond 2012-03-17.

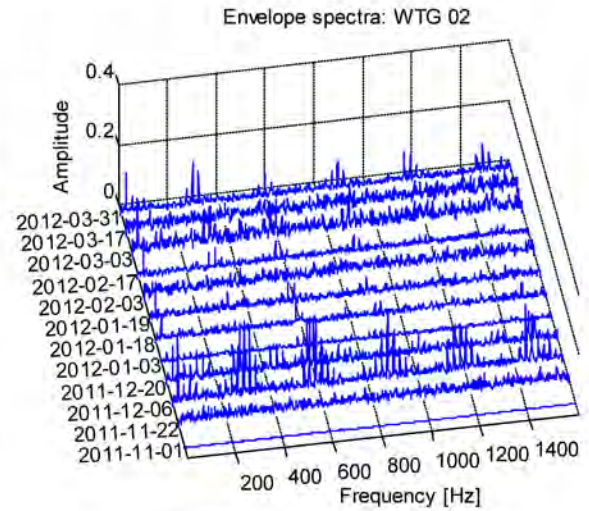


Figure 23: Envelope spectra and spectral kurtosis (WTG 02)

Figure 24 shows the envelope spectrum for the most recent measurement together with the most important frequency contents.

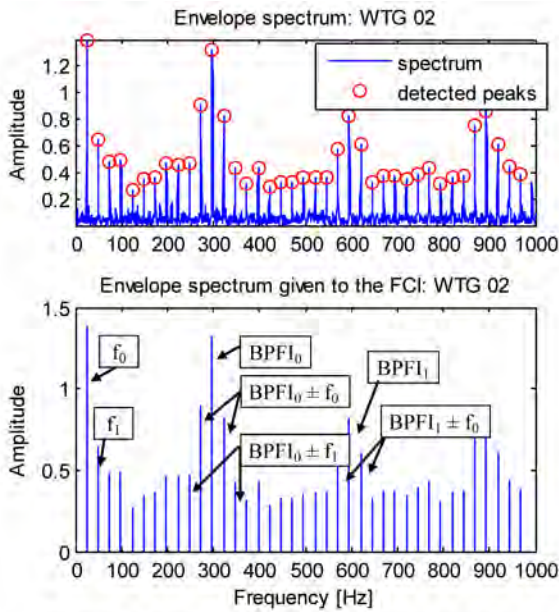


Figure 24: Envelope spectrum and detected peaks (top); envelope spectrum after peak filtering (#3, HSS, meas. 34, WTG 02)

The $BPFI_0$ is around 297.8 Hz, f_0 around 24.8 Hz and $BPFO_0$ 221.6 Hz. After FCI, the following entries in Table 4 contain a “1” for frequency content being present.

- | | |
|------------------------|-----------------------|
| • RPM_h0 (f_0) | • BPFO_h2 |
| • RPM_h1 (f_1) | • BPFO_h3 |
| • BPFI_h0 ($BPFI_0$) | • BPFO_h4 |
| • BPFI_h1 ($BPFI_1$) | • FTF_h4 |
| • BPFI_h2 | • BSF_h2 |
| • BPFI_h3 | • Mod_RPM_h0_BPFI_h0 |
| • BPFI_h4 | • Mod_RPM_h0_BPFI_h1 |
| • BPFO_h0 | • Mod_BPFO_h0_BPFI_h |
| • BPFO_h1 | • Mod_BPFO_h0_BPFI_h1 |

The following four rules were established based on common faults pattern of different failures:

- If BPFO_h0 & BPFO_h1 & Mod_RPM_h0_BPFO_h0 & Mod_RPM_h0_BPFO_h1 then outer race signature + unbalance or misalignment
- If BPFO_h0 & BPFO_h1 & BPFO_h2 then outer race signature + unbalance or misalignment
- If BPFI_h0 & BPFI_h1 & BPFI_h2 & Mod_RPM_h0_BPFI_h0 then inner race signature
- If FTF_h0 & FTF_h1 & FTF_h2 & then rolling element signature

Next to the speed and BPFI components also BPFO and BSF components have been detected. The BPFO is around

216 Hz and thus close to $BPFI_{h0} - RPM_{h2}$ (219 Hz). $BPFO_{h2}$ (649 Hz) is close to $BPFI_{h1} + RPM_{h2}$ (653 Hz). A similar explanation can be found for the $BPFO_{h3}$, the $BPFO_{h4}$ and the BSF component.

In classical logic, this fact would give problems in diagnosis, since there is the risk of collision with rule b). However, with FL each rule can be assigned a weight. Although in this example the outer race signature rule would not have fired, it proved useful assigning simple rules (based on simple patterns) a lower weight to avoid misdiagnosis. In such cases the rule with the higher weight would fire.

Applying the rules, rule c) fired indicating an inner race fault at the HSS bearing.

4.2. Example 2

The automated analysis highlighted five more turbines with inner race signature and inner race damage. Figure 25 shows a waterfall plot of the envelope spectra together with the SK trend of the most recent ten measurements of WTG 05.

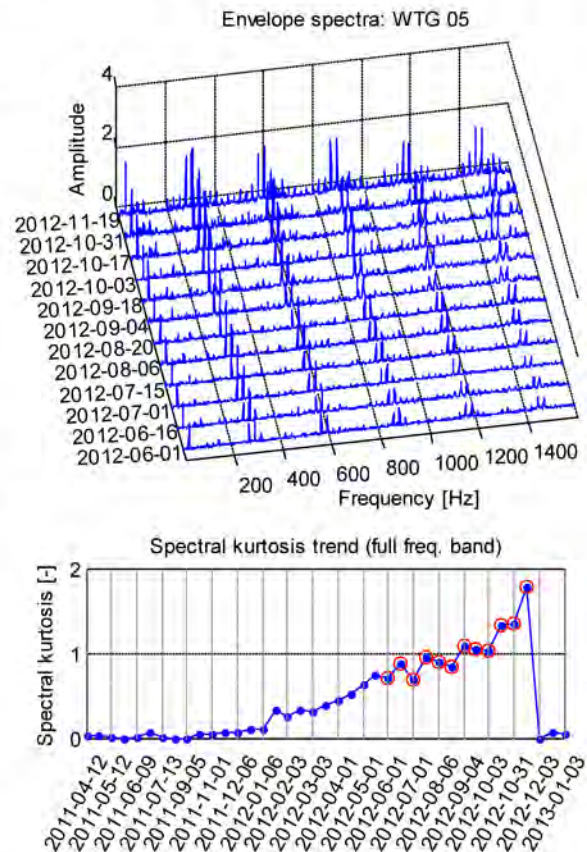


Figure 25: Envelope spectra and spectral kurtosis (WTG 05)

At this turbine a very clear trend in SK is visible since January 2012, with the recent measurements slightly

fluctuating. The envelope spectra follow this trend, with rising amplitudes and more harmonics showing up. Repair took place in November 2012. Plotting the SK trend after MED gives the results presented in Figure 26. Also for this turbine the SK has a high variance in the healthy period. A trend first establishes at 2011-03-18, which is two month later than with the SK values without MED. Hence, for the data sets available, it appears preferable using the SK values without MED for trending. However, this is in contrast to the findings of Sawalhi et al. [5] in which a case is shown where SK trending was not possible without MED.

Figure 27 shows the most recent envelope spectrum, the detected peaks and the filtered spectrum passed to the FCI.

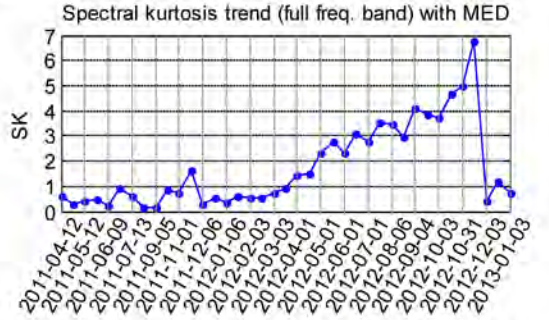


Figure 26: Spectral kurtosis trend after MED (WTG 05)

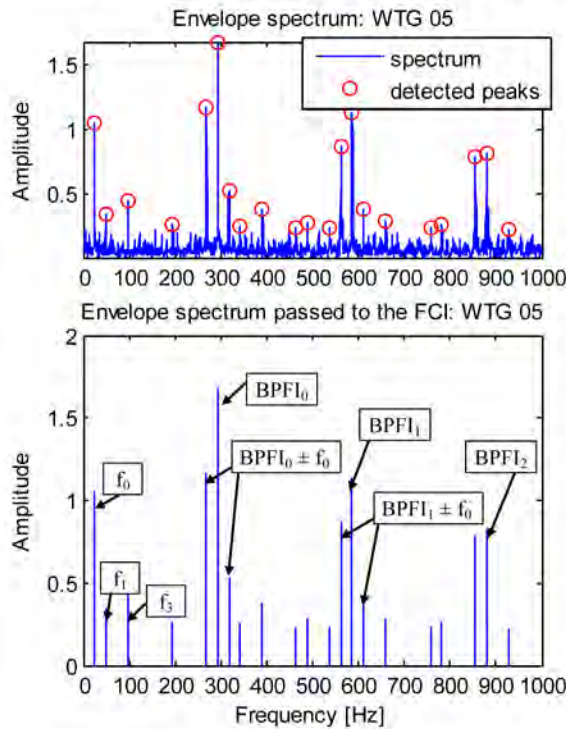


Figure 27: Envelope spectrum and detected peaks (top); envelope spectrum after peak filtering (#3, HSS, meas. 31, WTG 05)

The identified frequency contents by the FCI are:

- RPM_h0 (f_0)
- RPM_h1 (f_1)
- RPM_h3 (f_3)
- BPF1_h0 (BPF1₀)
- BPF1_h1 (BPF1₁)
- BPF1_h2 (BPF1₂)
- BPF1_h3
- BPF1_h4
- BPFO_h3
- BSF_h2
- Mod_RPM_h0_BPF1_h1
- Mod_RPM_h0_BPF1_h1
- Mod_BPFO_h0_BPF1_h2

The present frequency contents evaluated via the rules implemented highlight the inner race signature and the potential damage fully automatized, together with the correct diagnosis.

4.3. Example 3

Another turbine with detected inner race damage is WTG 04. Again the most recent ten envelope spectra and the spectral kurtosis trend is illustrated in Figure 28 and the SK trend with MED in Figure 29.

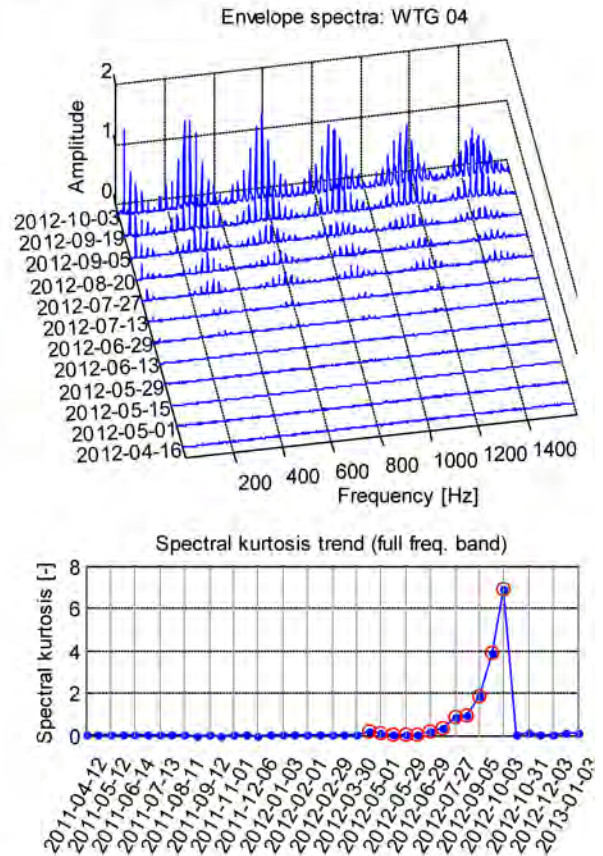


Figure 28: Envelope spectra and spectral kurtosis (WTG 04)

At this turbine many measurements in a healthy state exist, as the crack develops comparably late. However, the SK

sharply increases 2012-06-29 again well in line with the envelope spectra amplitude rise. Bearing replacement took place in October 2012.

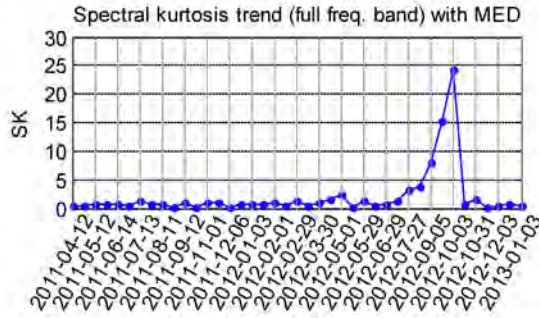


Figure 29: Spectral kurtosis trend after MED (WTG 04)

Also the SK after applying MED shows the steep rise. However, the healthy state is also for this turbine characterized by a larger variance. Because of this a trend is reliably detectable first at 2012-07-27, in contrast to trend establishment in the SK values without MED, where the trend can be identified two weeks earlier.

Figure 30 visualizes the most recent envelope spectrum and the detected beaks together with the envelope spectrum passed to the FCI and important frequency components.

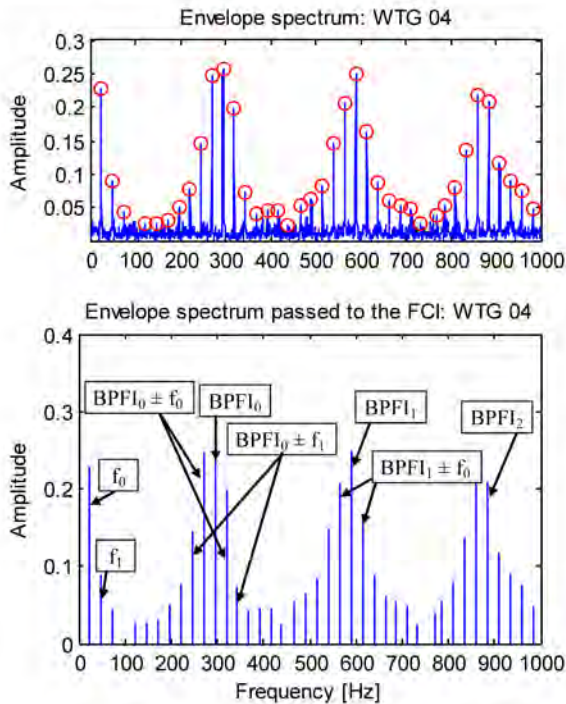


Figure 30: Envelope spectrum and detected peaks (top); envelope spectrum after peak filtering (#3, HSS, meas. 31, WTG 04)

The detected frequency content is:

- RPM_h0 (f_0)
- RPM_h1 (f_1)
- RPM_h2
- BPFI_h0 (BPFI₀)
- BPFI_h1 (BPFI₁)
- BPFI_h2 (BPFI₂)
- BPFI_h3
- BPFI_h4
- BPFO_h0
- BPFO_h2
- BPFO_h3
- BPFO_h4
- BSF_h2
- Mod_RPM_h0_BPFI_h0
- Mod_RPM_h0_BPFI_h1
- Mod_BPFO_h0_BPFI_h0
- Mod_BPFO_h0_BPFI_h1

Based on the frequency content present, an inner race signature was automatically detected, which is the correct diagnosis for the pattern existing.

5. Results and discussion

The given examples emphasize the successful application of the proposed procedure to real measured data and fault diagnosis.

Moreover research showed the benefits of using the CEP based signal PW for signal separation into deterministic and non-deterministic components. The performance of PW is close to the LP residual, but without the need of parameter optimization, which makes it an efficient method and very fast to compute.

Using MED as a further pre-process step prior envelope calculation proved useful to emphasizing present fault pattern both in the time and frequency domain. Attention must be paid, to the choice of filter order. It was found that a too high filter order can lead to start-up and end transients in the time series preventing successful diagnosis. For practical applications, it is thus recommended using a rather low filter order, accepting that the impulse amplification may be less effective than with a higher filter order, but avoiding the risk of transients.

Furthermore the effect of the spectral kurtosis based optimal filter band selection was studied. Results show that in many cases the combination of PW & MED lead to the situation that band pass filtering gives no improvement. However, in some cases a frequency band can be found that gives higher spectral kurtosis values and accordingly a clearer fault pattern in the envelope spectrum. By applying the fast kurtogram (proposed by Antoni [22]) the search for an optimal frequency band is fast to compute and should thus supplement the automated procedure.

Finally the effect of MED on the SK values trend (used as a fault size indicator) was under research. A higher variance of the SK values was observed for bearings in a healthy state when MED is used prior SK value calculation. This variance lead in some cases to a later establishment of observable trends, but both methods (with and without MED) show trends in the fault state, potentially linked to the crack size. The SK without MED appears to be superior to use, because of the more direct link between present patterns in the envelope spectra and high SK values as well as lower variance in the values.

6. Future research

Identifying HSS bearing faults in wind turbine gearboxes is a comparably simple task, because of the high rotational speed. Future research will focus on the efficiency of the proposed method to automatically detect and diagnose faults in the IMS and the planetary stage of wind turbine gearboxes. At least for the IMS stage (rotating with approx. 350 rpm) it is expected that the proposed procedure will give good results. However, the available data sets did not contain any bearing damages of the IMS or the planetary stage. Hence the efficiency for intermediate and low speed bearings must still be evaluated and is left for future research.

7. Conclusion

The proposed procedure proved capable of automatically process and analyse vibration time series for bearing defect frequencies in the envelope spectra. The diagnosis is based on rules which are set up using common knowledge about typical bearing fault signatures. The rules are implemented using fuzzy logic, building the basis for flexible and reliable fault diagnosis.

The proposed technique is an extension of the semi-automated procedure for bearing diagnosis (originally proposed by Sawalhi and Randall [4]) with regards to the application of a frequency content identification and rule based diagnosis. Moreover state of the art pre-processing techniques, namely CEP based signal PW and MED have been successfully applied and their effectiveness shown. In comparison to LP models for separating random from discrete components, CEP based signal PW proved to be an efficient and fast to compute method for separation.

References

- [1] Crabtree, C.J. and Tavner, P.J. (2011). "Condition monitoring algorithm suitable for wind turbine use. *Renewable Power Generation (RPG 2011)*, Edinburgh.
- [2] Randall, Robert B. (ISBN:978-0-470-74785-8, 2011). "Vibration-based Condition Monitoring. *Wiley*.
- [3] Schlechtingen, M., Santos, I.F. and Achiche, S. (2012). "Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 2: Application examples. *Applied Soft Computing*.
- [4] Sawalhi, N. and Randall, R. B. (2007). "Semi-automated bearing diagnostics - three case studies. *Proceedings of the Comadem Conference, Faro, Portugal*.
- [5] Sawalhi, N., Randall, R. B. and Endo, H. (2007). "The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. *Mechanical Systems and Signal Processing Vol.21*, pp.2616-2633.
- [6] Barszcz, T. and Sawalhi, N. (2012). "Fault detection enhancement in rolling element bearings using the minimum entropy deconvolution. *Archives of Acoustics, Vol. 37-2*, pp.131-141.
- [7] Randall, R. B. and Antoni, J. (2011). "Rolling element bearing diagnosis - A tutorial. *Mechanical Systems and Signal Processing Vol.25* pp.485-520.
- [8] Antoni, J. and Randall, R.B. (2004). "Unsupervised noise cancellation for vibration signals: part II a novel frequency-domain algorithm. *Mechanical Systems and Signal Processing Vol.18*, pp.103-117.
- [9] Randall, R. B. and Sawalhi, N. (2011). "A new method for separating discrete components from a signal. *Sound and Vibration Vol. May*, pp.6-9.
- [10] Randall, R. B. and Sawalhi, N. (2011). "Use of the cepstrum to remove selected discrete frequency components from a time signal. *IMAC conference, Jacksonville*.
- [11] Sawalhi, N. and Randall, R.B. (2011). "Signal pre-whitening using cepstrum editing (liftering) to enhance fault detection in rolling element bearings. *Proceedings of the 24th International Congress on Condition Monitoring and Diagnostics Engineering Management*, pp.330-336.
- [12] Niaoqing, H., et al. (2012). "Enhanced fault detection of rolling element bearing based on cepstrum and stochastic resonance. *Journal of Physics: Conference Series Vol.364*, pp.1-8.
- [13] Dyer, D. and Stewart, R.M. (1978). "Detection of rolling element bearing damage by statistical vibration analysis. *Journal of Mechanical Design Vol.100*, pp.229-235.
- [14] Bonnardot, F., Randall, R. B. and Antoni, J. (2004). "Enhanced unsupervised noise cancellation using angular resampling for planetary bearing fault diagnosis. *International Journal of Acoustics and Vibration Vol.9-2* pp.51-60.
- [15] Barszcz, T. (2009). "Decomposition of vibration signals into deterministic and nondeterministic components and its capabilities of fault detection and identification. *International Journal of Applied Mathematical Computation Vol.19-2* pp.327-335.
- [16] Oppenheim, A.V. and Schaffer, R.W. (2004). "From frequency to quefrency: a history of the cepstrum. *IEEE Signal Processing Magazine, September*, pp.95-104.
- [17] Randall, R.B. and Hee, J. (1981). "Cepstrum analysis. *Brüel & Kjaer, Technical Review Vol.3*, ISSN:0007-2621.
- [18] Wiggins, R. A. (1978). "Minimum entropy deconvolution. *Geoexploration Vol.16*, pp.21-35.
- [19] McDonald, G. *Minimum Entropy Deconvolution*. [Online] 04 2012. <http://www.mathworks.com/matlabcentral/fileexchange/29151-minimum-entropy-deconvolution-med-1d-and-2d/content/med2d.m>.
- [20] Dwyer, R. F. (1983). "Detection of non-Gaussian signals by frequency domain Kurtosis estimation. *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '83*, pp.607-610.
- [21] Antoni, J. and Randall, R. B. (2006). "The spectral kurtosis: application to the vibratory surveillance and

diagnostics of rotating machines. "Mechanical Systems and Signal Processing Vol.20, pp.308-331.

[22] **Antoni, J.** (2006). "Fast computation of the kurtogram for the detection of transient faults. "Mechanical Systems and Signal Processing Vol.21, pp.108-124.

[23] **Antoni.** Programmes. [Online] 04 2012. <http://www.utc.fr/~antoni/programm.htm>.

[24] **Howard, I.** (1994). "A Review of rolling element bearing vibration "detection, diagnosis and prognosis". "Aeronautical and Maritime Research Laboratory, DSTO-RR-0013.

[25] **Akaike, H.** (1973). "Information theory and an extension of the maximum likelihood principle. "Second International Symposium on Information Theory, pp.267-281, Budapest.

[26] **Palshikar, G. K.** (2009). "Simple algorithms for peak detection in time-series. " Proceedings of 1st IIMA International Conference on Advanced Data Analysis, Business Analytics and Intelligence, Ahmedabad, India.

[27] **Sawalhi, N. and Randall, R. B.** (2008). "Helicopter gearbox bearing blind fault identification using a range of analysis techniques. "Journal of Mechanical engineering Vol.5, pp.157-168.

[28] **Sawalhi, N. and Randall, R.B.** (2008). "Novel signal processing techniques to aid bearing prognosis. "IEEE PHM Conference, Denver, Co, USA, Oktober.

Appendix

List of abbreviations

AIC	=	Akaikes entropy based information criterion
ANC	=	Adaptive noise cancellation
AR	=	Autoregressive
BPFI	=	Ball pass frequency inner race
BPFO	=	Ball pass frequency outer race
BSF	=	Ball spin frequency
CEP	=	Cepstral editing procedure
DRS	=	Discrete/random separation
FCI	=	Frequency content identifier
FL	=	Fuzzy logic
FFT	=	Fast Fourier transformation
FTF	=	Fundamental train frequency
FRF	=	Frequency response function
HSS	=	High speed shaft
IMS	=	Intermediate speed shaft
LP	=	Linear prediction
MED	=	Minimum entropy deconvolution
MF	=	Membership function
OT	=	Order tracking
PSD	=	Power spectral density
PW	=	Pre-whitening
RMS	=	Root mean squared
RPM	=	Revolutions per minute
SANC	=	Self adaptive noise cancellation
SCADA	=	Supervisory control and data acquisition
SK	=	Spectral kurtosis
SRM	=	Stochastic resonance model
STFT	=	Short time Fourier transform
TS	=	Time series
TSA	=	Time synchronous average
WTG	=	Wind turbine generator

4.1.1. Extended research and development - Additional information

Research showed that based on the spectral kurtosis a suitable bandwidth for filtering can be identified without prior knowledge. This property is well suited to wind turbine applications, where usually no prior knowledge exists. The spectral kurtosis evolution prior MED allows fault identification at an early stage by using this simple parameter.

However, it has been shown in publication [P5] (example 1) that if the defect proceeds the SK begins decreasing. This behavior is unsatisfactory as at a later stage of defect, the fault may be classified wrongly, when only considering the actual SK value. Further research was performed to avoid the SK decrease. It has been found, that calculating the Kurtosis of the Envelope Amplitude Spectrum (KEAS) (signal separated by signal Pre-Whitening (PW)) gives a more stable parameter than monitoring the SK. Hence, the parameter is calculated as:

$$KEAS = \frac{\sum_{i=1}^M (B_i - \bar{B})^4}{(M-1)\sigma^4} \quad (5)$$

σ is the standard deviation; M is the number of spectral data points; B_i is the i th frequency component of the envelope amplitude and \bar{B} is the mean of the envelope amplitude spectrum. The main difference between the SK parameter of the full frequency band the KEAS is the performed transforms. While the SK here is calculated from the filtered (full band filter) pre-whitened signal, the KEAS is calculated from the envelope of this signal. Therefore SK (as it is used here) is a measure of the peakiness of the filtered signal, while the KEAS is a measure of the peakiness of the envelope spectra.

Based on six HSS bearing defects (introduced in [P5]) and five IMS bearing defects the effectiveness of this parameter in comparison to the SK is discussed in the following paragraphs. During normal operation, the approximate rotational speed of the HSS shaft is 1450 rpm, respectively 370rpm for the IMS shaft. At slower rotational speeds, the impulse from defect overrollment contains little energy reducing the reliability of defect detection. For the IMS bearing defects, also the effect of signal PW is investigated. Xiaofei [32] reported good performance of the algorithm for a rotational speed of 628 rpm. The performance of signal PW was not yet reported for bearings of lower rotational speed. Hence the result presents a novelty and aid to further validation of the method.

4.1.1.1. HSS bearing defects

The bearing defects present in the available data sets are inner race defects. At the concerned wind turbines straight axial (white edging) cracks occurred at the raceway. Figure 4.1 shows the axial cracks at the concerned wind turbines. At some of the WTGs additionally pitting of the surface is observed.

In Figure 4.2 a comparison is made between the SK and the KEAS evolution for bearing defects at the six WTGs. For analysis, the data of the Time Series (TS) 16000 (see Table 1.3) is used. This data corresponds to the measurements labeled #3 in [P5]. Due to the short sample length, the speed fluctuations are negligible. Analogously to the procedure in publication [P5] the SK is calculated from the full frequency band after signal PW.

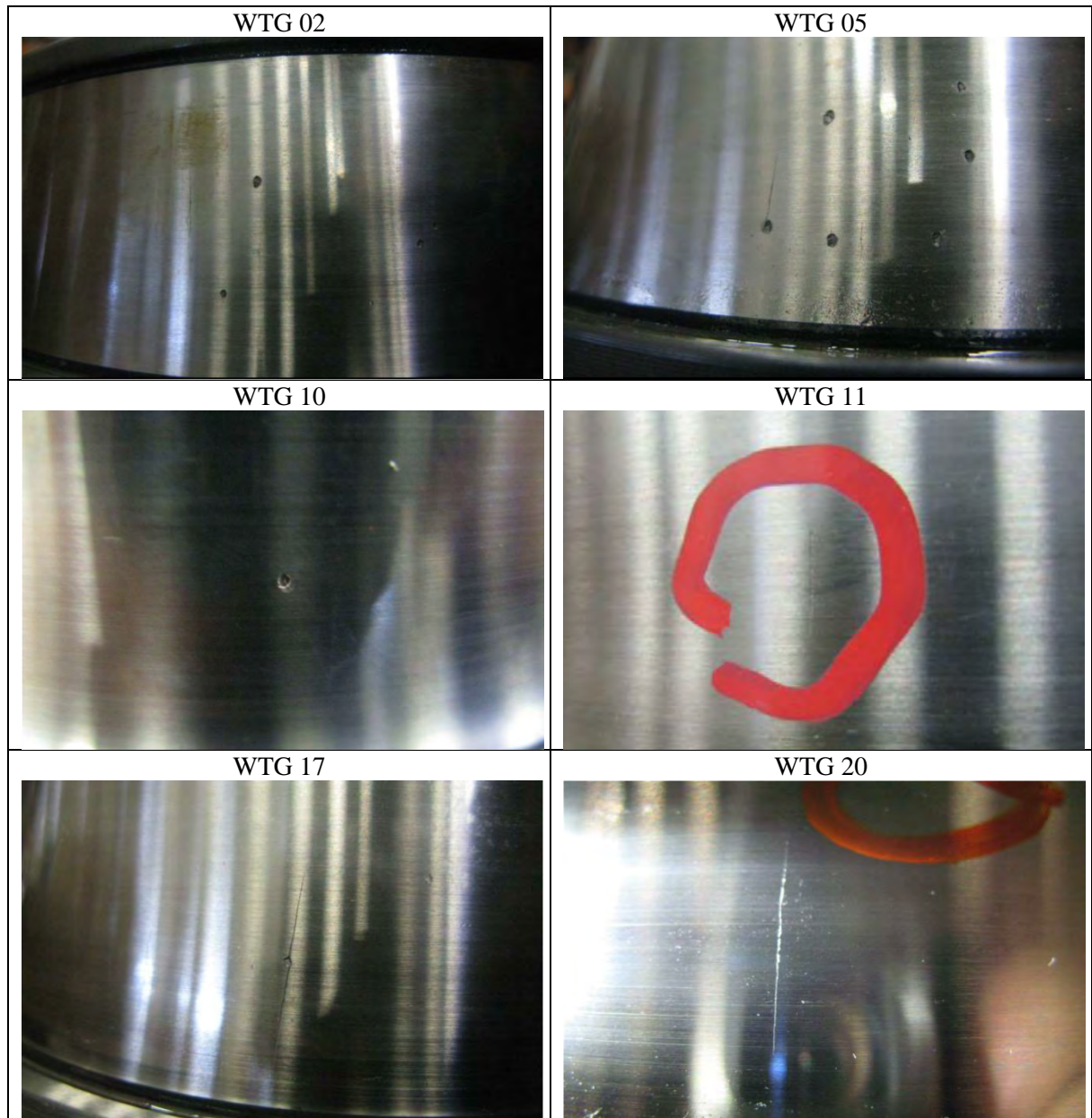


Figure 4.1: Axial cracks of the concerned wind turbines

At WTG 02, the remarkable difference between the two parameters evolutions is that the KEAS values do not decrease beyond 2012-08-19, in contrast to the SK values. The decrease in SK is potentially caused by the fact, that while the fault extends the individual component amplitudes (defect frequencies and its harmonics) tend to diminish in the filtered signal (see [P5] for further explanation) together with an increase in noise. This leads to decreasing kurtosis values. With inner ring defects, rich modulations appear, while the faults progresses. The reason for modulations occurring is the defect passing through the load zone, which modulates the component amplitude (the amplitude is high in the load zone and low at its opposite position). In the envelope spectrum, the modulations are present as distinct peaks and hence contribute to a high kurtosis value. However, in the actual case of WTG 02 the high kurtosis value is caused by an increase in the fundamental shaft component. Figure 4.3 visualizes the squared envelope spectra of WTG 02 at two different defect stages.

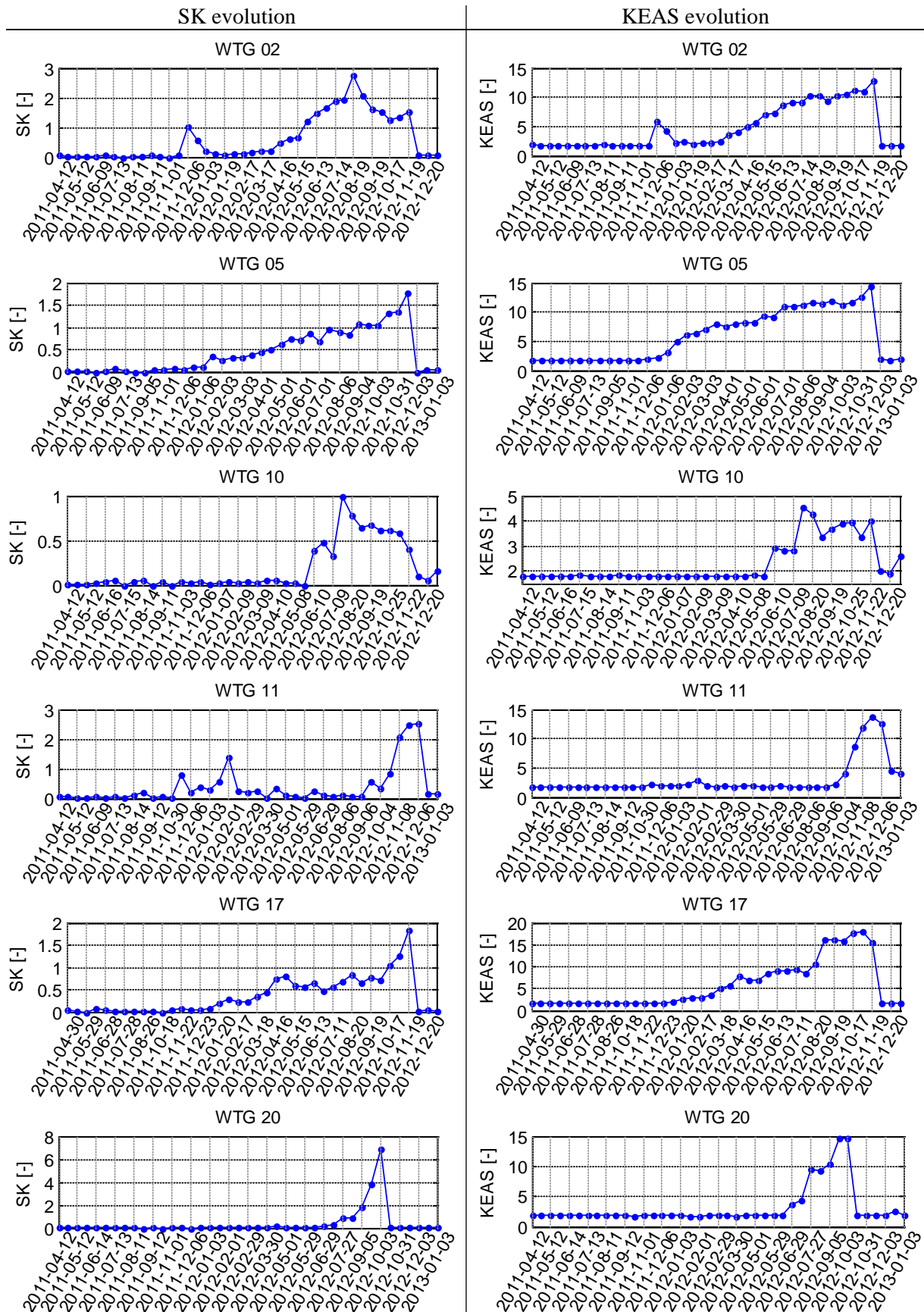


Figure 4.2: Comparison between SK and KEAS evolution (HSS stage: TS 16000)

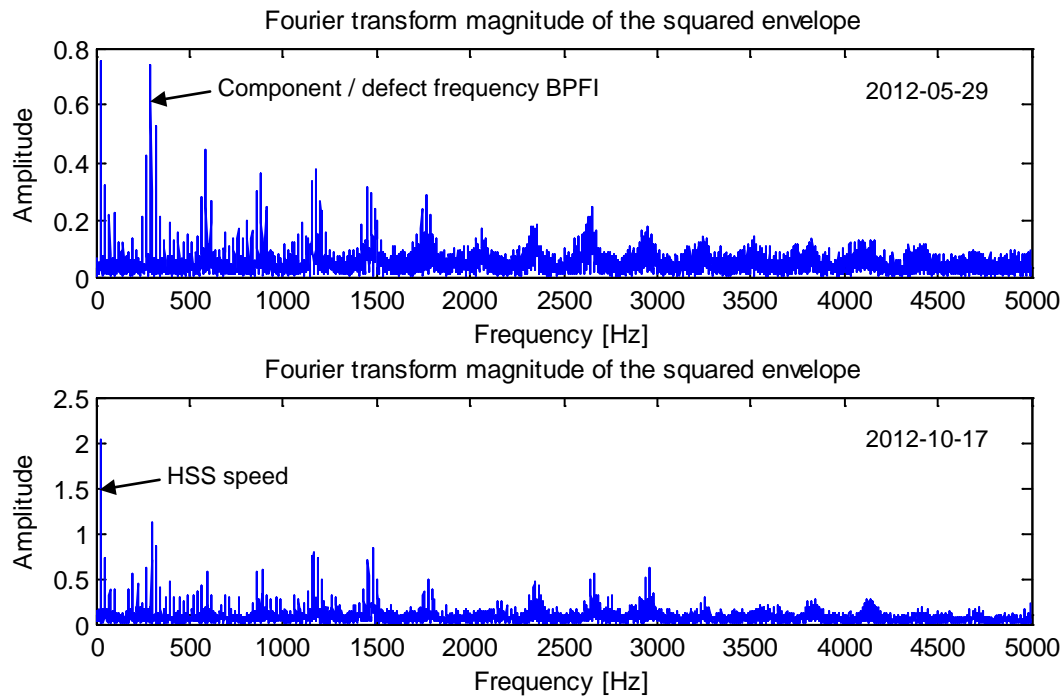


Figure 4.3: Squared envelope spectrum of WTG 02; top: at 2012-05-29; bottom at 2012-10-17

From the discussion it is clear that the component frequency amplitude in the envelope will also diminish at a later defect stage, but in terms of the kurtosis this effect is to some extent counteracted by the modulations and (in case of WTG 02) an increase in the fundamental shaft component. In the consequence the KEAS is likely to continue increasing or keeping a high level for a longer period than the SK value.

At WTG 10 the KEAS roughly maintains its level, while the SK values decrease beyond 2012-08-20. Here the same argumentation applies, but the increase in the fundamental shaft component is not as pronounced as at WTG 02.

At WTG 11, the KEAS evolution shows fewer fluctuations in the measurements between 2011-10-30 and 2012-02-29. In the envelope spectra of this period, the present peaks cannot be clearly dedicated to a fault and hence the KEAS parameter appears more trustable.

AT WTG 20 the KEAS grows at 2012-06-29, whereas the first high SK value is present at 2012-07-27, i.e. four weeks later. Figure 4.4 illustrates the squared envelope spectra at 2012-06-29 and the present defect frequency pattern, allowing identification of the inner race defect.

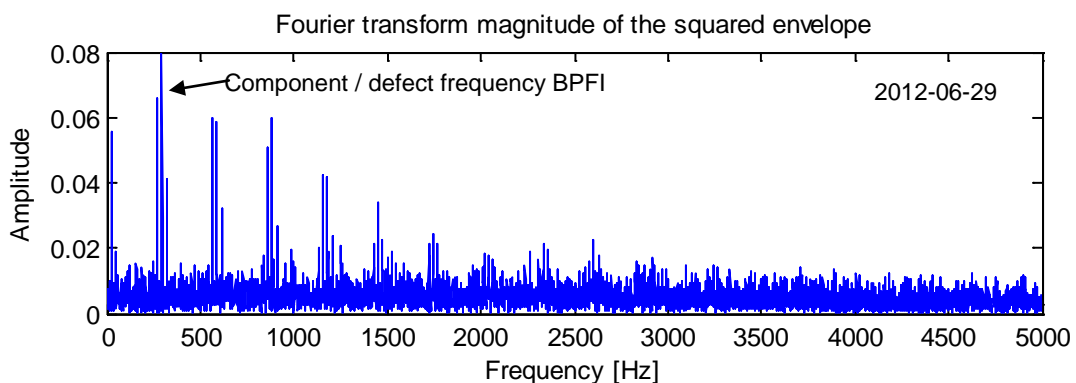


Figure 4.4: Squared envelope spectrum of WTG 20 at 2012-06-29

At WTG 05 and WTG 17, the performance of the two parameters is similar.

In summary, it can be concluded from the surveyed cases, that the KEAS value gives at least similar results as the SK value. Some improvements concerning the value decrease at a later defect state and earlier fault identification could be observed. However, it should be noted, that the surveyed cases represent only one particular fault pattern. Hence in the future further performance verification is required with different defect types, e.g. spalling or outer race defects, which was not possible in this research due to lack of cases.

4.1.1.2. IMS bearing defects

The IMS bearing defects considered in this work are also exclusively inner race defects with straight axial cracks on the raceway.

The effect of PW

Figure 4.5 shows a comparison between the squared envelope spectra for the following three cases analogously to Figure 10 presented in publication [P5], but for the IMS defects: no processing, linear prediction and signal PW. The involved preprocesses are PW or LP.

The unprocessed envelope spectra are dominated by deterministic frequency components (a). The characteristic fault pattern, i.e. high BPFI (77.6 Hz) surrounded by sidebands (± 6.4 Hz) is not visible. At the HSS defects surveyed, this pattern was also visible in most envelope spectra without processing although less pronounced. With IMS defects, as mentioned previously, the impulses contain less energy and preprocessing techniques gain importance. The performance of PW and linear prediction is comparable, with little difference visible in the spectra (b & c). All deterministic components have been removed successfully. At WTG 08, the fault pattern is shown at a later stage, with the component frequency (BPFI) already decreased and rich modulation sidebands present. The kurtosis values of the preprocessed signal are summarized in Table 4.1 for the IMS and in Table 4.2 for the HSS defects.

Table 4.1: Kurtosis of preprocessed time series (IMS inner race defect)

WTG	Linear prediction	Pre-whitening
01	4.42	4.93
08	4.91	5.75
11	3.80	4.22
21	3.62	3.97

Table 4.2: Kurtosis of preprocessed time series (HSS inner race defect)

WTG	Linear prediction	Pre-whitening
01	4.58	3.92
02	12.89	8.53
05	9.19	7.09
10	6.34	5.18
17	7.63	5.53
20 ¹⁾	9.13	6.24

¹⁾ Corresponds to WTG 04 in [P5]

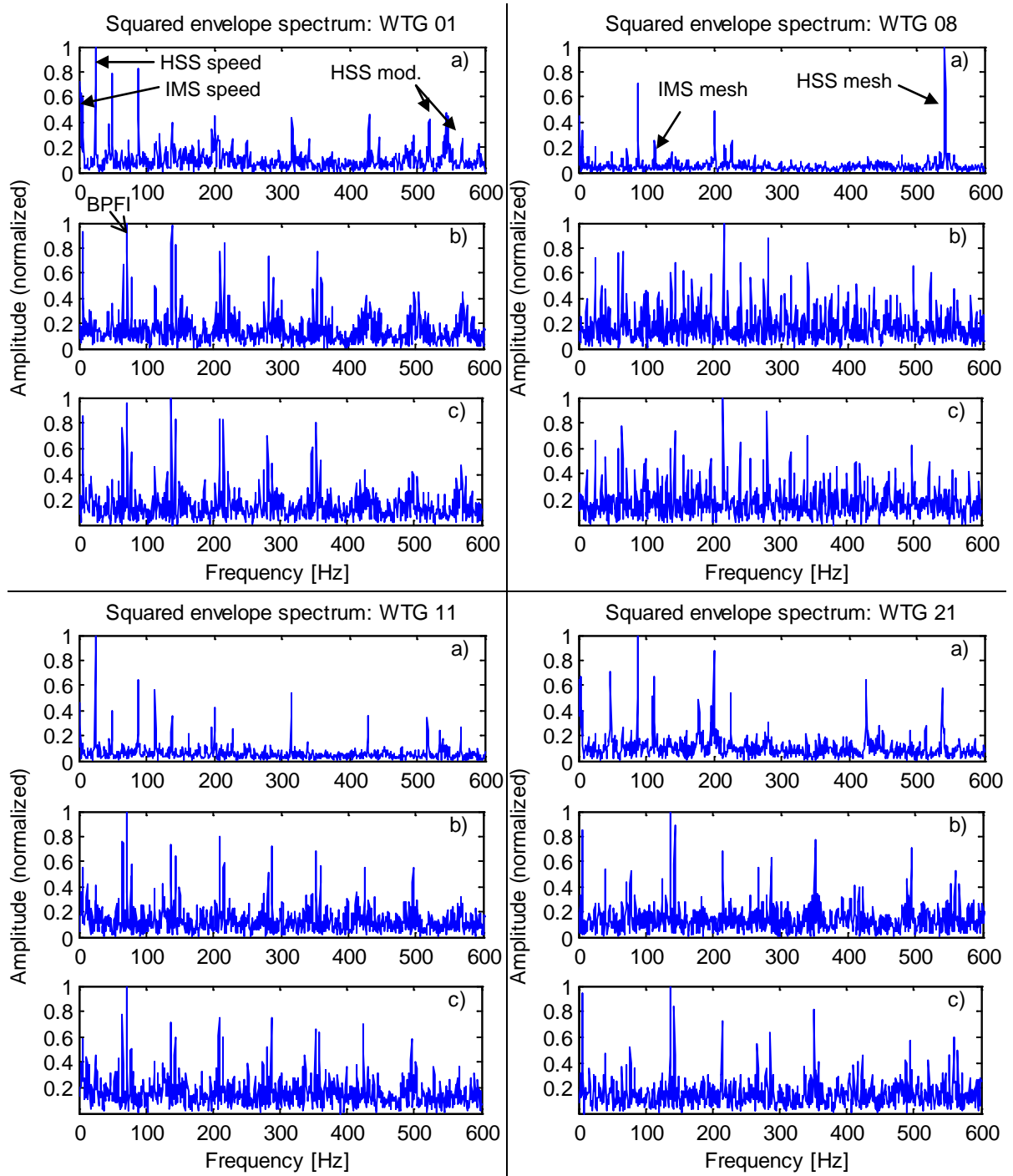


Figure 4.5: The effect of signal PW; a) no processing, b) linear prediction, c) PW; (TS 16000, IMS stage, meas. 40)

With the IMS defects, the PW procedure leads to higher kurtosis values. This is in contrast to the performance of PW with the HSS defects, where lower kurtosis values can be observed. A similar observation is made by Sawalhi and Randall [33] for the case of a helicopter gearbox. A further difference can be observed at the LP model order. While for the HSS defects Akaike's entropy based Information Criterion (AIC) [34] lead to an optimal model order of 598, it is around 796 for the IMS defects. This higher model order indicates lower deterministic frequency components in the IMS TS, which is due to the lower rotational speed.

Sawalhi and Randall [33] argue that the higher kurtosis values achieved by LP method in their case is caused by the optimization procedure used, which was set to maximize kurtosis. Here it is shown that the same behavior can be seen when using AIC to optimize the model order. AIC compromises the model complexity and model performance or accuracy.

The higher model order at the IMS defects thus leads to a high penalty for model complexity and a lower model performance.

However, the comparison of envelope spectra gave no major difference in performance of PW and LP, leading to the conclusion that the higher kurtosis achieved in the processed time signal does not give benefits in the envelope spectra. Furthermore, it can be concluded that the kurtosis of the TS is not a good measure to evaluate the procedure performance. Instead, the visibility of the fault pattern in the envelope spectra should be used as criteria. In summary it can be concluded from this research that PW is a very efficient algorithm with no need of prior knowledge and good performance with respect to the surveyed IMS and HSS defects.

SK and KEAS evolution as fault indicator

From the five WTGs with IMS defect, three can be used for SK and KEAS comparison, since at the other two HSS bearing defects are present at the same time, interfering the result. The comparison is made in Figure 4.6.

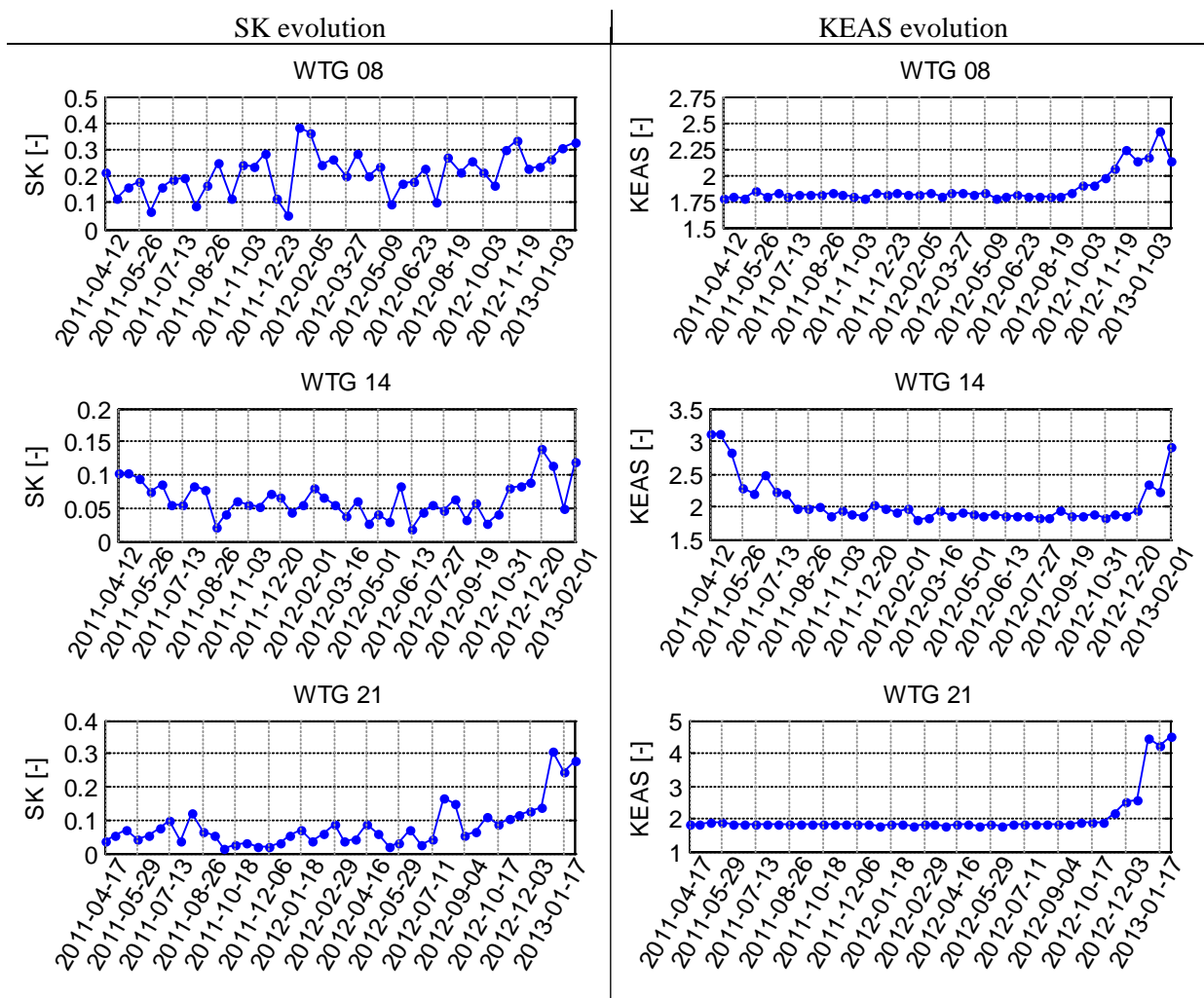


Figure 4.6: Comparison between SK and KEAS evolution (IMS stage TS 6000_Tr)

The comparison emphasizes that the SK parameter is rather insensitive to this type of IMS bearing fault. The large variance in a fault free situation prevents reliable early fault detection. An exception here builds WTG 21, at which the latest three measurements are exposed. The KEAS value on the other hand is characterized by a low variance in a fault free situation, compared to the increase in parameters, when a fault is present.

At WTG 14, the first eight measurements are contaminated with the FTF of the HSS bearing and its shaft speed component, causing the high KEAS values. After some run in period, these components vanish.

The results obtained from the comparison of the SK and KEAS parameter with the HSS bearing defects can thus be confirmed. The KEAS parameter proves useful in defect detection and allows earlier fault detection than the SK value.

4.2. Spur gears

In this chapter, the research performed to identify suitable algorithms to monitor wind turbine spur gear defects is summarized. Wind turbine gearboxes are usually built of multiple stages. One or more planetary stages are combined with spur gear stages in order to achieve high transmission ratios of about 1:90. Fault diagnosis in planetary gears based on vibration measurements requires different algorithms and should be a topic of future research work due to the problem complexity. Moreover, the available data sets are believed to be too short to pick up faults. It should be stated here one more time that this section is included for work completeness sake only, since the parameters derived are used by the global fuzzy expert system of chapter 5. Due to missing faults in the period of research, no complete validation could be performed. For this reason, this section is kept brief and unfortunately does not allow for generalization.

A performance comparison is made between different algorithms based on false brinelling (early scuffing¹) present at the HSS pinion and the IMS gear at WTG 18 of power plant 2. The WTG is reported to be very noisy. Figure 4.7 shows the false brinelling present at this turbine.

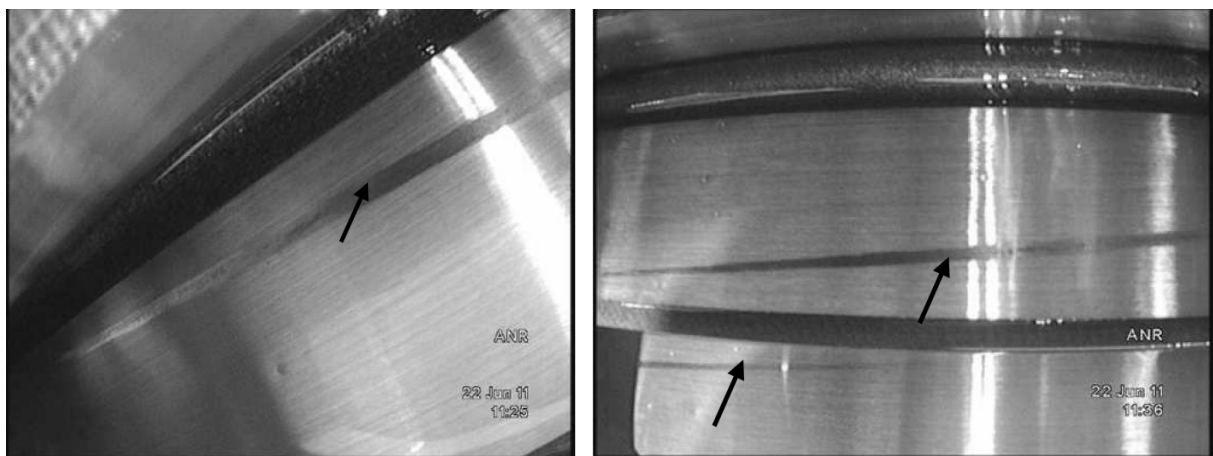


Figure 4.7: False brinelling at HSS pinion (left); at IMS gear (right) (pictures taken during endoscopic survey at 2012-08-14)

¹ Scuffing is a damaging condition that can happen when relative movement occurs between gear tooth contact surfaces upon loss of lubricant film [60].

There are different methods developed in the last decades to identify gear defects. In 2000 Dalpiaz, Rivola and Rubini [35] compared some of the most effective techniques experimentally using test rig data. The methods compared are:

- Cepstrum Analysis (CA) [36] [31] [35] [37] [38]
- CycloStationary Analysis (CSA) [35] [39] [40] [41]
- Time Synchronous Averaging (TSA) [42] [31] [35] [37] [38] [43] [44]
- Wavelet Transform (WT) [45] [35] [37] [46] [47]

They conclude, that a successful diagnostic tool should integrate several of these techniques [35]. Each of the techniques has its advantages and disadvantages, which will be further discussed in this section. In 2013, Siegel [38] compared several methods for WTG bearing and gear defect analysis. For gears the surveyed methods were frequency domain methods, CA, TSA and a planet separation method. The gearbox was exposed to a loss of oil event and many defects were present. With regard to spur gears, the faults under research were misalignment, scuffing, fretting corrosion² and polishing wear³. The drawn conclusion from this comparison is that for gear wheels in particular, the use of multiple algorithms appears necessary, considering the different number of gear failure modes [38].

In addition to the compared algorithms, the following algorithms can be found in literature for gear diagnosis.

- Transmission Error (TE) [31]
- Sideband Energy Ratio (SER) [48]

As the available data set contains only one type of local defect, a full validation is impossible and left for future research. Here some of the most promising algorithms based on the false brinelling present are implemented serving as further information sources in the global condition monitoring system.

In the following, each of the algorithms is briefly introduced and its application and effectiveness with regard to the false brinelling discussed. Note that the comparison is based on one type of defect at one turbine only. Hence, the results obtained are not universal and require further investigation once more cases exist.

4.2.1. Cepstrum analysis

The Cepstrum Analysis (CA) and its properties were already introduced in publication [P5], when describing the pre-whitening procedure for bearing diagnosis. For gear diagnosis, the CA can be useful in interpreting the spectrum, as a tool for detection of periodic structures [35]. It efficiently collects deterministic signal components, such as the meshing components and its sidebands. For gear diagnosis, the sidebands are particular useful as their spacing localizes the faulty component (sideband spacing equals the shaft rotational speed).

² Fretting corrosion is a surface defect of tightly contacting surfaces caused by repeated extremely small movements of one contact surface on another in the presence of corrosion [61].

³ Polishing is a specific kind of abrasive wear during which microirregularities of contact surfaces are worn gradually and surfaces acquire a mirror luster [61].

During progression of a local fault, the sideband amplitudes usually rise, as the fault gives a once per revolution modulation. Each family of sidebands, emphasized by the logarithmic scale, produces a peak in the cepstrum at a quefrequency⁴ corresponding to the reciprocal of the spacing of the spectrum components, as well as other peaks at multiple quefrequencies, called rahmonics [35]. Thus, the quefrequency of the fundamental cepstrum component represents the average sideband level over the whole spectrum [35].

The higher-order rahmonics contain information about the shape of the spectrum, which depends not only on the type of modulation phenomenon, but also on the dynamic properties of the mechanical system and transmission path between vibration sources and transducer location [35]. As a consequence rahmonics are generally less meaningful for monitoring and diagnostics than the fundamental component; this fact gives the obvious advantage that the averaged information concerning a whole sideband family is basically represented with high accuracy by one cepstrum peak, that is generally rather insensitive to transmission path and other secondary effects [49]. More information about the cepstrum can be found in reference [36] [50] [31] [51].

The diagnostic parameter derived from this is the quefrequency amplitude evolution for each gear component. This parameter is the amplitude height in the cepstrum of the quefrequency representing the sideband energy. For this purpose a narrow band around the quefrequency of interest is laid and the maximum amplitude value in this band selected. The bandwidth is $\pm 0.5\%$ of the expected quefrequency. In the actual case of the vibration data available, this parameter is based on the TS 6000_Tr, in which speed variations are encountered. Figure 4.8 shows the real cepstrum amplitude evolution of the HSS pinion and IMS gear for WTG 18.

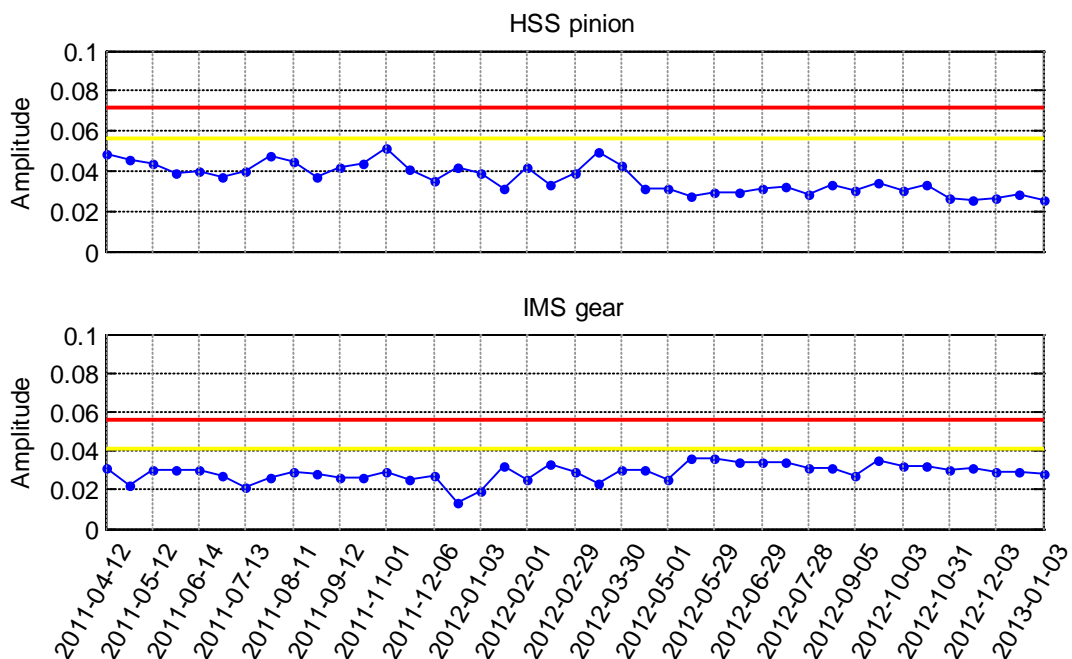


Figure 4.8: Quefrequency amplitude evolution of the HSS pinion and IMS gear of WTG 18

⁴ The word quefrequency originates from the word frequency. The inverse of a quefrequency can be related to a frequency content in a spectrum.

Given enough data sets, a certain amount of measurements (e.g. 20) can be used to calculate the SD, which in turn can be used to define relative thresholds.

$$thres = \bar{q} + \gamma\sigma(q) \quad (6)$$

With t being the threshold, q being the reference quefreny amplitudes, γ being a multiplier (here three and six) and σ the SD.

The quefreny amplitudes indicate no defect progression or change in condition, as the amplitude evolution shows no upwards trend. The reason for this the false brinelling occurring prior commissioning, hence no fault free TS exist. However, the absolute amplitude level can aid in issue identification, as the power plant average corresponding amplitude level is around 0.02 for both gear components.

4.2.2. Cyclostationary analysis

Stationary signals maintain their statistical properties over time. An example for such a signal is the vibration signal of an unbalanced rotor disc running at constant speed. Though in reality true constant speed is hard to maintain, because of minor speed fluctuation virtually always being present. At vibration measurements of mechanical systems, true stationary is seldom seen. Instead, the measurements often contain some form of cyclic periodicity, i.e. their statistical properties change over time with some cyclic pattern. Vibration measurements obtained from gears, are cyclostationary as the pattern, e.g. originating from tooth meshing induces cyclic impacts, causing a system response.

Cyclostationary analysis aims at extracting the cyclic features. A useful tool here is the spectral correlation. However, due to its high complexity and time limitations the cyclostationary analysis procedure was not implemented and evaluated in this research. More information about cyclostationary analysis can be found in reference [41] [40] [52].

4.2.3. Time synchronous average

Time Synchronous Averaging (TSA) is a powerful tool to extract periodic components from a signal. Based on the period length of the desired periodic content, synchronous averaging can be performed to extract the average vibration pattern. For gear diagnosis, it can be used to average the signal over one revolution. In particular, the TSA shows the pattern of the tooth-meshing vibration over one revolution, including any modulation effects, whilst its spectrum contains the meshing components and the sideband family associated with the gear of interest [35]. To encounter speed fluctuations in the time signal, which would smear the result, order tracking is usually necessary.

However, with the given data sets order tracking is not feasible, as no once per revolution trigger or other signal indicating the angular position is available. Instead, the data sets contain tracked time series. These time series are still in the time domain.

An algorithm has been developed to obtain the TSA signal without exact information about the angular position. The algorithm flow diagram is visualized in Figure 4.9. In this diagram N is the number of teeth and y the TS.

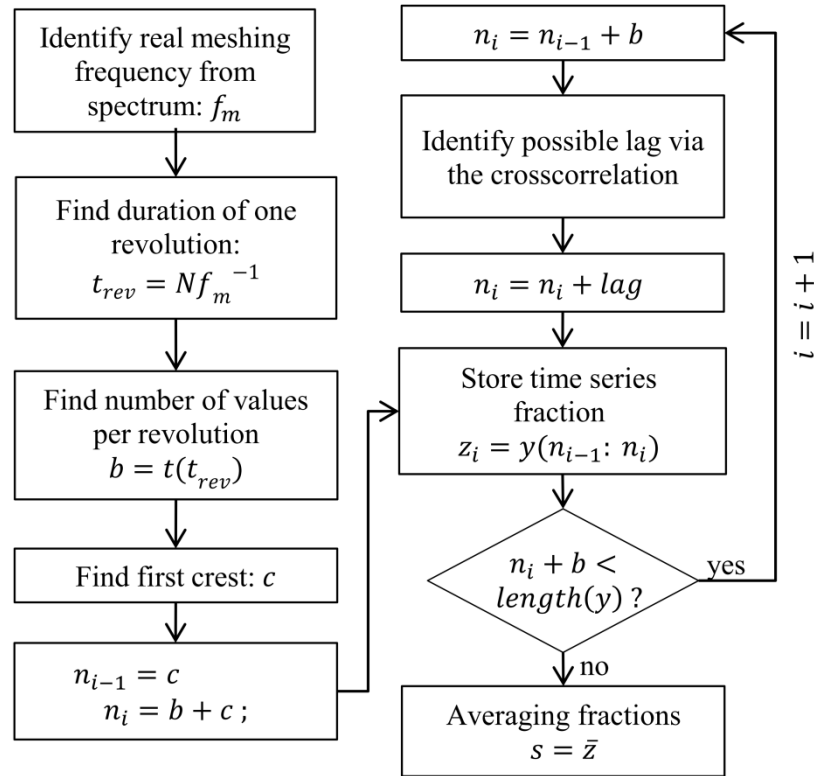


Figure 4.9: TSA procedure of tracked time series without angular position signal

The lag removal via the cross correlation function is useful to prevent a phase shift developing, caused by inaccuracies in period length determination. In the resulting TSA, the impact of each tooth is visible. All other components average out (provided a sufficient number of revolutions are averaged). Figure 4.10 illustrates the raw time series and the TSA signal synchronised with the rotation of the high-speed shaft gear with 22 teeth for an example WTG of power plant 2.

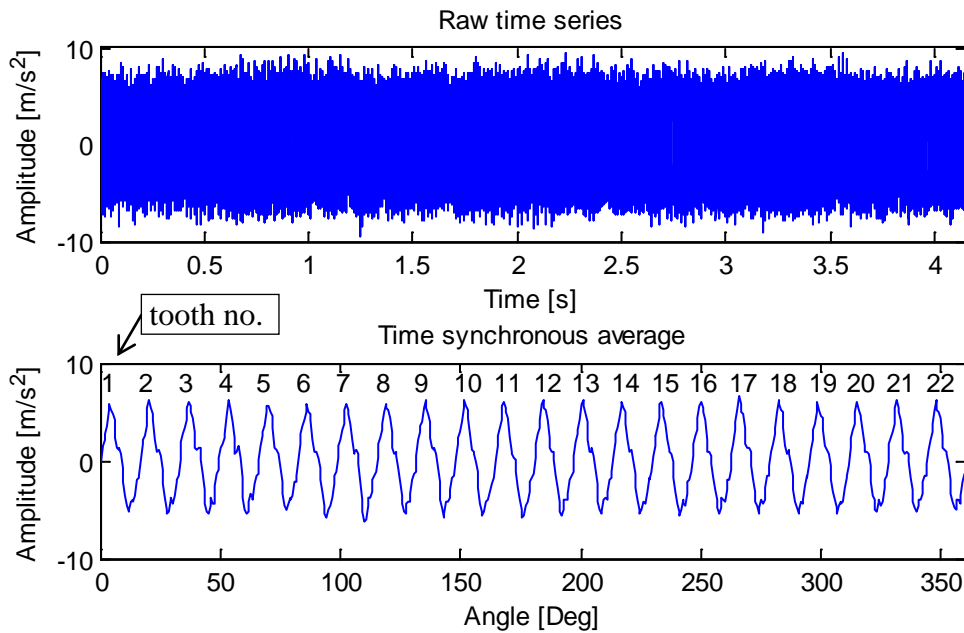


Figure 4.10: Raw time series and the resulting time synchronous average ($N = 22$)

Based on a once per revolution trigger and order tracking, it would be possible to maintain the angular position of each tooth for the different measurements. Without such a trigger, the tooth numbering is only valid for one measurement, i.e. a fault will visually be located at different angular positions for different measurements. In turn, this means, that in case of fault, the whole contour of the wheel must be inspected for faults, as the fault cannot be located. This drawback cannot be counteracted with the measurements at hand.

For local faults in a quite advanced stage of development, the simple visual inspection of the TSA may allow fault detection and localization: the faulted teeth will be characterized by alternations of vibration pattern with respect to the remaining sound part of the gear [35].

To enhance fault visibility more sophisticated analysis methods based on the TSA were developed in the past. They build up on the fact, that for fault analysis the sidebands and modulations contain more information than the regular meshing frequency by which the TSA is dominated.

Hence, techniques have been developed to eliminate the regular meshing pattern to obtain a residual, which indicates deviations from the regular meshing pattern. Popular techniques are:

- Ideal Filtering (editing of the spectrum)
- Linear prediction

In 1986, McFadden [53] proposed a further technique based on amplitude and phase demodulation of the residual. Each of the techniques will be evaluated in the following.

4.2.3.1. Ideal filtering

After time synchronous averaging the resulting signal approaches true periodicity (if a sufficient number of revolutions is averaged). Hence filtering can be achieved by editing the spectrum, without the need of windowing. For this purpose, the spectrum is calculated using FFT. After removing the undesired discrete components, the signal is converted back to the time domain axis by reverse mapping, i.e. the inverse FFT. This procedure does result in the least disruption of the residual signal, though this is rarely necessary [54]. It is limited to the removal of harmonics only and will not, for example, remove modulation sidebands unless these are also harmonics of one of the fundamental frequencies [54]. The process also only works over the full frequency band, including zero frequency, and cannot be used for partial bands (zoom) [54]. More information can be found in [31] and [55].

Using ideal filtering (spectral editing) of the meshing component and its harmonics, Figure 4.11 shows the TSA signal of WTG 18 for the HSS pinion and the residual after spectral editing of the meshing frequency and its harmonics.

The regular meshing frequency pattern has been successfully removed from the synchronous average. What remains is the departure from this normal pattern. Although, already in the TSA signal disturbances are visible, the fault visibility is enhanced by ideal filtering. The spectrum before and after filtering is visualized in Figure 4.12.

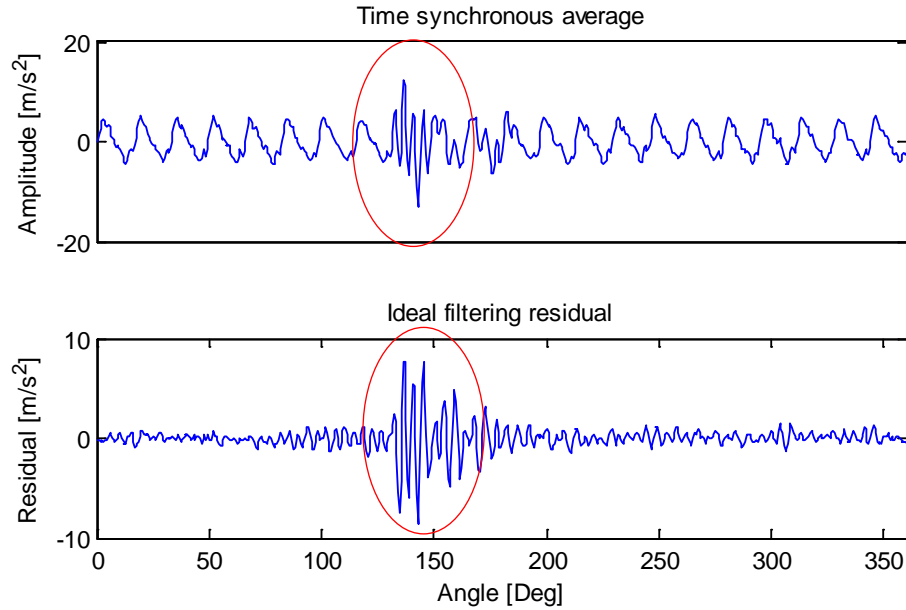


Figure 4.11: Time synchronous average and ideal filtering residual (HSS pinion)

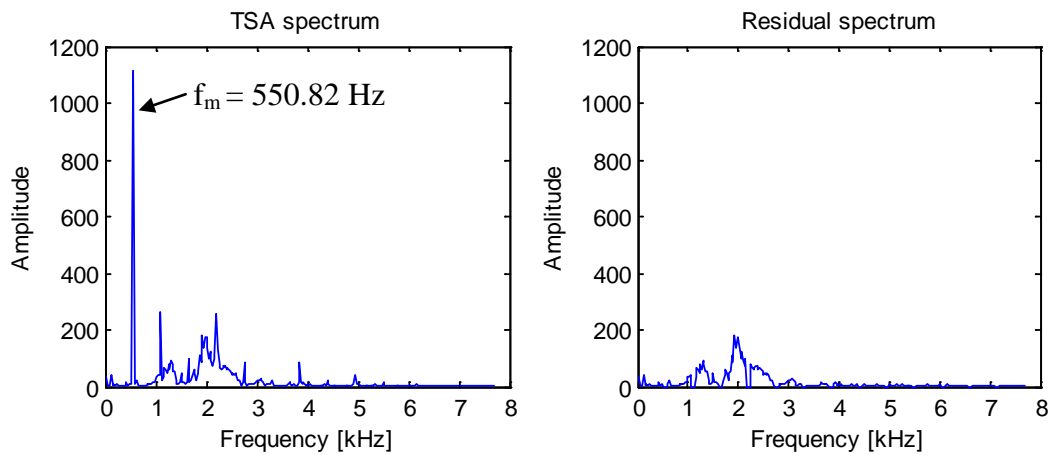


Figure 4.12: Spectrum of the TSA signal before and after ideal filtering (HSS pinion)

After ideal filtering no dominant discrete frequency contents exist, leading to the good results obtained.

However, Randall [31] states that filtering only the meshing components often resulted in quite large modulation effects at the first and second harmonics of the rotational speed, which were not related to local faults and so one or two sidebands around each toothmesh harmonic may also be removed to obtain the residual signal. These sidebands are not pronounced in the spectra, which is why their removal is not necessary here.

Investigations showed, that in some cases dominant frequency components may be left in the spectrum after editing. In those cases the peak detector, presented in publication [P5] can be used to blindly identify those components for zeroing.

Figure 4.13 shows the TSA signal and the residual after filtering of the meshing frequency and its harmonics for the pinion of WTG 18 of power plant 2.

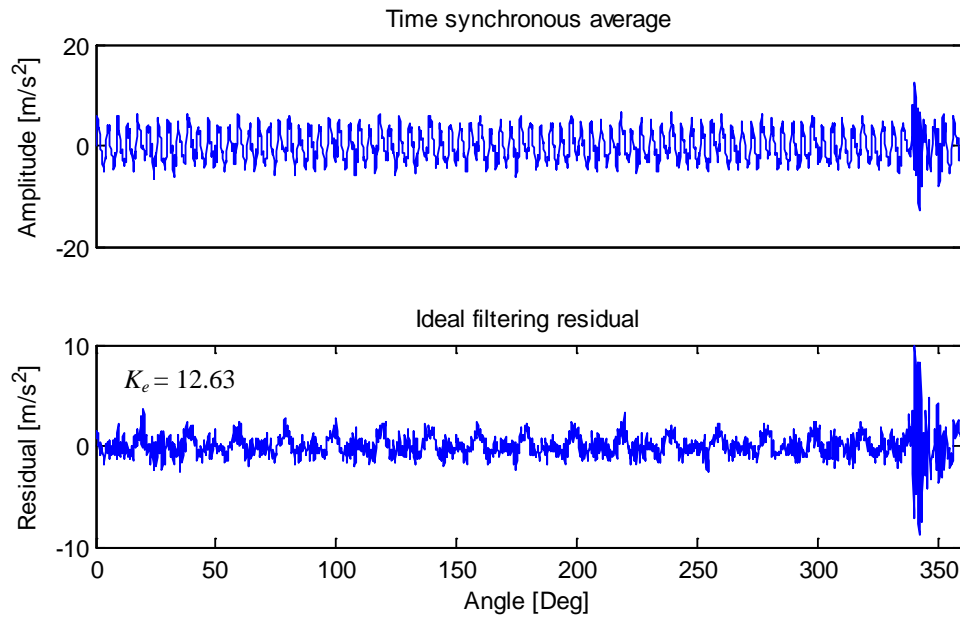


Figure 4.13: Time synchronous average and ideal filtering residual (IMS gear)

Also on the IMS gear, the false brinelling is already visible in the TSA, but gains prominence in the ideal filtered signal. The residual still contains some variations, which appear being regular. Figure 4.14 shows the spectrum of the TSA signal and the residual.

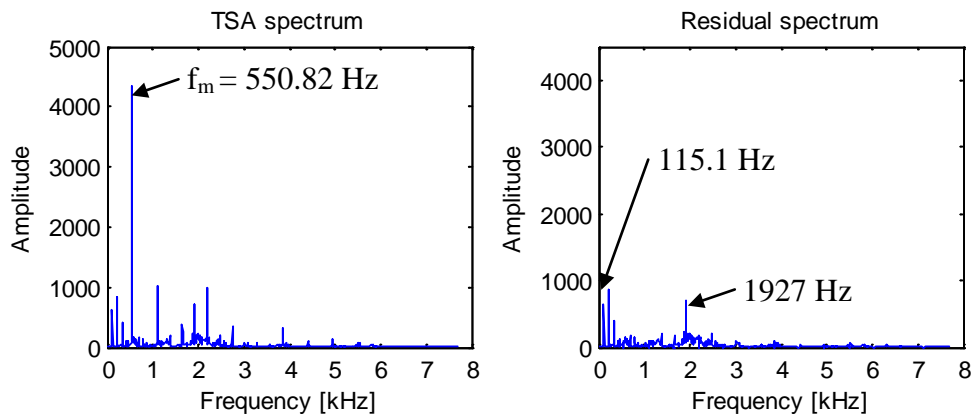


Figure 4.14: Spectrum of the TSA signal before and after ideal filtering (IMS gear)

After removal of the meshing components, there are still some distinct frequency peaks visible in the spectrum. They are the meshing frequency of the IMS/LSS components (115.1 Hz) and its harmonics and a sub harmonic of the IMS/HSS meshing frequency ($2202 \text{ Hz} - 550 \text{ Hz} / 2 = 1927 \text{ Hz}$). In general, local defects manifest themselves as deviations from the regular pattern and hence are modeled by the noise part in the spectrum. In turn, this means that all distinct frequency components can be removed. Applying the peak detector (window max method) with a bandwidth of ten values identifies the distinct frequency components as shown in Figure 4.15.

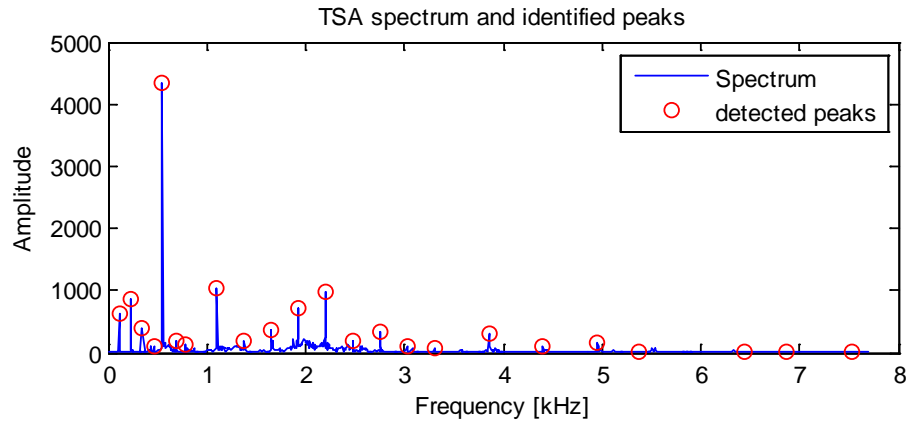


Figure 4.15: Spectrum of the TSA signal before ideal filtering (IMS gear) and detected distinct frequency components.

Ideal filtering the TSA signal by removing the detected distinct frequency components leads to the residual depicted in Figure 4.16. Note that independent of the detected peaks, also the meshing frequency and its harmonics are eliminated from the spectrum. This is a more general approach than the suggested removal of one or two sideband pairs by Randall [31]

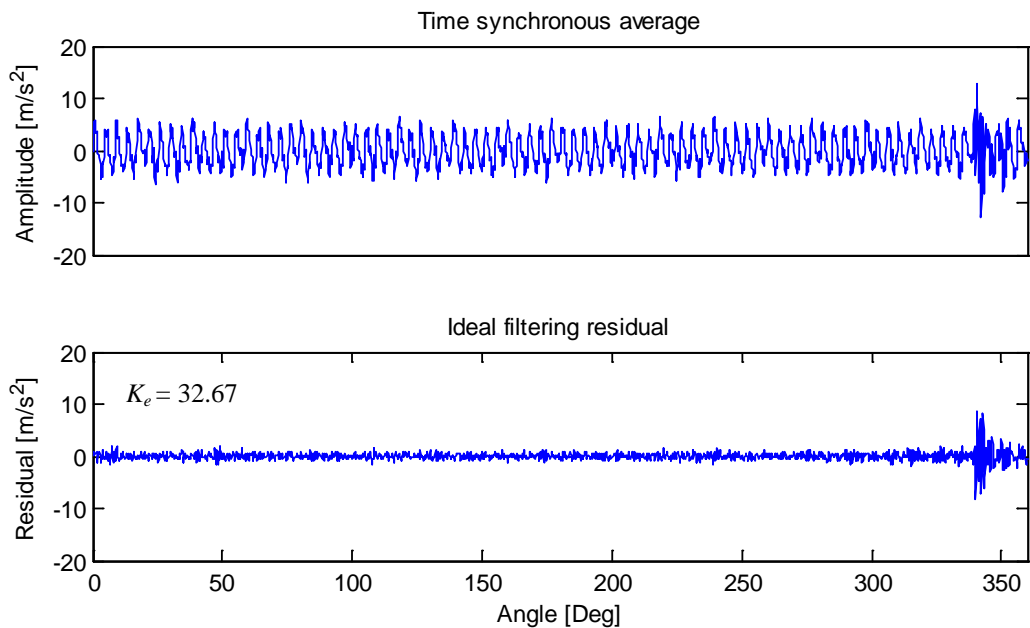


Figure 4.16: Time synchronous average and ideal filtering residual (IMS gear) (filtering distinct frequency components)

The removal of the distinct frequency components leads to noise level reduction and no obvious remaining regular pattern. The fault visibility is improved, which also expresses in the kurtosis value, which increased to 32.67 (from 12.63).

The diagnostic parameter for monitoring derived from this is the residual kurtosis evolution for each gear component.

4.2.3.2. Linear prediction

Linear Prediction (LP) is basically a way of obtaining a model of the deterministic (i.e. predictable) part of a signal, based on a certain number of samples in the intermediate past, and then using this model to predict the next value in the series [56]. After fitting a linear model to the data, the residual can be obtained by subtraction of the prediction from the measurement. For modeling, an AutoRegressive (AR) model can be used as introduced in publication [P5], which takes the form:

$$\hat{x}(j) = \sum_{k=1}^p a(k)x_m(k-j) \quad (7) [31]$$

$\hat{x}(j)$ is the predicted current value, p is the number of previous values (model order), $a(k)$ is the weighting coefficient and $x_m(k-j)$ is the previous value. The weighting coefficients can be obtained using the Yule-Walker equations model and Akaike's entropy-based Information Criterion (AIC) [34] to estimate the model order p .

Figure 4.17 shows the resulting residuals for the HSS pinion and the IMS gear of WTG 18.

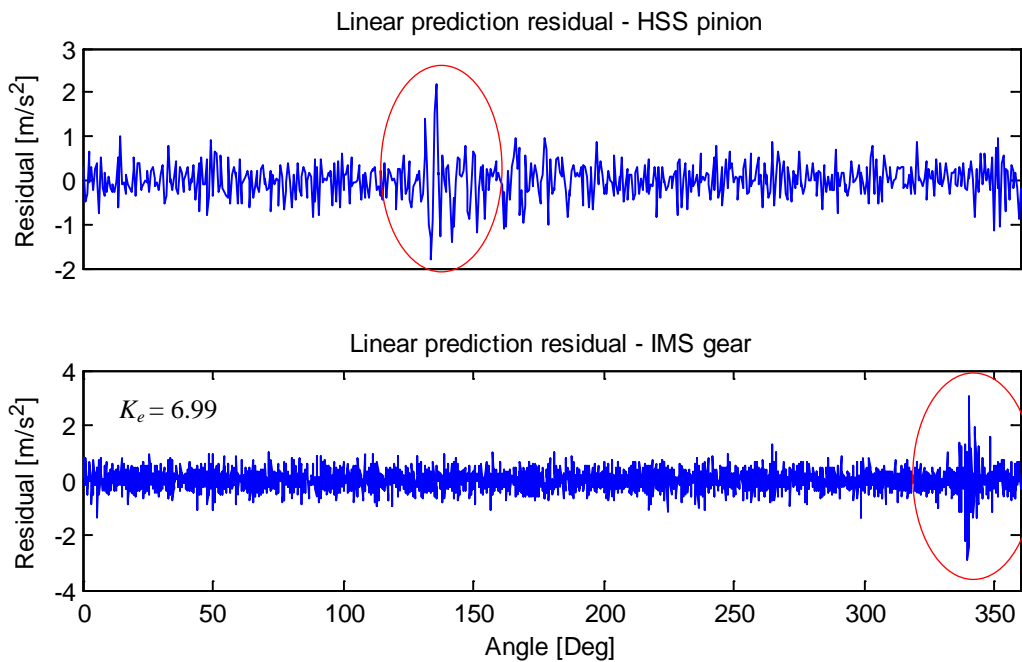


Figure 4.17: Time synchronous average and linear prediction residual (HSS pinion)

The performance of the LP model is not as good as the performance of the ideal filter. The noise part of the residual has larger variance, which also expresses in the residual kurtosis values (see Table 4.3).

Table 4.3: Kurtosis value comparison between ideal filtering and linear prediction

	Ideal filtering	Linear prediction
HSS pinion	16.60	5.18
IMS gear	32.67	6.99

The power plant average kurtosis for the gear and the pinion is around three, using ideal filtering. In comparison: the kurtosis observed at WTG 18 this is five to ten times this value.

It has been realized that the algorithm performance is dependent on the shaft speed of the gear of interest. There is a tendency that with slow rotating shafts (e.g. the LSS pinion), the performance of linear prediction is better than with ideal filtering. This is caused by the low number of revolutions in the given sample size for slow rotating shafts, leaving many distinct components in the frequency spectra of the TSA, which cannot be removed without loss of desired information. Furthermore, the assumption of true periodicity does only hold for a large number of averages.

More general information about LP can be found in [31].

4.2.3.3. Amplitude and phase demodulation

The examples given indicate that local faults may be already visible in the raw TSA signal in the form of a once per revolution modulation at the faulty tooth. Furthermore, it was shown that fault visibility is enhanced using ideal filtering of the meshing harmonics, sidebands and other distinct peaks from the spectrum or by applying LP techniques to obtain a residual. The techniques are known as broadband techniques, since they operate on the full frequency band. In 1986, McFadden [53] proposed a narrow band analysis technique, based on amplitude and phase demodulation. The TSA signal is bandpass filtered about one of the larger meshing harmonics. This meshing harmonic is also filtered and the remaining signal enveloped using the Hilbert transform. By this procedure, a signal is obtained which can be related directly to the amplitude and phase modulation present in the original signal average [53]. The results achieved using this technique implied that the phase modulation may allow for earlier warning of a local defect such as a fatigue crack in a gear than does the amplitude modulation [53]. This result has been confirmed by Dalpiaz [35]. However, in their research it was found that the technique is sensitive to the transducer location and the selection of a proper frequency band.

Given the measurement setup in wind turbines, the transducer location is usually fixed, and cannot be optimized to achieve better results. Moreover, the choice of a proper frequency band has to be automatized due to the large amount of data and components to be monitored. McFadden [53] suggests band pass filtering the TSA signal around one of the larger meshing frequency harmonics to obtain reasonable results, with a bandwidth of $2mf_s$. Where m is the number of sideband pairs and f_s is the shaft rotation frequency.

In the practical application of the wind turbine data available, selection of a good frequency band is a challenge, as it was found that the frequency band heavily influences the fault visibility. Actually, the phase modulation does not indicate any signs of fault, if a bad frequency band is chosen.

Using a data set of WTG 18 a comparison is made of the effectiveness of amplitude and phase modulation for different filter bands.

Figure 4.18 depicts the amplitude and phase modulation for a frequency band around the second meshing harmonic.

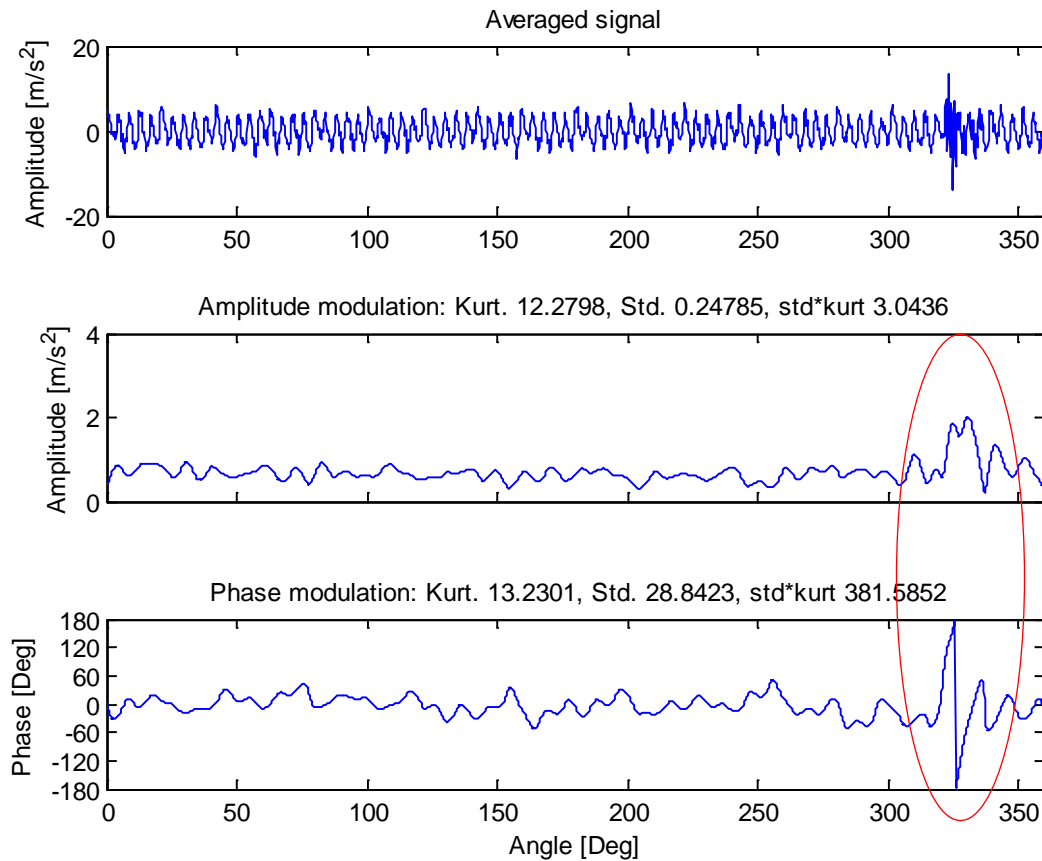


Figure 4.18: Amplitude and Phase modulation (band around: $2 \times f_{\text{mesh}}$)

It is visible that the amplitude shows increased values around the defect and the corresponding phase shows a lead. Figure 4.19 shows the same analysis, but for a frequency band around the meshing harmonic (dominant component in the spectrum).

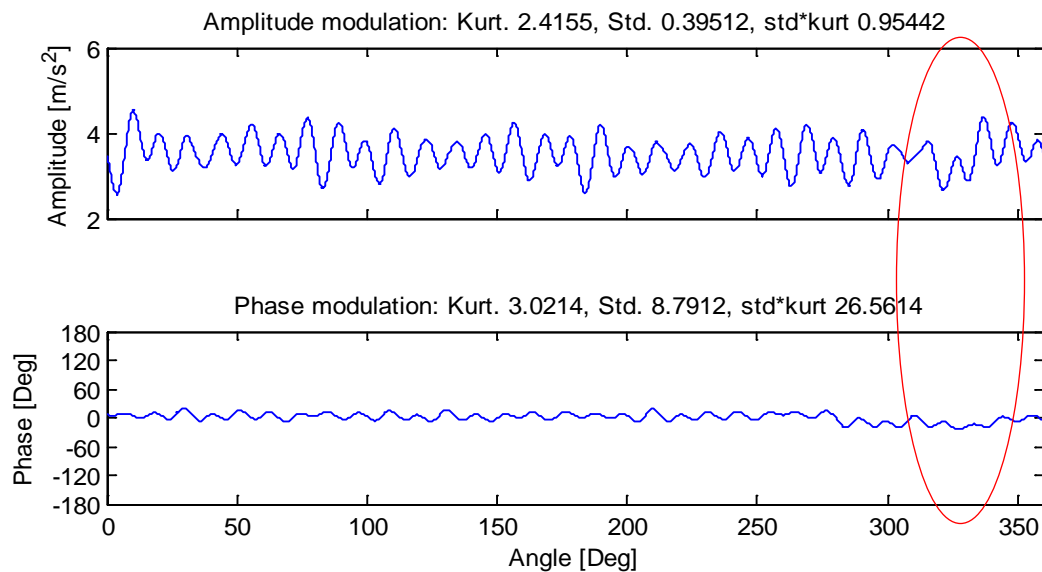


Figure 4.19: Amplitude and Phase modulation (band around: $1 \times f_{\text{mesh}}$)

In both modulations (amplitude and phase), the potential fault is practically invisible, giving no fault indication. The kurtosis values are with 2.42 and 3.02 low in comparison to the values obtained around the second meshing harmonic (12.28 and 13.23). Filtering the TSA

signal around the fourth meshing harmonic (2202 Hz) gives very good fault visibility in the amplitude modulation ($K_e = 45.50$), and many phase jumps in the phase modulation with a kurtosis value of 5.39 (see Figure 4.20).

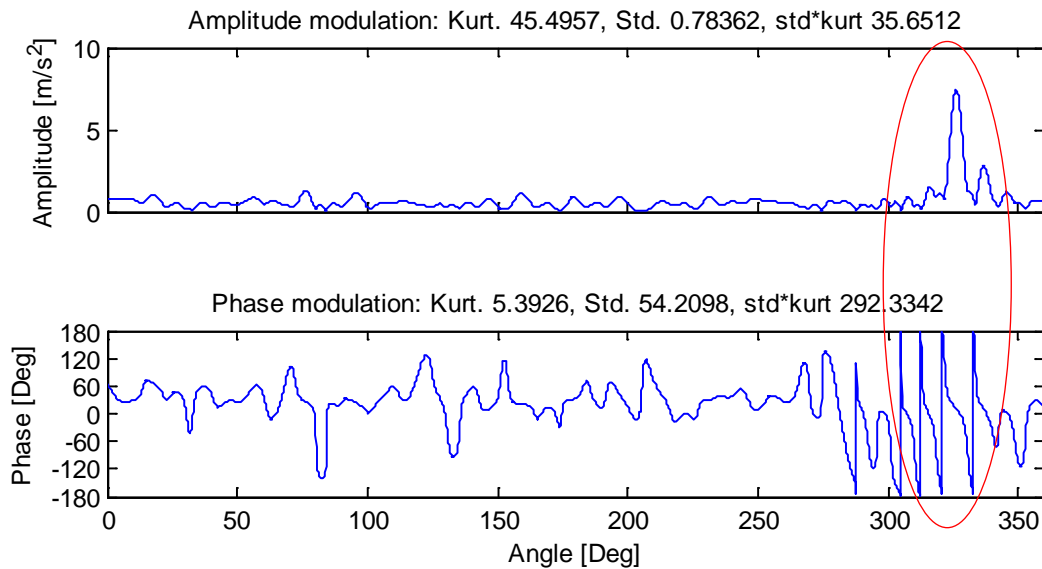


Figure 4.20: Amplitude and phase modulation (band around: $4 \times f_{\text{mesh}}$)

This brief comparison shows that for each filter band different results are obtained, that lead to different fault visibilities and patterns, confirming the research results of Dalpiaz [35].

In a practical application, the optimum frequency band cannot be known before hand, which makes this procedure not well suited for automated analysis.

4.2.3.4. TSA conclusions

The TSA appears as a useful tool for false brinelling defect detection. If a position signal is available, TSA allows exact localization of the fault. Deriving the TSA residual, by means of linear prediction or ideal filtering as proposed in literature increases the fault sensitivity. The example showed best performance of the ideal filtering approach to receive the TSA residual for fast to intermediate rotating shafts, while at slow rotating shafts linear prediction was superior. The diagnostic parameter identified is the TSA residual kurtosis, which can be monitored in terms of both the absolute level (in comparison to the power plant average) and relative level. Figure 4.21 illustrates the residual kurtosis evolution of the HSS pinion and the IMS gear for WTG 18.

The thresholds are computed analogously to the procedure described in section 4.2.1. By knowing the power plant average kurtosis level of the components, which is around three, all measurements show signs of abnormality, but no trend is present indicating fault progression.

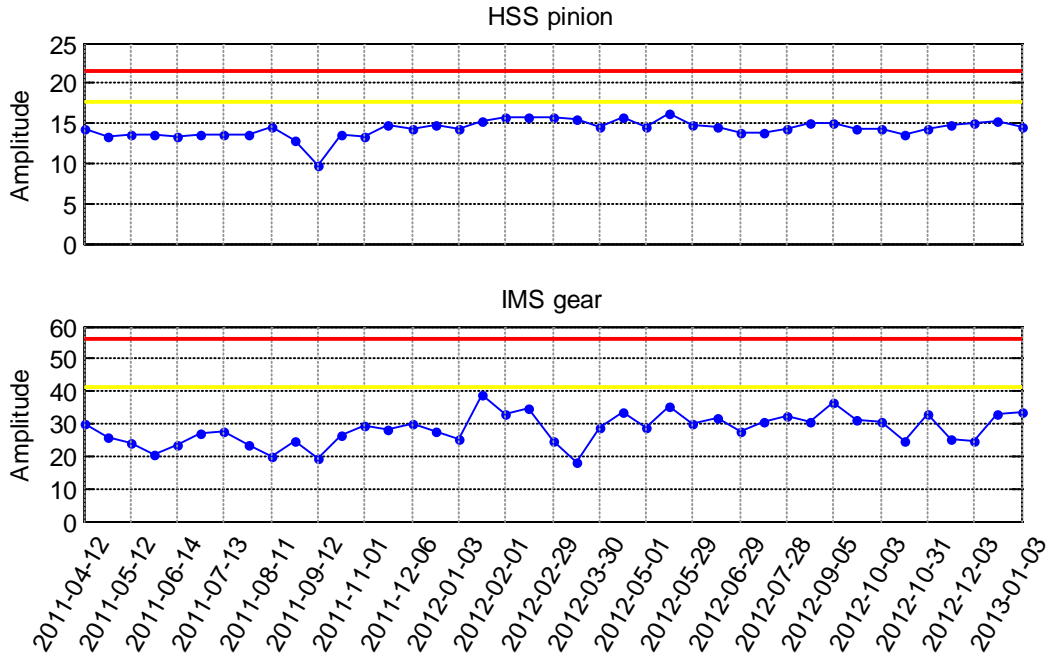


Figure 4.21: Residual kurtosis evolution (ideal filtering) of the HSS pinion and IMS gear for WTG 18 of power plant 2

4.2.4. Wavelet transform

A more recent technique to detect local gear defects is wavelet analysis and the application of the WT. It was first applied to gearbox spur gear defect diagnosis by Staszewski [45] in the early 1990's. The signal is decomposed in terms of a family of wavelets, which have a fixed shape, but can be shifted and dilated in time [31]. The WT reads.

$$W(a; b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (8)$$

Where $\psi(t)$ is the mother wavelet, b is the time, a the scale and $x(t)$ is the time variant signal. The WT provides a good means of studying how the frequency content changes with time and consequentially is able to detect and localize short-duration phenomena [35]. There is a broad variety of mother wavelet functions available. They can be real valued or complex. The choice of the wavelet function influences the analysis result. It means that different wavelets produce different separation results [46]. To obtain the best fault separation, a careful selection of wavelets needs to be carried out [46]. Dalpiaz [35] achieved good performance, by applying the Morlet wavelet. Elbarghathi [46] suggested based on their study selecting a wavelet that produces a higher RMS value of wavelet coefficients when it is applied to baseline data. Here three mother wavelets were compared based on this suggestion. The wavelets compared are:

- Morlet [57]
- Derivative of Gaussians (DOG) [57]
- Paul [57]

Using data from a fault free turbine, the Morlet wavelet gave the highest RMS of wavelet coefficients (Morlet 22.1; DOG: 13.7; Paul 13.5). The Morlet wavelet is defined as [31]:

$$\psi(t) = \frac{\sigma}{\sqrt{\pi}} e^{i2\pi f_0 t} e^{-\sigma^2 t^2} \quad (9)$$

With σ and f_0 being the Gaussian shape parameters and t the time instant. The shape of the Morlet wavelet is shown in Figure 4.22.

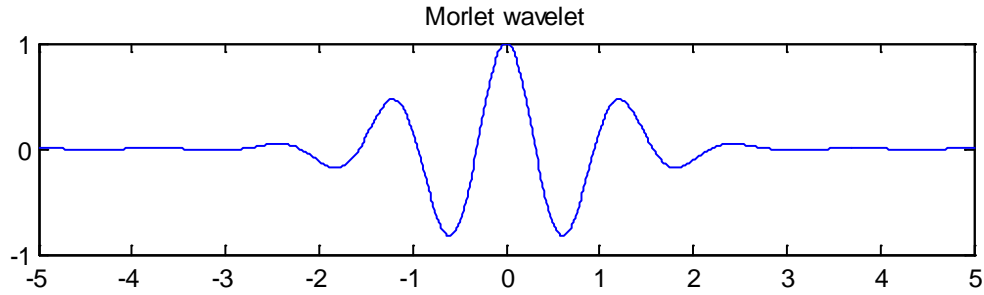


Figure 4.22: Morlet wavelet

Generally, the Morlet wavelet is known to have a good frequency resolution, due to its comparably large number of oscillations. Wavelets with fewer oscillations, like the Paul wavelet, give a good time resolution but compromising the frequency resolution. The Morlet wavelet $\psi((t - b)/a)$ family has the property that with increasing scale a , the frequency of the wavelet decreases, but at the same time the period length increases, keeping the number of oscillations constant. Thus, in the frequency domain, the WT has a good resolution for low frequencies and in time domain good resolution for high frequencies; the latter property makes the WT suitable for detection of transient signals [35]. A further property of the Morlet wavelet is that its spectrum is Gaussian. More information about the wavelet properties can be found in reference [57].

The WT can be achieved by multiplication in the frequency domain of the signal spectrum with that of the Morlet wavelet and then inverse Fourier transformation [31]. The result of the wavelet transformation can be presented in time-scale plot, i.e. a scalogram. The scalogram for WTG 18 is visualized in Figure 4.23.

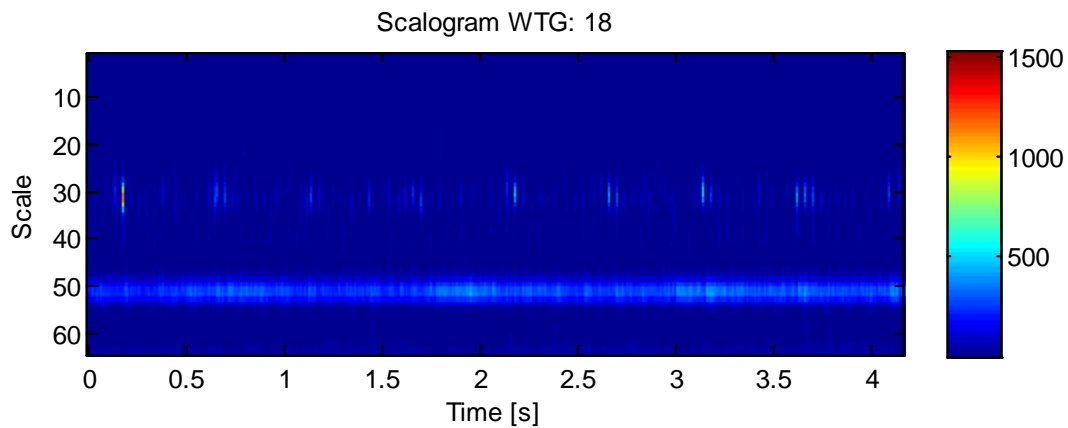


Figure 4.23: Scalogram of WTG 18

The scalogram shows a repeating pattern about a scale of 30. The colored bar at a scale around 53 is the regular meshing pattern. However, from the raw TS, it is not directly possible to identify the faulty component, e.g. HSS pinion or IMS gear. Dalpiaz [35] found the WT on the raw data to be practically unaffected by their studied small crack. Hence, the WT is applied to the TSA signal. In Figure 4.24 the raw signal is averaged synchronous to the HSS pinion and in Figure 4.25 synchronous to the IMS gear.

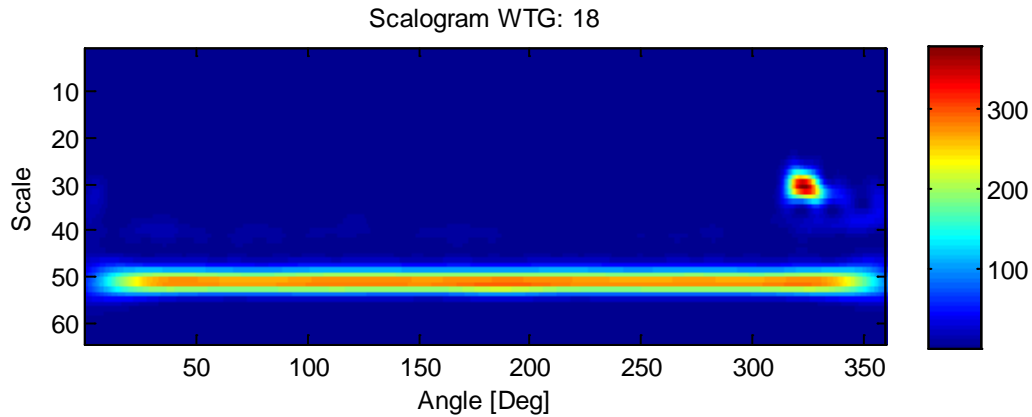


Figure 4.24: Scalogram for the TSA signal of WTG 18, HSS pinion

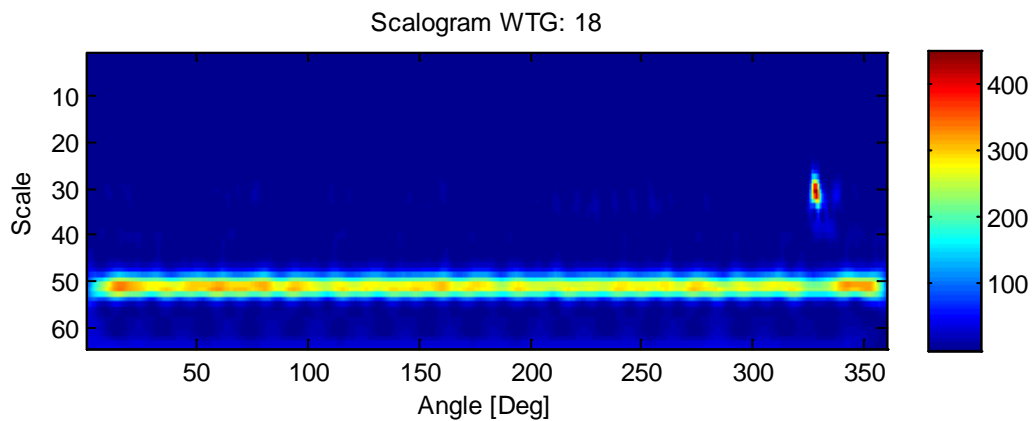


Figure 4.25: Scalogram for the TSA signal of WTG 18, IMS gear

Both scalograms show that the pattern around a scale of 30 is synchronous with the individual shaft rotations. Clear “hot spots” are visible around 325°. The dominating meshing pattern can be eliminated by applying linear prediction or ideal filtering to the TSA signal (see Figure 4.26), further increasing the fault visibility.

The appearance of faults in the scalogram can be quite diverse, depending on how well the mother wavelet can represent the fault pattern. For the case surveyed, the fault expresses itself as clear spot, which is easy to identify visually. Dalpiaz [35] suggest calculating the mean across the scales (along each angular position) for feature extraction. This suggestion arose from their observation that the surveyed fault established as a vertical line across several scales. Applying this procedure here would smear the information contained in the scalogram. In the need of deriving a diagnostic parameter from the wavelet analysis, for this particular fault type the maximum kurtosis value is chosen. For this purpose, the kurtosis along each scale is calculated and the maximum of the values is taken. It should be noted, that the

scalogram can be very useful, for an expert to identify the fault type, but reliable and automated feature extraction is still an open field of research.

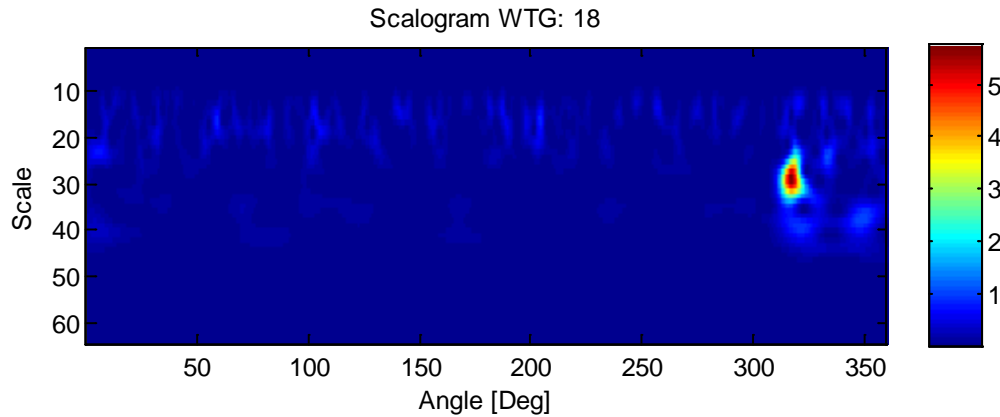


Figure 4.26: Scalogram for the TSA signal of WTG 18 filtered by LP, HSS pinion

The wavelet coefficients of the TSA residual signal must be calculated for each wheel individually. By doing so, it can be noticed that the general fleet kurtosis level of different components can be quite different. This is caused by the different characteristics of the TSA for each component, e.g. different number of blocks to be averaged which influences the quality of the TSA, or different meshing error contributions for each wheel. Therefore, there is no single global kurtosis value suitable as a threshold above which a local fault is present.

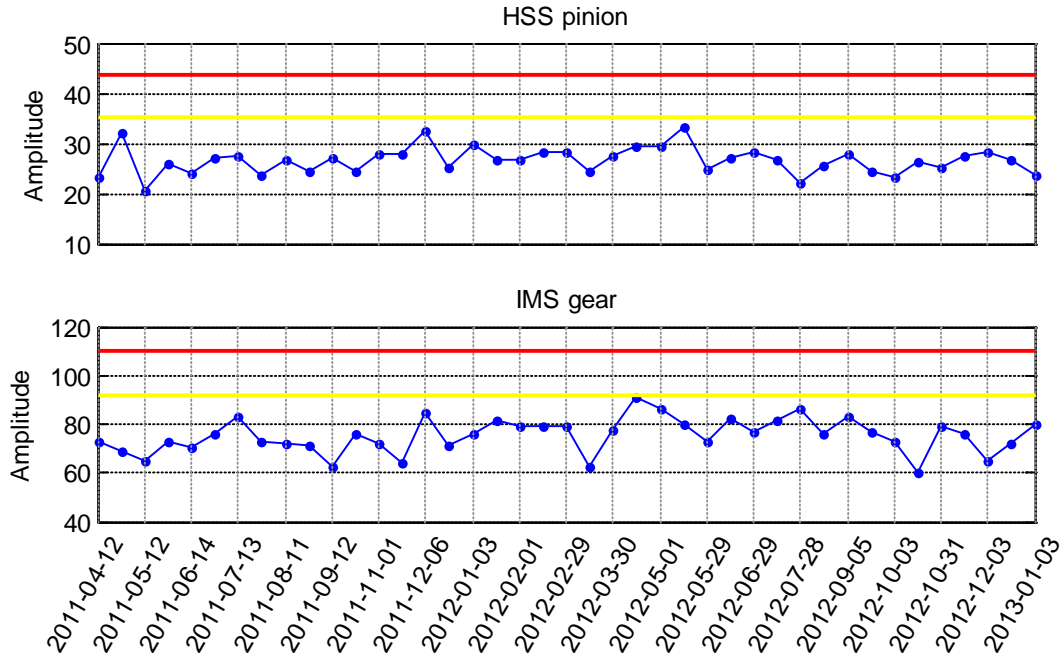


Figure 4.27: Maximum wavelet coefficient kurtosis evolution of the HSS pinion and IMS gear for WTG 18 of power plant 2

As with TSA the residual kurtosis, also the maximum wavelet coefficient kurtosis evolution shows no trend or decreasing component condition. What is remarkable is the kurtosis level. The power plant average kurtosis level of inconspicuous turbines is around ten for the HSS

pinion and the IMS gear and thus the kurtosis values of WTG are unusually high, indicating the local discontinuity at the gear wheels (false brinelling).

In general, the normal meshing error (e.g. caused by the chipping wheel during manufacturing) makes it necessary to treat each turbine individually with its characteristics. A comparison to the fleet average is still useful to identify turbines like WTG 18, for which very high kurtosis values are achieved since commissioning.

4.2.5. Transmission error

The transmission error represents the difference between the angular motion of a driven gear and that which it would have if the transmission were perfectly conjugate, that is constant speed out for constant speed in [31]. For determination of the transmission error, accurate shaft encoders are required. One at the gear and one at the pinion shaft. After scaling the measured speed (or angular motion) of the gear or pinion shaft by the gear ratio, the difference in angular motion can be determined. This method seems interesting as with accurate encoders, the deflection of each tooth (expressing itself as difference in angular motion) can be determined. However, for the available data sets the encoder information is not available. In any case, if available, it would only exist as gearbox input and output speed and thus requires some extra processing to extract information for each gear pair.

4.2.6. Sideband energy ratio

The Sideband Energy Ratio (SER) is a patented algorithm evaluating the energy contained in the sidebands, normalized by the energy of the center frequency (meshing frequency).

$$SER = \frac{\sum_{i=1}^6 Sideband Amplitude_i}{Center mesh frequency amplitude} \quad (10) [48]$$

It is based on the expectation, that in case of defect the meshing frequency is amplitude modulated and consequentially the sidebands corresponding to the shaft rotational speed increase. Hanna et.al [48] show in their technical note the efficiency of this algorithm, using two examples of broken HSS pinion teeth. However, the defect was quite severe, with large parts of the broken tooth already missing. The authors state that for healthy state gears the SER is less than one. In the two test cases, SER values of 5.6 and 6.6 were achieved.

The procedure is sensitive to speed fluctuations, which is why order tracking or tracking is required to assure that the mesh frequencies and related sidebands are at fixed orders or frequencies.

In 1989, Szczepanik [58] presented the sideband index for first estimation of the pinion quality. The sideband index is defined as:

$$M = \frac{1}{N} \sum_{i=1}^N S_{\max i} \quad (11)$$

where M is the sideband index, N is the number of sidebands and $S_{\max i}$ is the i th maximum linear amplitude of the sideband [58].

This algorithm is very similar to the SER presented by Hanna et.al [48]. Differences exist in normalization and sideband processing. While Hanna et.al [48] sum all sideband amplitudes on either side of the center frequency, Szczepanik sums the maximum amplitudes of each sideband pairs.

As a consequence, the sideband index is the average sideband amplitude, whereas the SER is a nondimensional ratio. The latter allows definition of fixed thresholds.

Applying this method to the wind turbine data at hand shows that SER values are dependent on the center mesh frequency. Figure 4.28 shows a comparison of the SERs for the fundamental meshing frequency and the 3rd harmonic for an inconspicuous WTG.

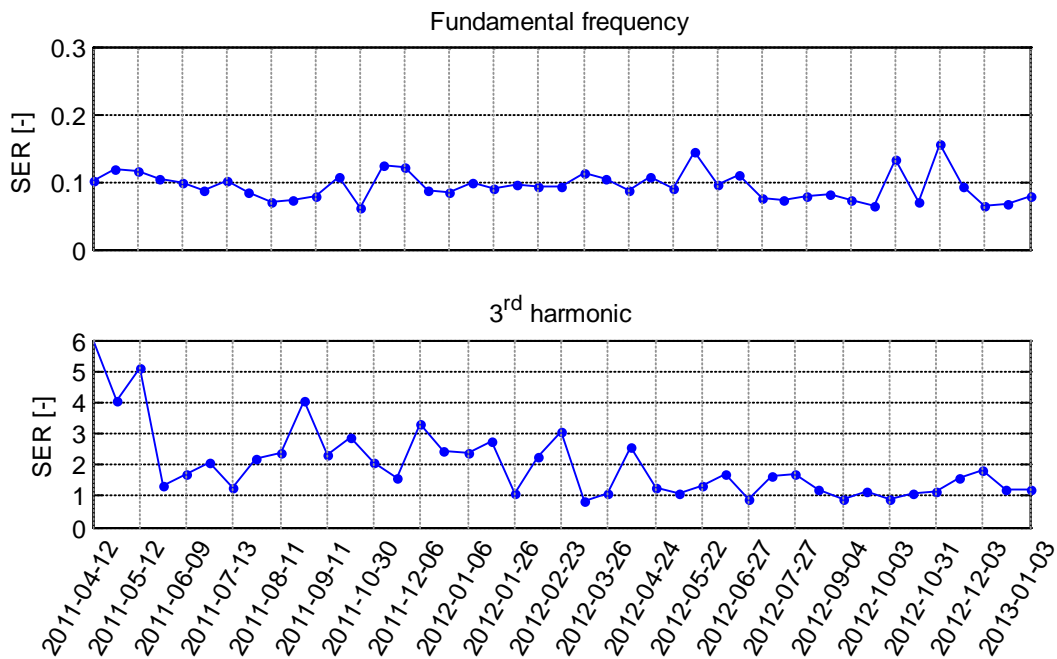


Figure 4.28: SER evolution of WTG 07 (HSS pinion); top: SER around the fundamental frequency, bottom: SER around 3rd harmonic.

The fundamental meshing frequency is the dominating meshing component and its amplitude is high. Consequentially the SER values obtained are low. Higher harmonics are lower in amplitude, but the sidebands are comparably high, leading to larger SERs. Especially for larger harmonics, the SER values are volatile.

The figure shows different SER levels (for $i = 6$) dependent on the meshing harmonic selected. Hence, a universal claim as made by Hanna et al [48] “The SER value is typically less than one for a healthy gear mesh” cannot be confirmed for all meshing harmonics. Figure 4.29 shows a similar comparison for WTG 18 with the existing false brinelling.

The SER level around the basic meshing frequency mesh frequency is almost the same for both turbines. However, the level for the 3rd harmonic of the meshing frequency differs. At WTG 18 the level is about 15 times higher, but fluctuates heavily throughout different measurements, hence it cannot be used as fault severity indicator. However, the overall level in comparison to other turbines may indicate a gear problem. Due to the large variance, trending is difficult. Since it is unknown prior defect around which meshing harmonic the largest sidebands occur, the algorithm proposed by Hanna et.al [48] is modified with a further summation over four meshing harmonics, to obtain only one SER value.

$$SER = \sum_{j=1}^4 \frac{\sum_{i=1}^6 Sideband\ Amplitude_i}{Center\ mesh\ frequency\ amplitude_j} \quad (12)$$

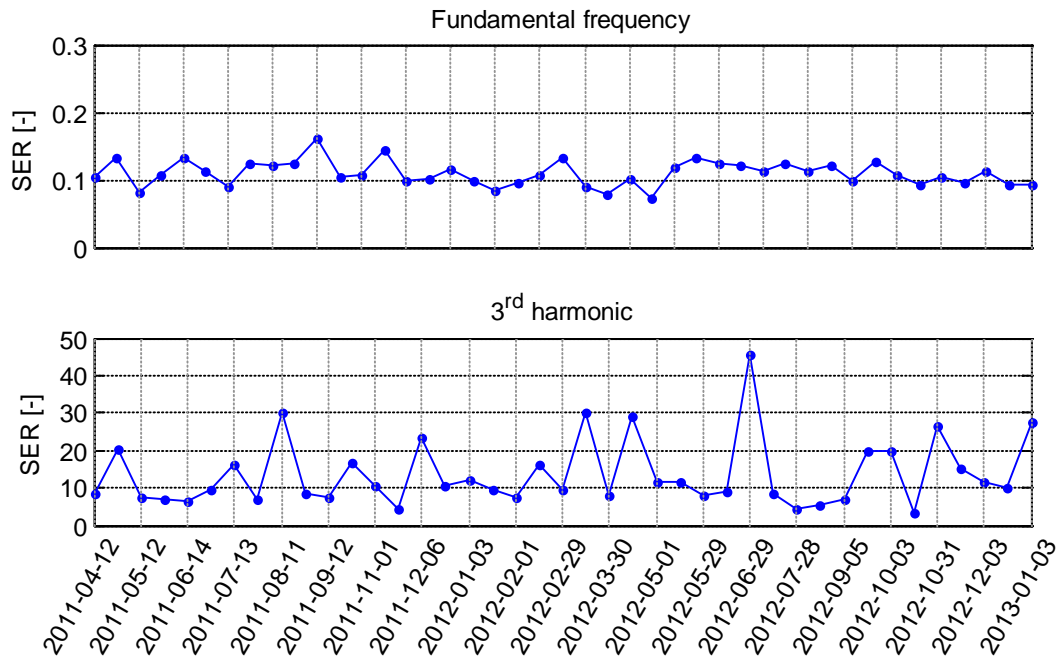


Figure 4.29: SER evolution of WTG 18 (HSS pinion); top: SER around the fundamental frequency, bottom: SER around the 3rd harmonic.

The modified SER evolution for the HSS pinion and the IMS gear of WTG 18 is presented in Figure 4.30.

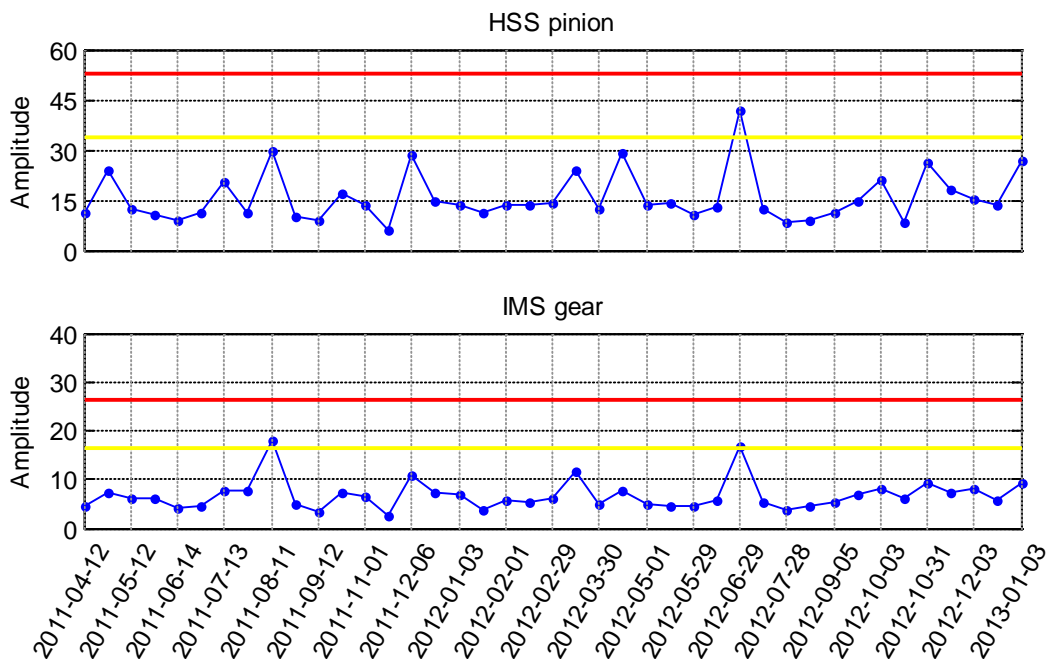


Figure 4.30: SER evolution of the HSS pinion and IMS gear for WTG 18 of power plant 2

The modified SER evolution still shows volatile behavior, with some outliers. The power plant average kurtosis level of inconspicuous turbines is around five for the HSS pinion and the IMS gear, which gives indication of anomaly here at the HSS pinion, but not at the IMS gear.

The SER must be calculated for each gear and pinion separately. Dependent on the sideband spacing and the mesh frequency considered, the defective component can be located.

The collection of the energy contained in the sidebands and the meshing frequencies can be efficiently performed via the cepstrum as described in section 4.2.1. However, the cepstrum bears the disadvantage, that the defective component cannot be uniquely determined. In complex gears (more than one stage), the sideband spacing (shaft rotational speed) only locates the shaft at which the defect is present. The information around which meshing frequency these spacings occur is lost by the transformation to quefrequencies.

4.2.7. Discussion

All surveyed algorithms were able to give some evidence about the false brinelling issue. Due to missing base line measurements, trend detection in the parameter evolutions did not indicate any anomaly. Instead, a comparison to power plant average values gave evidence. The variance of the monitored parameter is an important quantity, since it influences the trust an expert has in the parameter if an alarm is raised. With this metric, the SER performed worst, since its values were found to heavily fluctuate making a reliable condition classification difficult. The TSA residual kurtosis and the Wavelet maximum kurtosis appear good fault indicators, with high deviations from the power plant average values. The cepstrum and the SER give only a small deviation from power plant average values in case of the IMS gear.

It is important to note that these results require further validation and are by no means of global nature. Considering the conclusions drawn by Dalpiaz et.al [35] and Siegel et.al [38], monitoring of the TSA residual kurtosis, the wavelet maximum kurtosis and the cepstrum quefrequency evolution is monitored in the CMS, to cover a broad variety of possible failure modes in future.

5 Global fuzzy expert system

Most of the fuzzy expert system has already been introduced in publication [P3], [P4] and [P5]. So far the different parameters to be monitored have been analyzed in isolation. The global fuzzy expert system described in this section is characterized by combining the fuzzy expert systems of publication [P3] & [P5]. Hence, it allows the combination of the analysis results from SCADA and vibration data.

The experience gained with fuzzy logic throughout this research has shown that fuzzy logic is a powerful tool for automating analysis and supporting decisions. It is able to mimic human intuition and reasoning well, provided the right set of rules is available. Especially the ability to output results in cases where rules only partially apply proves useful in real world problems. Unforeseen events and external influences can cause the actual fault symptoms to be manifold. Classical logic (crisp logic), is likely to fail in those cases, since it can not cope with the real worlds imprecision.

A further useful property of fuzzy logic is that the output can be fully backtracked to the rule triggered and the individual input contribution. This ability puts an expert in the position to adapt rules in case a undesired result is obtained. Fuzzy expert systems have rarely been applied to wind turbine condition monitoring in the past. In 1997 Caselitz, Giebhardt and Mevenkamp [12] described a CMS development combining different sensor information with a fuzzy classifier. However, no details about the involved fuzzy classifier are published. Furthermore their development requires a specific sensor setup and additional hardware. In fact Sanz-Bobi et.al [17] were the first to describe the application of fuzzy expert systems for wind turbine SCADA data analysis automation in 2006.

In publication [P3] & [P4] a fuzzy expert system is presented that is similar to the approach taken by Sanz-Bobi [17] allowing generic and specific rule implementation to analyse normal behaviour modeled wind turbine SCADA data. Their research was applied to the health condition of a wind turbine gearbox. The main original contribution of publication [P3] & [P4] in this context is the prove of applicability of fuzzy expert systems to a broad variety of different wind turbine components and the introduction of additional Membership Functions (MFs). Furthermore a new set of inputs has been introduced carrying the same information, but with less MFs. This refers to the two types:

- Inputs with 5 MFs: very low, low, normal, high, very high
- Inputs with 3 MFs: low, normal, high

This proposed set up reduces the number of rules required describing signal anomalies and gives the expert more freedom in deriving rules. For more information please refer to publication [P3].

In publication [P5] a fuzzy logic based interpretation of envelope frequency contents for bearing diagnosis is presented. In this approach the benefits of fuzzification remain unused as the interpreted output of the frequency content identification is crisp, i.e. a frequency content is either present or not present. In 2008 Ebersbach and Peng [59] presented an expert system development for vibration analysis in machine condition monitoring using fuzzy logic. It consists of a frequency confidence factor using a triangular MF and an amplitude confidence factor using a linear MF. The frequency confidence MF is centered about the fault frequency and the amplitude confidence factor is applied for amplitudes between alarm and severe alarm thresholds as a linear relationship. The confidence factor is the sum of the frequency and amplitude confidence factors [59].

There are two problems occurring with this approach, which is why a different line is followed in this research. The first concerns the frequency confidence factor. In practical applications, the kinematic defect frequencies seldom match the theoretical values, since no slip is considered in theory. The slip is typically of the order of 1-2%, both as a deviation from the calculated and as a random variation around the mean frequency [31]. Thus, the frequency confidence factor will often be lower than the numerical value of one although the frequency content can clearly be related to a defect frequency. The second concerns the amplitude confidence factor. Due to the preprocessing procedure (signal PW) the scaling of the amplitudes is somewhat arbitrary [33]. This fact also makes definition of fixed envelope amplitude thresholds difficult, which are used as a severity indicator by Ebersbach and Peng [59].

Signal PW is a new method for discrete random separation. Although effective, it requires different approaches for fault pattern analysis. The development here consists of a frequency content identifier able to cope with the imprecision of the defect frequencies. The output is the information if a certain frequency content is dominant in the envelope spectra. In publication [P5] this output is analysed by a simple fuzzy logic procedure.

The fuzzy logic systems under research to analyse SCADA data and the frequency content are different in their approach and must be lifted to a level, not restricting the expert in the process of rule (and thus knowledge) implementation. Providing a simple but effective methodology to link all available information presents a novelty in the field of wind turbine condition monitoring. Although in many cases, the separation of the two fuzzy expert systems developed may be no harm, from publication [P1] it can be imagined that cases exist, where linking SCADA data with vibration measurements may be of benefit in terms of confidence in reasoning. Combining the information of two different information sources may lead to a higher confidence in diagnosis than treating each information source individually.

5.1. System overview and description

Core elements of the global fuzzy expert system are linguistic terms and MFs, able to describe the input pattern. Using the following linguistic terms allows for the input pattern characteristic to be well described:

- Present
- Not present
- Very low
- Low
- Normal
- High
- Very high

The corresponding MFs are described in publication [P3] & [P5]. Figure 5.1 emphasizes the global fuzzy expert system structure.

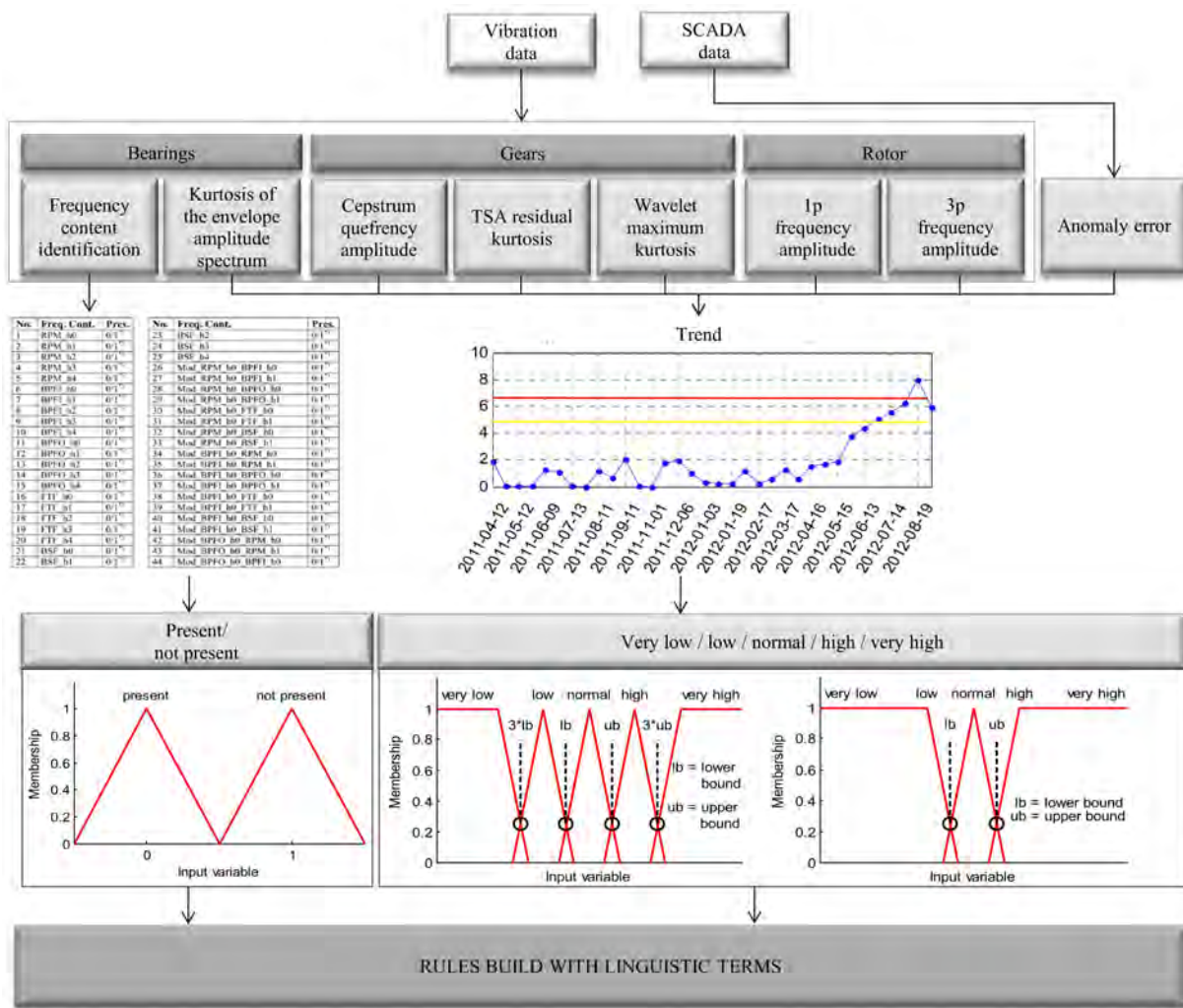


Figure 5.1: Global fuzzy expert system schematic

The proposed setup allows combination of all kinds of analysis results. With the procedure presented in publication [P5] the automated analysis is limited to the statement whether or not a certain frequency pattern is present. With the proposed structure in Figure 5.1, bearing fault severity can be quantified by use of the KEAS value, whose condition thresholds (yellow and red) are determined analogously to chapter 4.2.1. Each rule has the outputs:

- Condition

- Diagnosis
- Potential root cause

With the MFs defined in the same manner as the fuzzy expert system described in publication [P3]. Rules to diagnose bearing inner race defect can be for instance (for terminology see publication [P5]):

If KEAS **high** & BPFI_h0 **present** & BPFI_h1 **present** & BPFI_h2 **present** & Mod_RPM_h0_BPFI_h0 **present** **then** Condition **yellow** & Diagnosis **bearing inner race defect** & Potential root cause **axial crack**.

If KEAS **very high** & BPFI_h0 **present** & BPFI_h1 **present** & BPFI_h2 **present** & Mod_RPM_h0_BPFI_h0 **present** **then** Condition **red** & Diagnosis **bearing inner race defect** & Potential root cause **axial crack**.

Other example rules to characterize bearing defect incorporating the bearing temperature may be:

If KEAS **high** & Anomaly error (3 MF's) Gearbox bearing temp. HSS gen. **high** & BPFI_h0 **present** & BPFI_h1 **present** & BPFI_h2 **present** & Mod_RPM_h0_BPFI_h0 **present** **then** Condition **yellow** & Diagnosis **bearing inner race defect** & Potential root cause **axial crack**.

If KEAS **very high** & Anomaly error (3 MF's) Gearbox bearing temp. HSS gen. **high** **then** Condition **red** & Diagnosis **bearing inner race defect** & Potential root cause **axial crack**.

From the example rules it becomes evident, that the global fuzzy expert system gives the expert full freedom in implementing his knowledge, provided the required analysis and information source are available. Future developments of further analysis methods, e.g. to monitor the planetary stage of the gearbox or slow rotating bearings as well as oil quality analysis or oil particle counter information can be easily incorporated and their output implemented in terms of modifications of existing or by setting up new rules.

Each component in the turbine is assigned individual rules as long as condition monitoring signals exists. This incorporates the expertise that fault patterns can be different e.g. for medium and fast rotating bearings. Moreover, rules can be established at different levels (see Figure 5.2).

This procedure is advantageous, since some parameters monitored do not allow backtracking to the root. For instance, a cabinet temperature shows an anomaly, but the temperatures of the components in the cabinet behave normal. Another example is the monitoring of the KEAS value, which in itself does not allow identification of the root cause, without consideration of the envelope spectra. Those general parameters can be monitored by rules at the component level, reducing the required number of rules. This fact is further stressed by the application examples presented in the following.

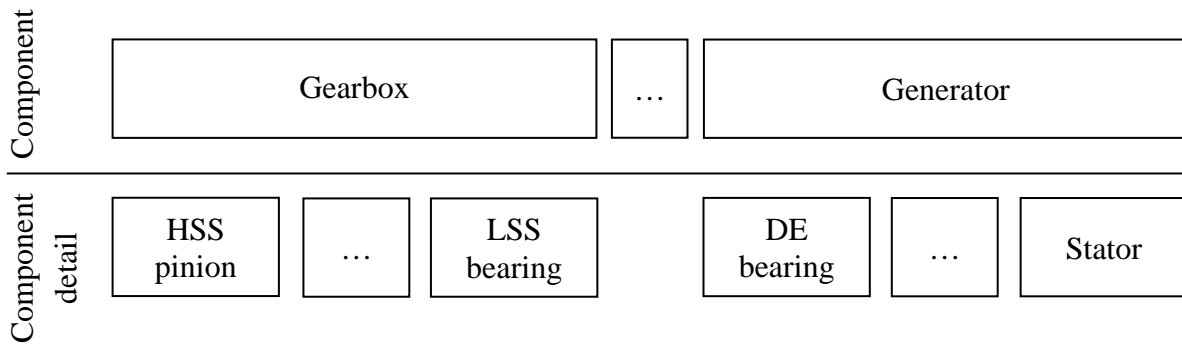


Figure 5.2: Levels for rule establishment

5.2. Application examples

In this section, three examples are given to emphasize the working principle of the global fuzzy expert module. Furthermore, it is briefly shown how the analysis result may be visualized to ease diagnosis. Due to the tremendous amount of data generated by each power plant on a daily basis, data streaming and visualization is important in order to maintain an overview.

The developed visualization is structured in a layer based manner. The four layers implemented here are:

- Fleet layer
- Power plant layer
- Turbine layer
- Component layer (if available)

At the fleet layer the worst condition of each WTG in a power plant is visualized allowing a quick navigation to the power plant of interest. In addition, the power plant average power output is visualized for information. The fleet layer visualization is shown in Figure 5.3

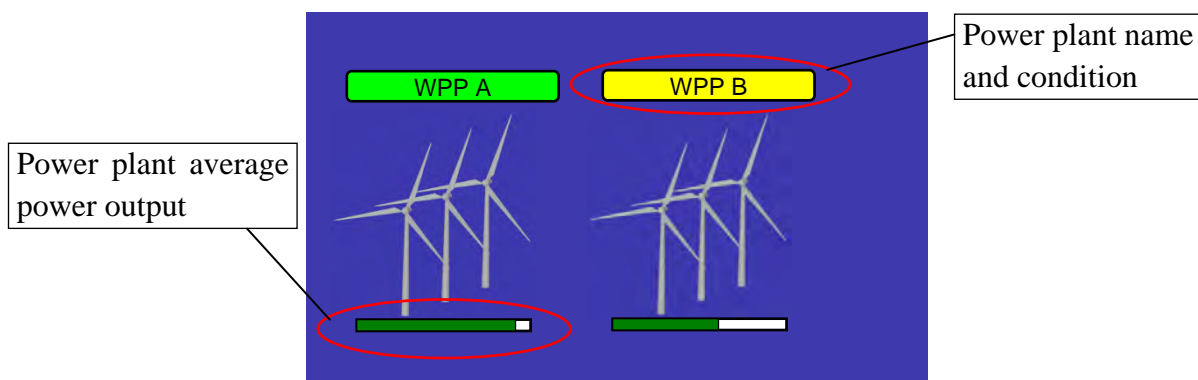


Figure 5.3: Condition visualization at the fleet level

The power plant layer displays the latest condition of the entire power plant turbines, together with the information about the average power output of each turbine at this day. The latter indicates under what load the turbine operated on average, while the anomaly or fault occurred. The power plant level for power plant 2 is visualized in Figure 5.4 exemplary.

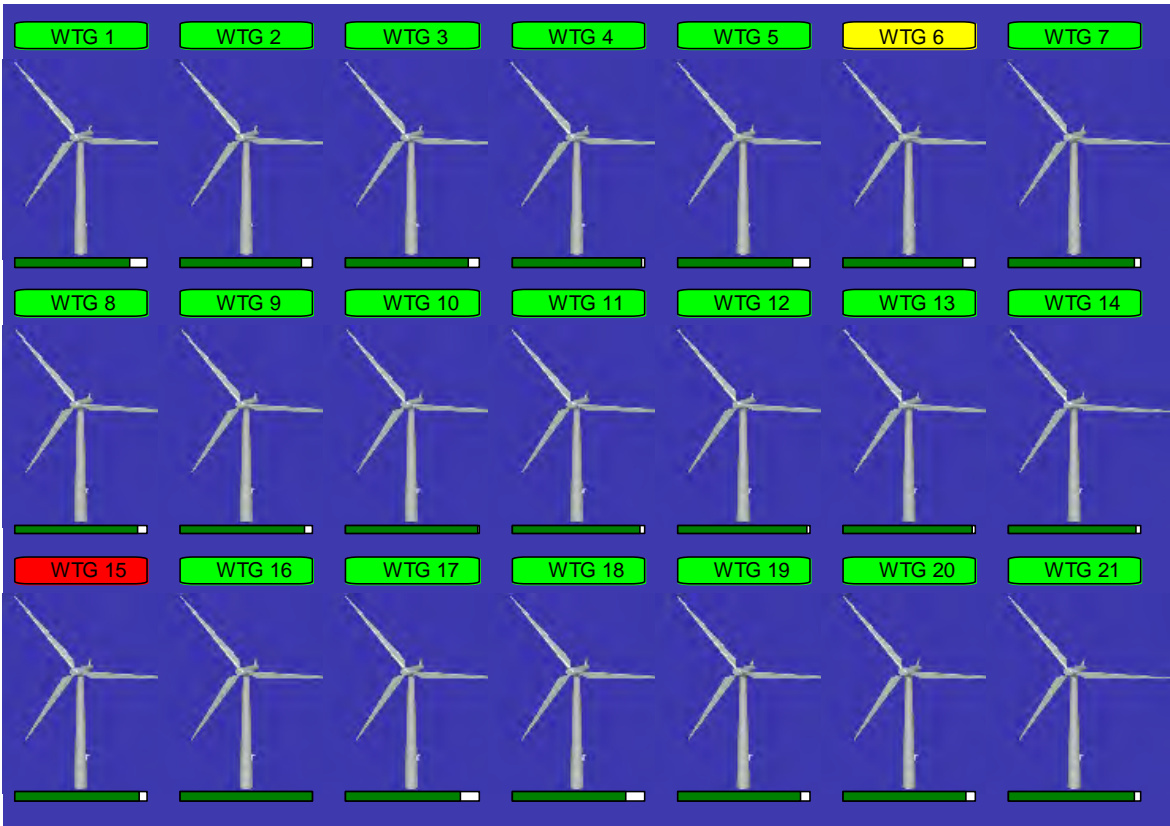


Figure 5.4: Condition visualization at the power plant level

The turbine layer can be accessed by clicking on the turbine of interest and the component layer by clicking on the component. The two layers are depicted in Figure 5.5. The description of the components can be found in section 1.4.

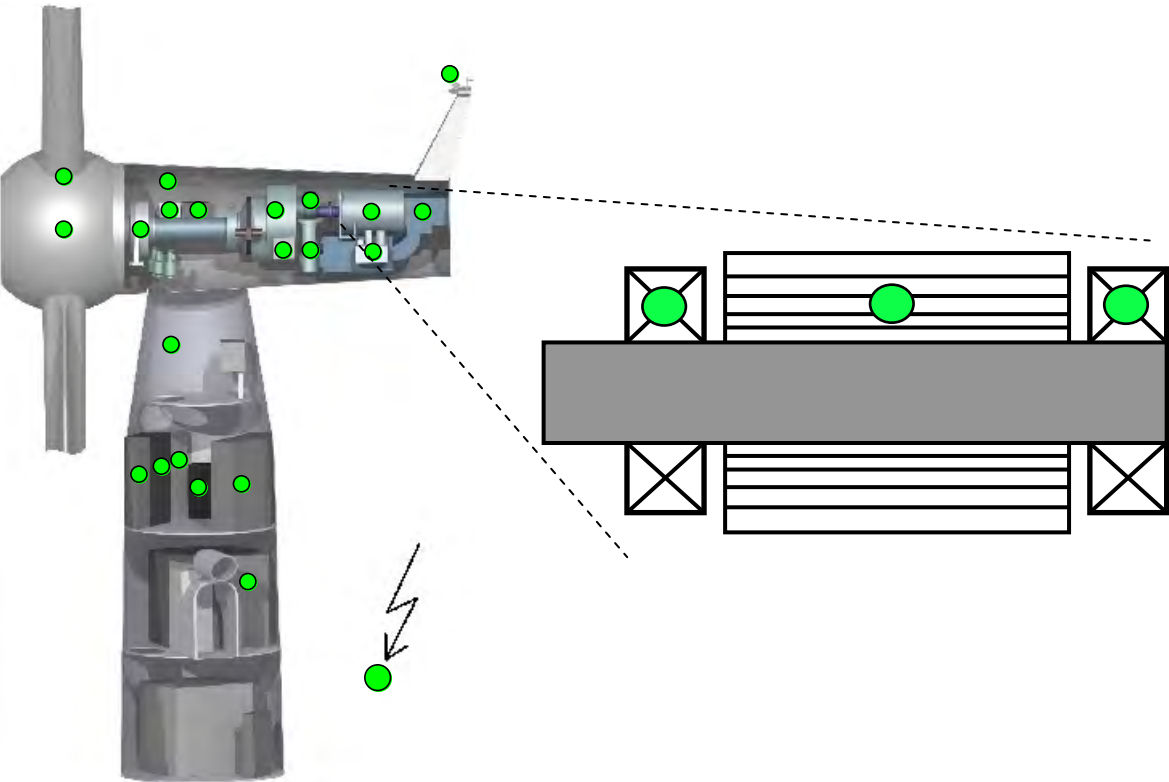


Figure 5.5: Condition visualization: Turbine level (left); Component level (generator)(right)

5.2.1. Example 1: HSS bearing defect

In this example a HSS bearing defect at WTG 01 is diagnosed by the global fuzzy expert system. On the bearing surface straight axial cracks occurred, which at a later stage lead to pitting at the crack edges. Figure 5.6 shows the defect after bearing replacement.

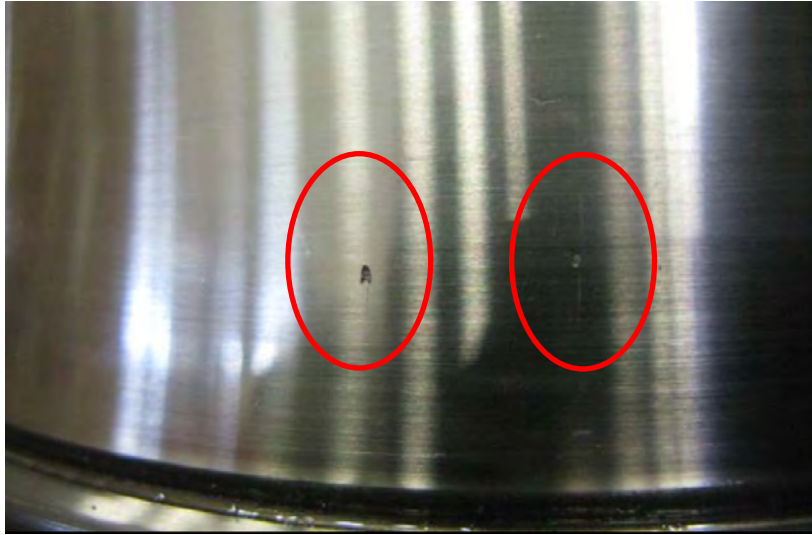


Figure 5.6: Axial cracks and pitting at the HSS bearing inner race at WTG 01

In section 4.1 it has been shown that the KEAS value is a good indicator of this type of fault. In publication [P5] the rules for diagnosis were established purely based on the output of the frequency content identifier, i.e. on the envelope spectrum analysis. In section 5.1 rule examples were presented which combine the information coming from the KEAS, containing information about the fault severity and the output of the frequency content identifier, containing information about the defective component. However, in this particular case first signs of fault were indicated by a rule monitoring the raw time series RMS value at the component level:

If Raw TS RMS **high** then Condition **yellow** & Diagnosis **Raw TS RMS high** & Potential root cause **ambiguous**.

The TS RMS evolution is shown in Figure 5.7.

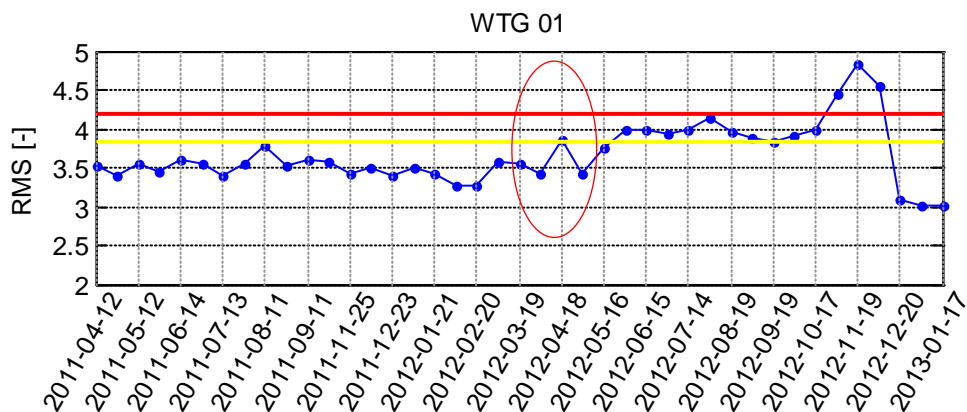


Figure 5.7: Raw TS RMS evolution (WTG 01: HSS stage TS 16000)

By the time the RMS value exceeds the threshold, the envelope spectrum did not have any distinct frequency components allowing component identification. A first threshold violation (for definition see section 4.2.1) occurs at 2012-04-18. In the CMS the fault is highlighted by a warning at the component level. Figure 5.8 shows the alarm status of the turbine at this date.

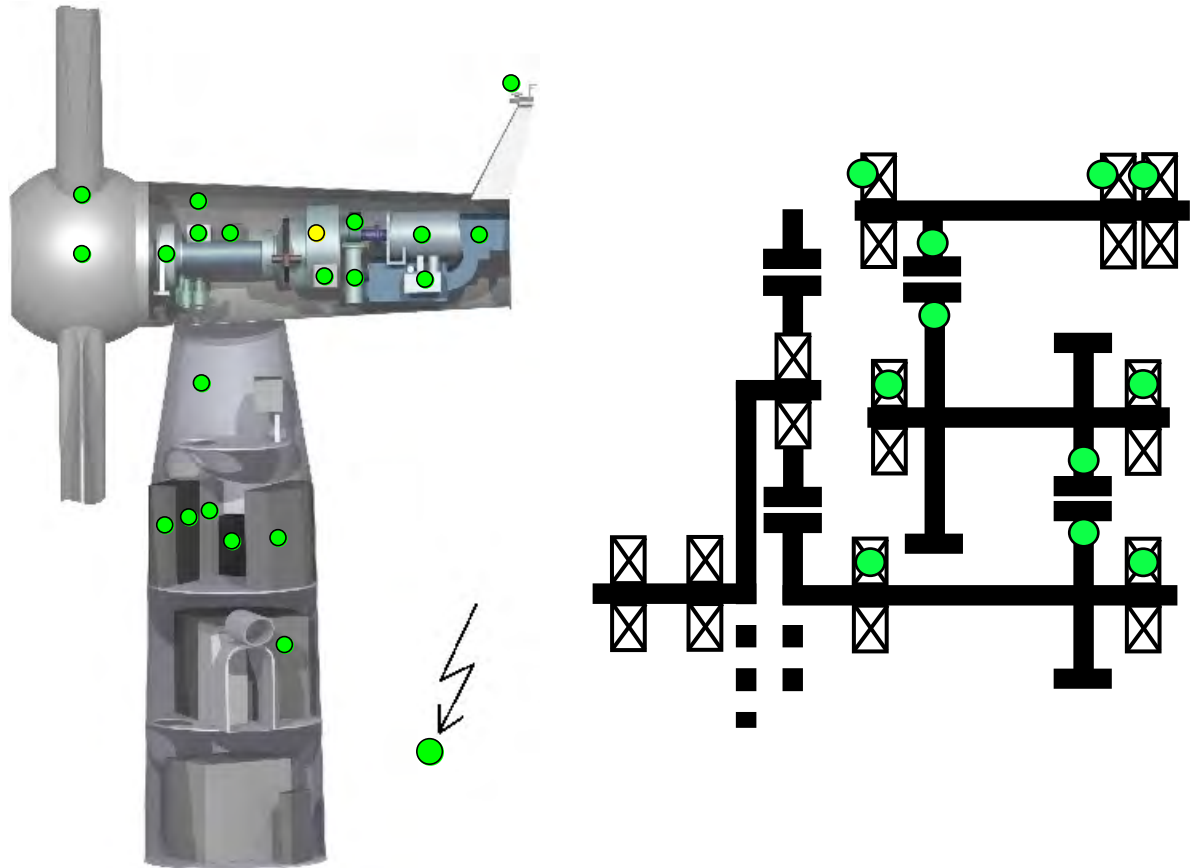


Figure 5.8: Condition visualization at 2012-06-01: Turbine level (left); Component level (Gearbox)(right)

It is visible, that the alarm is highlighted at the component level only, due to the missing identification of the root cause. On the basis of the raw TS RMS value alone the defective component cannot be identified, since the reason for an increased RMS value can be manifold. The RMS value at the next sequential measurement is considered normal again, before another threshold violation occurred at 2012-06-01. From this date onwards the RMS value stabilizes at a high level.

Beginning on 2012-06-29 (2.5 month later) the KEAS starts increasing (see Figure 5.9).

The standard thresholds separating the terms high (condition yellow) and very high (condition red) are 1.84 and 1.88 respectively. In comparison to the maximum value of 9.4 occurring during fault propagation, these thresholds are close together, which is caused by the low variance in the reference values. The general rule indicating the very high KEAS level at the component level is:

If KEAS **very high** then Condition **red** & Diagnosis **KEAS very high** & Potential root cause **ambiguous**.

The characteristic defect frequencies are not visible in the envelope spectra before 2012-07-28. Figure 5.10 illustrates waterfall plot of the squared envelope spectra.

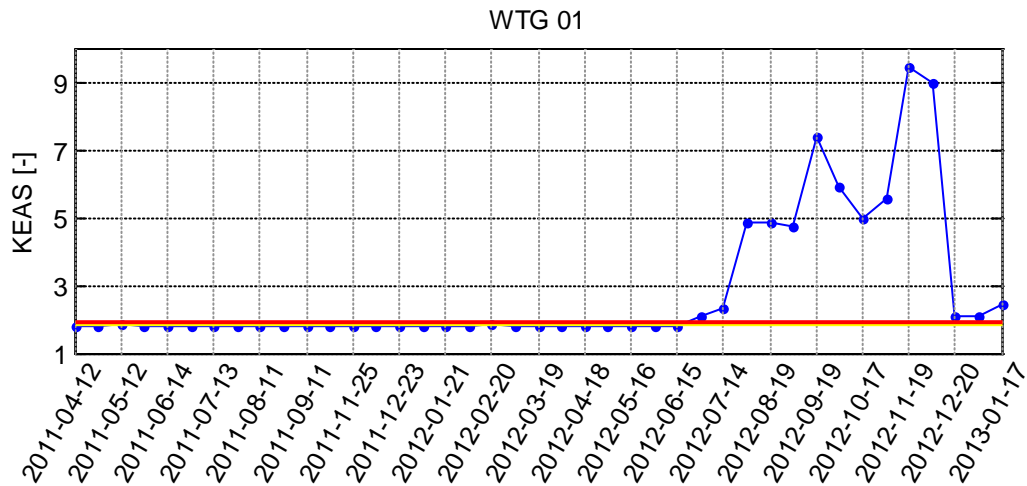


Figure 5.9; KEAS evolution (WTG 01: HSS stage TS 16000)

With presence of characteristic defect frequencies, the global fuzzy expert system finds a rule at the component detail level that suits the present pattern at 2012-07-28. It is:

If KEAS very high & BPFI_h0 present & BPFI_h1 present & BPFI_h2 present & Mod_RPM_h0_BPFI_h0 present then Condition red & Diagnosis bearing inner race defect & Potential root cause axial crack.

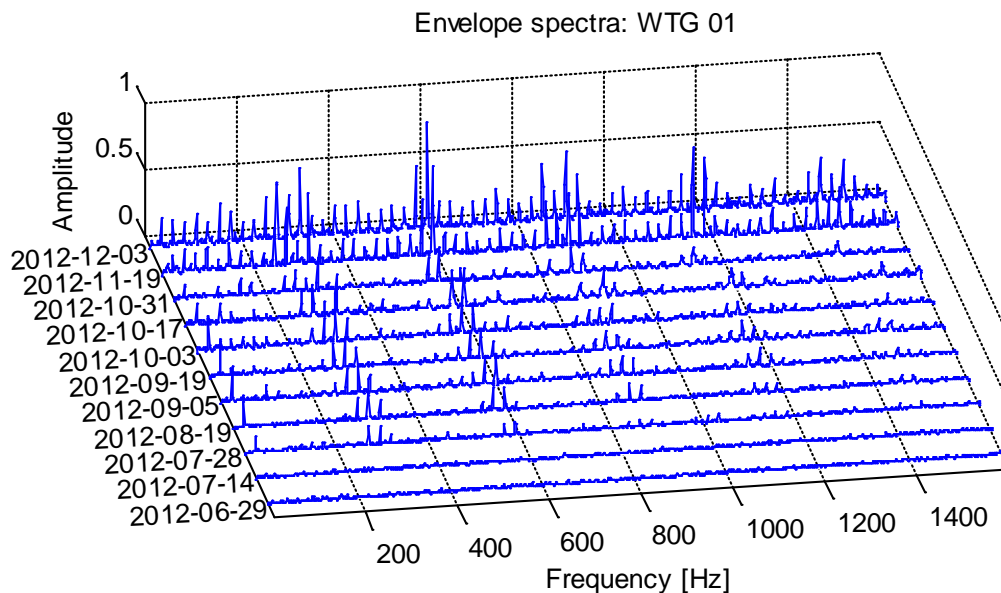


Figure 5.10: Envelope spectra during fault progression (WTG 01: HSS stage TS 16000)

In the CMS the output of the global fuzzy expert system is visualized as shown in Figure 5.11.

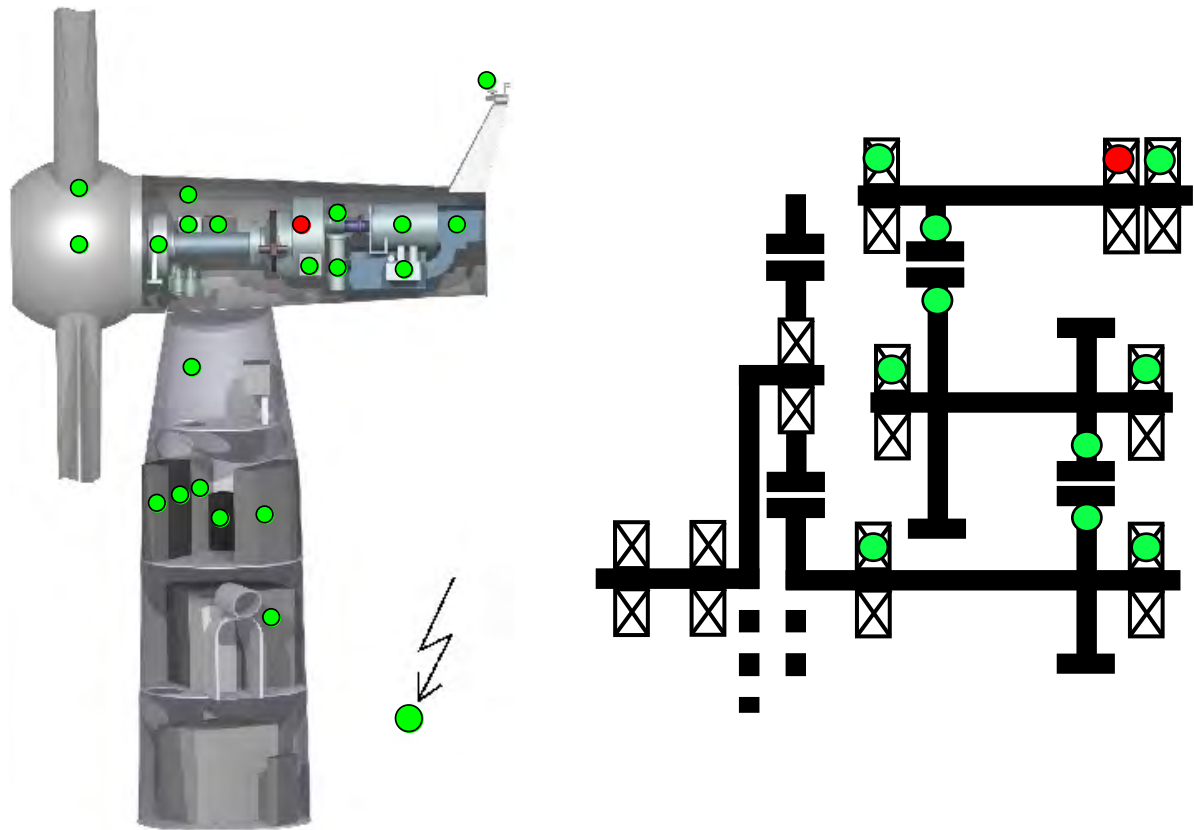


Figure 5.11: Condition visualization at 2012-07-29: Turbine level (left); Component level (Gearbox)(right)

After the fault severity has been identified (e.g. by endoscopic investigation or inspections after component replacement) the thresholds can be adapted to better represent the actual component conditions. In this example, the thresholds characterizing the condition red has been lifted to a KEAS value of five.

Bearing replacement took place 2012-12-06. Note that the KEAS value beyond this date does not fall back its original value (see Figure 5.9) because of a simultaneously developing IMS bearing defect from 2012-12-20 detected by the rules applied in the next example.

5.2.2. Example 2: IMS bearing defect

The IMS bearing defect example presented here is from WTG 21. No pictures of the fault can be presented, as the bearing has not been replaced yet. However, at this type of gearbox it is likely that the fault mechanism is similar to the observed HSS bearing defects, i.e. axial cracks with pittings at the crack edges at a later defect stage.

First sign of anomaly were highlighted by a rule at the component level at 2012-10-17. The rule which fired reads:

If KEAS high then Condition yellow & Diagnosis KEAS high & Potential root cause ambiguous.

The rule is the same as the one, which fired in example 1 before the characteristic fault frequencies occurred in the envelope spectra. Differences exist in which KEAS value is considered here. While for the HSS bearing defects the HSS stage sensor is evaluated, for the

IMS bearing defects the IMS stage sensor is considered. Hence, at the component level the KEAS of the TS 16000 as well as 6000_Tr is monitored. The specific rule at the component level considers on the KEAS 6000_Tr for the IMS bearings. The reason is for this is that the lower rotational speed of the bearing calls for longer TS records to capture more fault overrollments. Usage of the TS 6000_Tr is possible, since the resonance frequencies excited by fault overrollment, are of lower frequency. For the HSS bearings, the excited resonance frequency is about 20 kHz, while the IMS defect excites a resonance frequency of about 2 kHz. In the consequence, the HSS bearing defects are hardly visible in the TS 6000_Tr (sampling freq. about 15.38 kHz), while they are obvious in the TS 16000 (sampling freq. of 41.66 kHz).

The KEAS for the TS 6000_Tr evolution is shown in Figure 5.12.

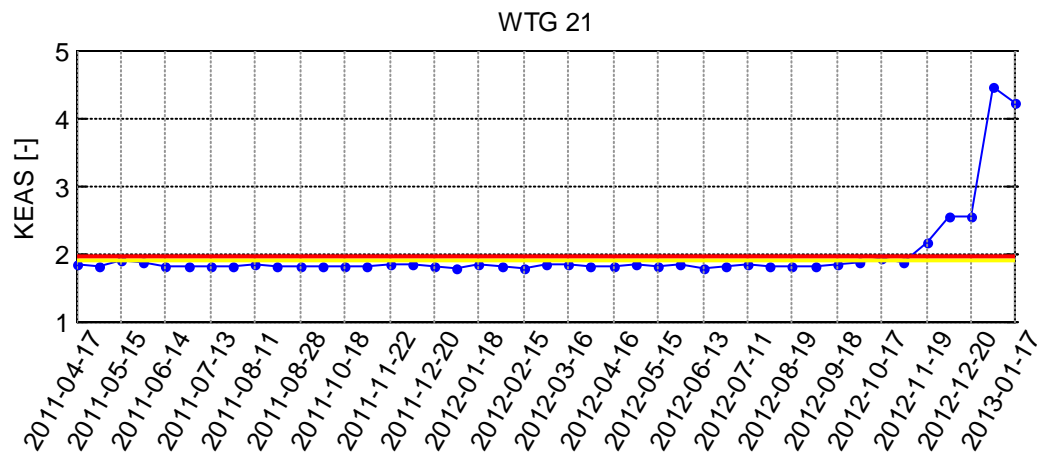


Figure 5.12: KEAS evolution (WTG 21: IMS stage TS 6000_Tr)

As in the previous example, the determined thresholds lie close together (yellow: 1.89; red: 1.97). For condition separation, these thresholds are not well chosen, but the low variance in the reference values considered lead to high trust in this parameter for fault detection.

Based on the experience gained with this type of fault at the other WTGs with IMS bearing defects, the rules on the component detail level have been modified to better capture the present fault signatures. It was observed, that the modulation amplitudes around the defect frequency sometimes exceed the center frequency. The frequency content identifier thus is unable to detect those peaks as modulations. The two rules at the component detail level for the IMS bearings are thus:

If KEAS **high** & BPFI_h0 **present** & BPFI_h1 **present** & BPFI_h2 **then** Condition **yellow** & Diagnosis **bearing inner race defect** & Potential root cause **ambiguous**.

If KEAS **very high** & BPFI_h0 **present** & BPFI_h1 **present** & BPFI_h2 **then** Condition **red** & Diagnosis **bearing inner race defect** & Potential root cause **ambiguous**.

The defect frequencies and its harmonics are identified by the frequency content identifier in the envelope spectrum already from 2012-10-17. Consequentially the defective component was identified from the first day of anomaly.

Figure 5.13 shows a waterfall plot of the amplitude envelope spectra during fault progression.

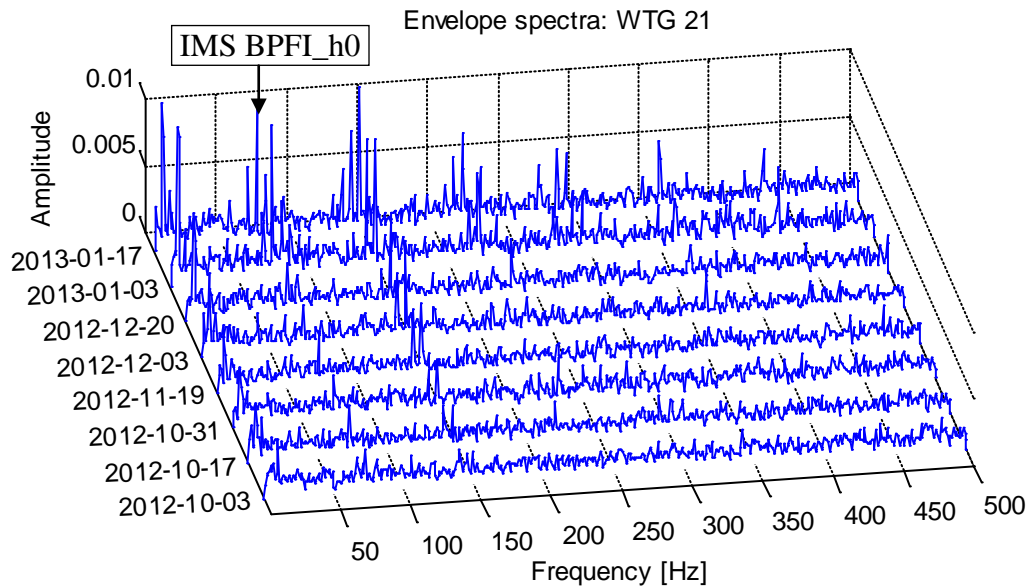


Figure 5.13: Envelope spectra during fault progression (WTG 21: IMS stage TS 6000_Tr)

The envelope spectra well correspond to the observed increase in the KEAS. The condition visualization at the turbine and the component level at 2012-10-07 is illustrated in Figure 5.14.

After the initial KEAS increase at 2012-10-17 the next measurement at 2012-10-31 is normal again, before a trend establishes from 2012-11-19. As a result of the KEAS being above the red threshold, the assigned condition is red.

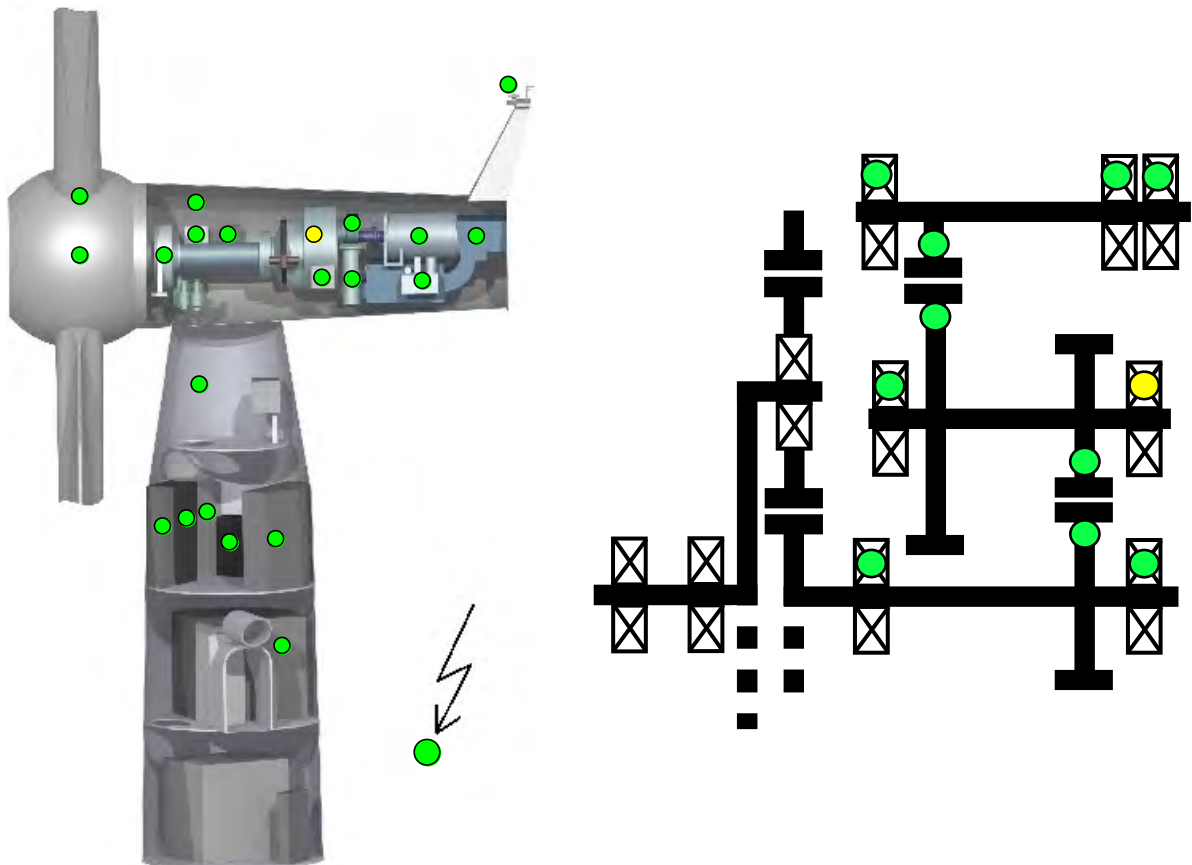


Figure 5.14: Condition visualization at 2012-10-17: Turbine level (left); Component level (Gearbox)(right)

A final modification of the threshold to match the actual condition is not made due to lack of information about the actual fault severeness. This can be done after bearing replacement and final investigations.

5.2.3. Example 3: Generator NDE bearing defect

This example emphasizes the benefit of combining the information coming from SCADA and vibration data. At WTG 06 a generator NDE bearing degradation is observed, which was first picked up by vibration measurements, but at the same time lead to an increase in bearing temperature. While the vibrations normalize a few measurements after their onset the temperature measurement remained abnormal, further indicating a bearing degradation. The generator NDE bearing is difficult to access, due the generator cooling components mounted around the generator NDE. Hence, a visual inspection was not yet instructed. The turbine is still operational.

The CMS highlighted rolling element defect at 2012-07-13 at the generator NDE bearing with the condition red (see Figure 5.16), based on the following rule:

If KEAS very high & FTF_h0 present & FTF_h1 present & Mod_BSF_h0_FTF_h0 present then Condition red & Diagnosis Rolling element signature & Potential root cause Rolling element defect.

The KEAS is very high at this date (see Figure 5.15), but decreases to its normal level already two measurements later.

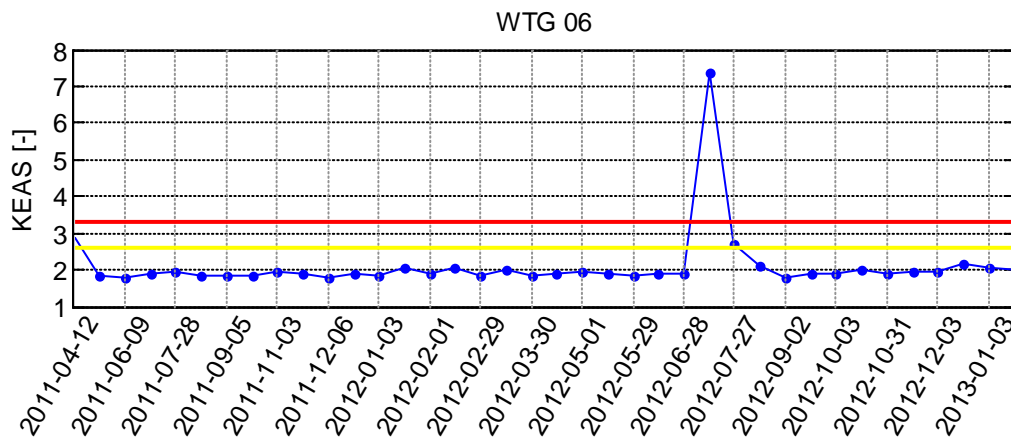


Figure 5.15: KEAS evolution (WTG 06: Generator NDE TS 6000_Tr)

The corresponding envelope spectra show rich harmonics of the FTF and the BSF. A waterfall plot of the envelope spectra is depicted in Figure 5.17.

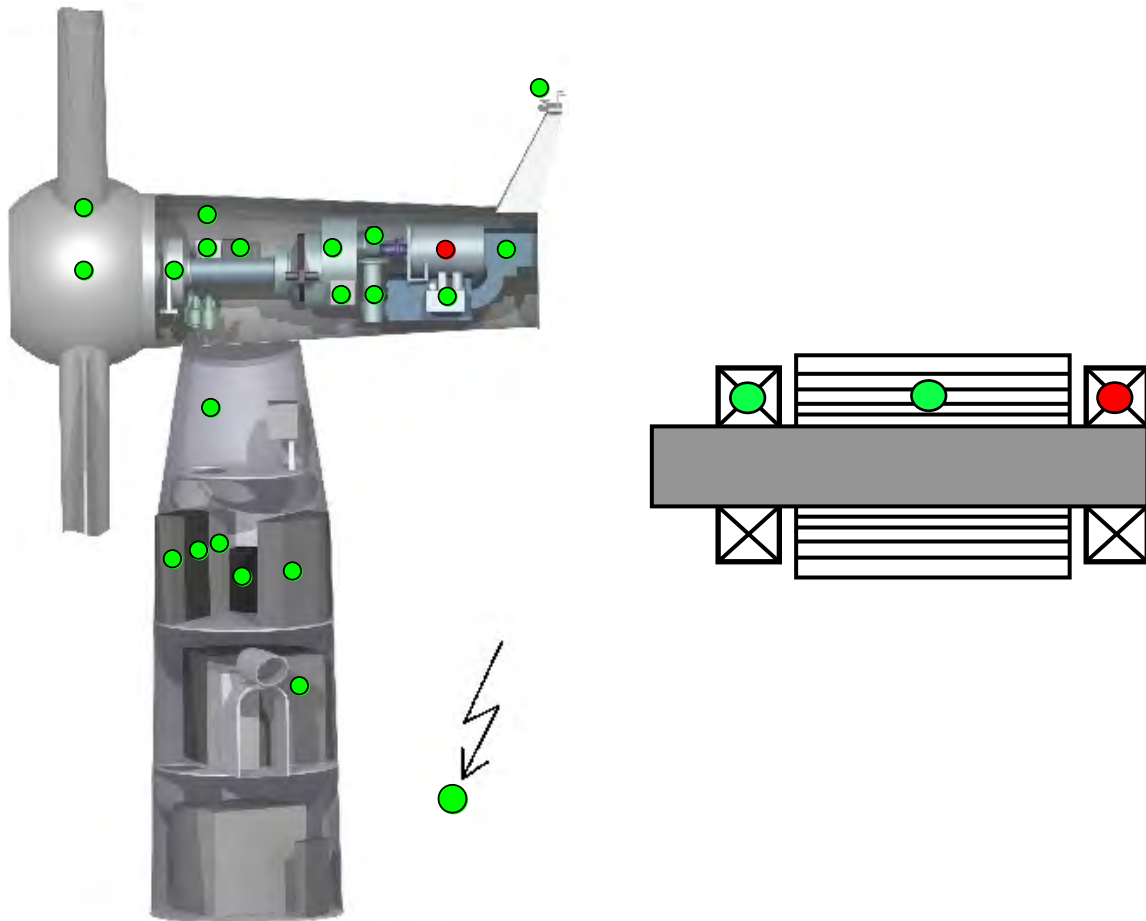


Figure 5.16: Condition visualization at 2012-07-13: Turbine level (left); Component level (Generator)(right)

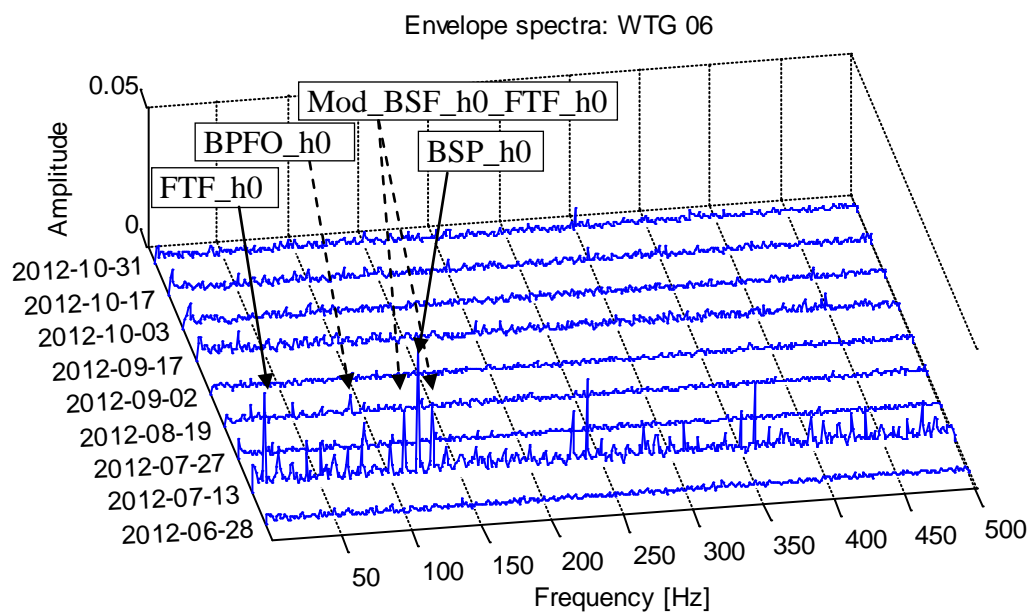


Figure 5.17: Envelope spectra during anomaly (WTG 06: Generator NDE TS 6000_Tr)

The fault pattern is a single event present at 2012-07-13 only. At the sub sequential measurements the BSF component vanishes together with the rich harmonics of the FTF. At the next measurement the BPFO frequency component appears in the spectrum, but also

decreases over the next measurements. One explanation for such a pattern can be that the bearing was contaminated by one or more particles, initially sticking to one of the roller elements. This particle potentially loosened and stuck to the outer race and was eventually flushed from the bearing at some point in time.

However, the event likely caused some defect to the bearing, expressing itself in a temperature increase of 2°C. The prediction error evolution of the generator NDE temperature is shown in Figure 5.18.

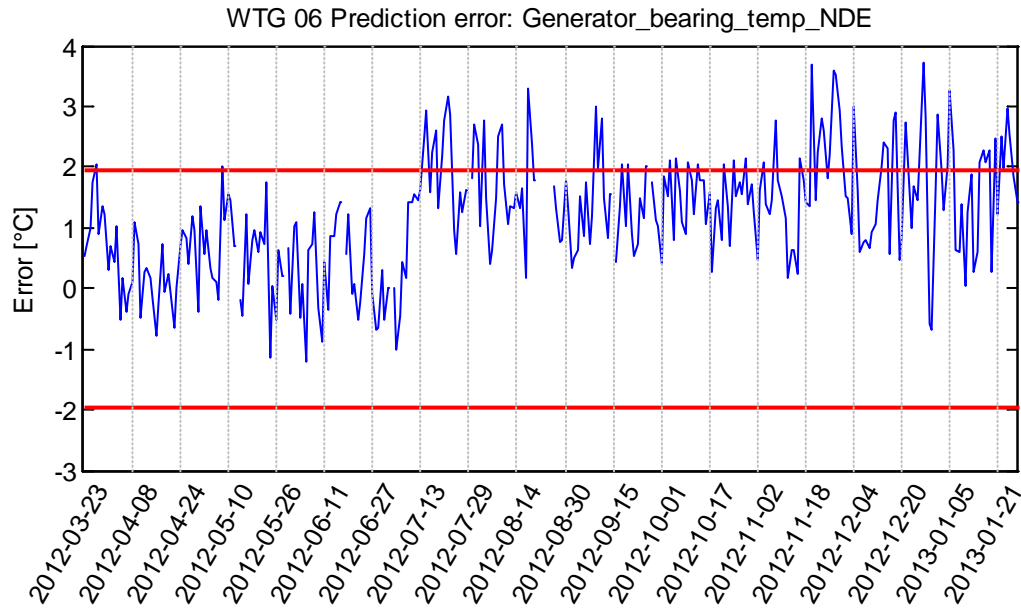


Figure 5.18: Generator bearing temperature NDE prediction error evolution (WTG 06)

The increase is highlighted at the component detail level via the following general rule.

If Anomaly error (5 MF's) Generator bearing temp. NDE **high** **then** Condition **yellow** & Diagnosis **Generator bearing temp. NDE high** & Potential root cause **ambiguous**.

The rule is assigned a lower weight as the rules considering the vibration data, as less specific information is obtained. Hence, the rule first fired after the frequency components vanish. The increase in temperature is, however, not large enough to cause a permanent violation of the threshold. In the global fuzzy expert system this behavior causes the assigned condition to fluctuate, between normal and yellow. The expert has the ability to counteract this by adapting the thresholds and thus setting new alarm limits, while appreciating that the bearing may have a reduced remaining lifetime. Hence, the CMS aided the expert in taking informed decisions. Note that such small temperature changes are not visible in the raw data, due to the annual temperature cycle and other fluctuations caused by the different operational modes, emphasizing the strength of NBM based approaches.

6 Conclusions and future aspects

6.1. Conclusion

The methodology presented allows monitoring of a broad variety of different signals to evaluate the turbine performance and its components condition. In the area of SCADA data analysis it has been shown that models developed with a FSRC approach allow earlier anomaly detection and give larger shifts in mean than models build with an autoregressive approach. This property gives advantages in condition classification, which relies on a distinct separation of normal and abnormal signal behavior.

It has also been shown that ANFIS models perform better than NN models, in terms of both the prediction error and the computational time required for model training. With the large amount of SCADA data coming from modern wind power plants, efficient algorithms are required. With 45 models developed to monitor the SCADA data of a turbine of power plant 1 and 80 for a turbine of power plant 2, 2490 models were required to be trained throughout this research emphasizing the need of efficiency. Due to service and maintenance or component replacements, models may require additional retraining several times throughout the turbine live time.

The application of the FSRC approach in combination with ANFIS models in the given examples proves that SCADA data contain information useful for condition monitoring. Although each turbine type has a different sensor setup, SCADA data are usually broadly available to operators and are often the only possibility of monitoring a turbine without installation of additional sensors. The proposed methodology thus provides a practical solution to the often required low cost monitoring of wind turbines. Dependent on the available budget for monitoring, operators can set up their own CMS to monitor the available data to the extent that suits their needs. However, it is recommended that at minimum the turbine power output should be monitored.

Difficulties with monitoring wind turbine SCADA data arise from the data quality. Service and maintenance actions performed frequently influence the SCADA data and may cause false alarms. The same is true for changes to the turbine control system in terms of updates or parameter adoptions. Signals that are frequently disturbed by service actions may need to be excluded from monitoring and may not be considered as input signal. However, the reliability and stability of the signals over time is not known beforehand, requiring some effort before good input output sets are identified. Research showed that for identification of suitable

model input and output sets at least one year of measurements is required to cover a full service cycle. In this research, different training levels have been used to allow early monitoring after one, three, six and nine month of data have been collected. Models based on unreliable input signals may need to be refined in these periods.

In the area of vibration data analysis signal PW proved a useful preprocess for discrete random separation required for bearing diagnosis. Research showed that this efficient method can be applied to the comparably complex vibrations coming from wind turbine gearboxes. Discrete components e.g. originating from gears are reliably removed from the signal improving the fault visibility in the envelope spectrum, which is otherwise dominated by these components. PW is proved applicable also for discrete random separation at longer time series, as they are required for defect detection at wind turbine gearbox bearings rotating with intermediate speed.

A disadvantage caused by PW is that the remaining signal amplitude scaling is somewhat arbitrary. Hence, the amplitude height is not stable preventing definition of global amplitude thresholds and other methods must be applied for condition classification.

In order to identify the fault severeness the KEAS value is proposed, which gives some advantage over the SK in terms of earlier fault detection and level stability at advanced defect stages. The latter causes condition classifications to be more stable which is desirable in a CMS. Diagnostic information can be obtained by combining the KEAS value with the information contained in the envelope spectrum pattern. Using a frequency content identifier known frequency components (e.g. originating from the system kinematics) can be identified and build the basis for automated fault diagnosis. Automated fault diagnosis is important considering the large amount of turbines requiring monitoring. It furthermore, brings down the monitoring cost, which in turns helps reducing the cost of energy.

Automation of fault analysis can be achieved by use of fuzzy logic. The different turbine configurations call for methods that can cope with a certain level of imprecision. Pattern present at one turbine maybe not be the exact same at another turbine, even if the turbines are of the same type and brand. The global fuzzy expert system proposed can deal with this imprecision and allows rules to be defined linking different information sources, e.g. information coming from SCADA and vibration data. It is expected that this possibility not only gives the expert the maximum amount of freedom in setting up rules representing the present pattern, but also improves diagnosis certainty.

Research has shown that cases exist, where reliable diagnosis is only possible when information from different sources is combined.

With the proposed methodology to monitor wind turbine SCADA data and the global fuzzy expert system presented, tools are available to evaluate the condition of the entire turbine, instead of focusing on single components or component groups only. If required, specialized algorithms can be developed/ applied and their output linked to the other information available. This gives a major contribution as it lifts the discussion about CMS for wind turbines to the turbine/system level, and thus should end the common association of wind turbine CMS with vibration drive train monitoring only. The latter although important should rather be seen as part of a global condition monitoring system.

6.2. Future aspects

The contribution of the thesis opens up for supplementary research in this multidisciplinary field. The proposed method for monitoring SCADA data proved useful to detect anomalies in these data once proper input-output sets have been identified. Further research is required with respect to aiding the input selection. Here, a combination of the physical understanding of the system to be modeled and advanced data reduction techniques has been applied. However, the data reduction technique often outputs sister signals (e.g. the voltage phase R to model voltage phase T). These signal pairs are undesirable in many cases and a manual selection process is still required.

With respect to the surveyed HSS and IMS bearing defects it should be noted that these defects were local defects for which automated fault diagnosis is feasible. Further research is required about the possibility of automated distributed defect diagnosis. In this context, also the SK and KEAS require further comparison for different faults to generalize the findings of this research. Moreover, the slow rotating and planetary bearings have not been in focus of the research here. Methods for defect diagnosis at these components are still an open area of research in particular for wind turbine applications, as the signal to noise ratio is usually low and turbines operate with varying speed and load. Apart from this, difficulties arise at planet bearings also from the fact that the planet location changes over time as the planets rotate about the sun. With a fixed point measurements (as usually applied), the signal picked up can thus not be assumed stationary. Furthermore, bearing defects at the planets need to be transmitted through the teeth in mesh to be picked up at the gearbox housing, representing a difficult and non-constant signal transfer path.

Likewise, it is necessary to further validate the described algorithms and parameters for wind turbine gear defect diagnosis. In particular with respect to their capability of early fault detection using real measured data from a defect period at an operating unit. So far comparisons either are based on laboratory tests or consider only a few measurements, neglecting the requirement of parameter stability in a healthy case. A promising method identified is wavelet analysis. However, problems exist in feature extraction as validation cases in literature show that fault pattern can be quite different in the wavelet scalogram. Deriving parameters to indicate the different types of defect remains a challenge.

At planetary gears, the major challenges are the change in signal transfer path caused by the movement of the planets and the low rotational speed. In the past methods have been proposed to monitor planetary gears, but these methods have not sufficiently been validated for modern type wind turbines operating at variable speed. Thus, a comparison of different methods is desirable using real measured wind turbine planetary gear faults.

The global fuzzy expert system should be extended to take inputs from oil analysis reports (or oil quality sensors), oil particle counters as well as blade vibration measurements into account. For this purpose, suitable algorithms must be identified and validated.

With a rising complexity of the global fuzzy expert system, it becomes important to have a reference database of fault pattern with known root causes. With such a database, it could be checked, if the implementation of a new rule has a backlash on the system performance to detect the already known faults. Research should focus on automated validation of the system performance and reporting of possible rule/ expert knowledge conflicts.

7 References

7.1. References cited in thesis

- [1] **Yang, W., Jiang, J.** (2011). "Wind turbine condition monitoring and reliability analysis by SCADA information. *"Second International Conference on Mechanic Automation and Control Engineering (MACE), 15-17 July, Blyth, UK, pp.1872-1875.*
- [2] **Petersen, K.E.** (1993). "Reliability based maintenance systems; A knowledge based approach. *"Proc. of the Fourth Scandinavian Conf. on Artificial Intelligence Electrum, Stockholm, Sweden, 4-7 May, pp.409-410.*
- [3] **Clayton, B.R., Dutton, A.G., Aftab, N., Bond, L.J., Irving, A.D., Lipman, N.H.** (1990). "Development of Structural condition monitoring techniques for composite wind turbine blades. *"Proc. of European Community Wind Energy Conference, Madrid, Spain, 10-14 September, pp.10-14.*
- [4] **Sturm, A., Billhardt, S.** (1991). "Envelope curve analysis of machines with rolling-element bearings. *"Proc. of Saveprocess 91 (2), pp.275-279.*
- [5] **Chow, M.Y., Mangum, P.M., Yee, S.O.** (1991). "A neural network approach to real time condition monitoring of induction motors. *"IEEE Transactions on Industrial Electronics 38 (6), pp.448-454.*
- [6] **Penman, J., Tavner, T.J.** (1989). "Condition monitoring of electrical machines. *"New York: Research Studies Press Ltd., Wiley.*
- [7] **Penman, C.M., Yin, J.** (1994). "Feasibility of using unsupervised learning, artificial neural networks for the condition monitoring of electrical machines. *"IEE Proc. Electrical Power Application 141 (6), pp.317-323.*
- [8] **Caselitz, P., Giebhardt, J., Mevenkamp, M.** (1994). "On-line fault detection and prediction in wind energy converters. *"Proc. of the European Wind Energy Conference, Thessaloniki, Greece, 10-14 October, pp.623-627.*
- [9] **Frank, P.M.** (1994). "Diagnoseverfahren in der Automatisierungstechnik. *"at-Automatisierungstechnik 42 (2), pp.47-64.*
- [10] **Sorsa, T., Koivo, H.N.** (1991). "Application of artificial neural networks in process fault diagnosis. *"Proc. of Saveprocess 91 (2), pp. 133-138.*
- [11] **Caselitz, P., Giebhardt, J., Krüger, T., Mevenkamp, M.** (1996). "Development of a fault detection system for wind energy converters. *"Proc. of the European Wind Energy Conference, Goteborg, Sweden, 20-24 May, pp.1004-1007.*
- [12] **Caselitz, P., Giebhardt, J., Mevenkamp, M.** (1997). "Application of condition monitoring systems in wind energy converters. *"Proc. of the European Wind Energy Conference, Dublin, Ireland, 6-9 October, pp.579-582.*

-
- [13] **Caselitz, P., Giebhardt, J.** (1999). "Advanced condition monitoring for wind energy converters. " *Proc. of the European Wind Energy Conference, Nizza, France, 1-5 March*.
 - [14] **Li, S., O'Hair, E., Giesselmann, M.G.** (1997). "Using neural networks to predict wind power generation. " *Proceedings of the International Solar Energy Conference, Washington D.C. pp.415-420*.
 - [15] **Verbruggen, T.W.** (2003). "Wind Turbine Operation & Maintenance based on Condition monitoring. " *ECN-C-03-047*.
 - [16] **McArthur, S.D.J., Booth, C.D., McFadyen, I.T.** (2005). "An Agent based anomaly detection architecture for condition monitoring. " *IEEE Transactions on Power Systems* 20 (4), pp.1675-1682.
 - [17] **Sanz-Bobi, M. A., del Pico, J., Garcia, M. C.** (2006). "SIMAP: Intelligent System for Predictive Maintenance Application to the health condition monitoring of a windturbine gearbox. " *Computers in Industry Vol.57, pp.552-568*.
 - [18] **Zaher, A.S., 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*.
 - [19] **Li, S., Wunsch, D.C., O'Hair, E., Giesselmann, M.G.** (2001). "Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation. " *Journal of Solar Energy Engineering Vol. 123 pp.327-331*.
 - [20] **Kusiak, A., Zheng, H., Song, Z.** (2009). "Models for monitoring wind farm power. " *Renewable Energy Vol.34 pp.583-590*.
 - [21] **Kusiak, A., Zheng, H., Song, Z.** (2009). "On-line monitoring of power curves. " *Renewable Energy Vol.34, pp.1487-1493*.
 - [22] **Xiang, J., Watson, S., Liu, Y.** (2009). "Smart monitoring of wind turbines using neural networks. " *Sustainability in Energy and Buildings Part 1, pp.1-8*.
 - [23] **Zaher, A., McArthur, S.D.J., Infield, D.G.** (2009). "Online Wind Turbine Fault Detection through Automated SCADA Data Analysis. " *Wind Energy Vol.12, pp.574-593*.
 - [24] **Gae, X. Z., Ovaska, S. J.** (2001). "Soft computing methods in motor fault diagnosis. " *Applied Soft Computing Vol.1, pp.73-81*.
 - [25] **Haykin, S.** (1998). "Neural Networks: A Comprehensive Foundation. " ISBN: 0132733501. Prentice-Hall.
 - [26] **Schlechtingen, M., Santos, I.F.** (2011). "Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. " *Mechanical Systems and Signal Processing Vol.25 No.5, pp.1849-1875*.
 - [27] **Tarassenko, L.** (1998). "Guide to Neural Computing Applications. " *Elsevier, ISBN:0340705892*.
 - [28] **Rafiq, M.Y., Bugmann, G., Easterbrook, D.J.** (2001). "Neural network design for engineering applications. " *Computers and Structures Vol.79, pp.1541-1552*.
 - [29] **Jang, J. S. R.** (1993). "ANFIS: Adaptive-Newtork-Based Fuzzy Interference System. " *Transactions on Systems, Man and Cybernetics Vol.23 No.3, pp.665-685*.
 - [30] **Jang, J. S. R., Sun, C.T., Mizutani, E.** (1997). "Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligence. " *Prentice Hall, ISBN: 0-13-261066-3*.

-
- [31] **Randall, R.B.** (2011). "Vibration-based Condition Monitoring. "Wiley, ISBN:978-0-470-74785-8,.
 - [32] **Xiaofei , Z., Niaoqing, H., Bin, F., Zhe, C.** (2012). "Enhanced fault detection of rolling element bearing based on cepstrum editing and stochastic resonance. "*Journal of Physiscs: Conference Series Vol.364*, pp.1-8.
 - [33] **Sawalhi, N., Randall, R.B.** (2011). "Signal pre-whitening using cepstrum editing (liftering) to enhance fault detection in rolling element bearings. "*Proceedings of the 24th International Congress on Condition Monitoring and Diagnostics Engineering Management*, pp.330-336.
 - [34] **Akaike, H.** (1973). "Information theory and an extention of the maximum likelihood principle . "*Decond International Symposium on Information Theory*, pp.267-281, Budapest.
 - [35] **Dalpiaz, G., Rivola, A., Rubini, R.** (2000). "Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears. "*Mechanical Systems and Signal Processing Vol.14*, pp.387-412.
 - [36] **Randall, R.B., Hee, J.** (1981). "Cepstrum analysis. "*Brüel&Kjær Technical Review Vol.3*, pp.3-40.
 - [37] **Wang, W.Q., Ismail, F., Golnaraghi, M.F.** (2001). "Assessment of gear defect monitoring techniques using vibration measurements. "*Mechanical Systems and Signal Processing Vol.15*, pp.905-922.
 - [38] **Siegel, D., Zhao, W., Lapira, E., AbuAli, M., Lee, J.** (2013). "A comparative study on vibration-based condition monitoring algorithms for wind turbine drive trains. "*Wind Energy*, doi: 10.1002/we.1585.
 - [39] **Zimroz, R., Bartelmus, W.** (2009). "Gearbox condition estimation using cyclostationary properties of vibration signal. "*Key Engineering Materials, Vol.413-414*, pp.471-478.
 - [40] **Antoni, J.** (2009). "Cyclostationarity by examples. "*Mechanical Systems and Signal Processing, Vol. 23*, pp.987-1036.
 - [41] **Antoni, J.** (2007). "Cyclic spectral analysis in practise. "*Mechanical Systems and Signal Processing, Vol.21*, pp.597-630.
 - [42] **McFadden, P.D.** (1987). "Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the meshing vibration. "*Mechanical Systems and Signal Processing Vol.1(2)*, pp.173-183.
 - [43] **McFadden, P.D., Smith, J.D.** (1985). "A signal processing technique for detecting local defects in a gear from the signal average of vibration. "*Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, Vol.199*, pp.287-292.
 - [44] **Bechhoefer, E., Kingsley, M.** (2009). "A review of time synchronous average algorithms. "*Annual Conference of the Prognostics and Health Management Society*.
 - [45] **Staszewski, W.J., Tomlinson, G.R.** (1994). "Application of the wavelet transform to fault detection in a spur gear. "*Mechanical Systems and Signal Processing, Vol.8(3)*, pp.289-307.

-
- [46] **Elbarghathi, F., Wang, T., Zhen, D., Gu, F., Ball, A.** (2012). "Two stage helical gearbox fault detection and diagnosis based on continuous wavelet transformation of time synchronous averaged vibration signals. *"Journal of Physics: Conference Series* 364, pp.1-11 .
- [47] **Halim, E.B., Shah, S.L., Zuo, M.J., Choudhury, M.A.A.S.** (2006). "Fault detection of gearbox from vibration signals using time-frequency domain averaging. *"Proceedings of the 2006 American Control Conference, Minneapolis, USA, June 14-16.*
- [48] **Hanna, J., Hatch, C., Kalb, M., Weiss, A., Luo, H.** (2011). "Detection of wind turbine gear tooth defects using sideband energy ratio. *"GE Energy.*
- [49] **Randall, R.B.** (1982). "A new method for modelling gear faults. *"Journal of Mechanical Design* Vol.104, pp.259-267.
- [50] **Brandt, A.** (2011). "Noise and vibration analysis. *"Wiley, ISBN: 978-0-470-74644-8.*
- [51] **Oppenheim, A.V., Schafer, R.W.** (2004). "From frequency to quefrency: a history of the cepstrum. *"IEEE Signal Processing Magazine, September, pp.95-104.*
- [52] **Serpedin, E., Panduru, F., Sari, I., Giannakis, G.B.** (2005). "Bibliography on cyclostationarity. *"Signal Processing, Vol.85, pp.2233-2303.*
- [53] **McFadden, P. D.** (1986). "Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration. *"Journal of Vibration, Acoustics, Stress and Reliability in Design, Vol. 108, pp.165-170.*
- [54] **Randall, R. B., Sawalhi, N.** (2011). "A new method for seperating discrete components from a signal. *"Sound and Vibration Vol.May, pp.6-9.*
- [55] **McFadden, P. D., Toozhy, M. M.** (2000). "Application of synchronous averaging to vibration monitoring of rolling element bearings. *"Mechanical Systems and Signal Processing Vol. 14(6), pp.891-906.*
- [56] **Randall, R. B., Antoni, J.** (2011). "Rolling element bearing diagnosis - A tutorial. *"Mechanical Systems and Signal Processing Vol.25 pp.485-520.*
- [57] **De Moortel, I., Munday, S.A., Hood, A.W.** (2004). "Wavelet analysis: the effect of varying basic wavelet parameters. *"Solar Physics, Vol.222, pp.203-228.*
- [58] **Szczepanik, A.** (1989). "Time synchronous averaging of ball mill vibrations. *"Mechanical Systems and Signal Processing Vol.3(1), pp 99-107.*
- [59] **Ebersbach, S., Peng, Z.** (2008). "Expert system development for vibration analysis in machine condition monitoring. *"Expert system with Applications, Vol.34, pp.291-299.*
- [60] **Conner, R.W., Beckman, K.O.** (2004). "API 613 fifth edition, special purpose gear units for petroleum chemical, and gas industry services-overview presentation. *"Proceedings of the turbomachinery symposium, Vol. 33, pp.135-142.*
- [61] **Starzhinskii, V.E.,** (2008). "Forms of defect to gear wheels: typology and recommendations on prevention. *"Journal of Friction and Wear, Vol.29-5, pp.340-353.*

7.2. References cited in papers [P1-P5]

- [62] **Süttmann (Schlechtingen), M.** (2010). "Master Thesis: Condition Monitoring in Wind Turbines - A Drive Train Monitoring System. "
- [63] **Blundell, S., K. M., Blundell.** *Concepts in thermal physics*. 2006. 978-0198567691.
- [64] **Swingler, K.** (1996). "Applying Neural Networks - A Practical Guide. "ISBN: 0340705892. Elsevier.
- [65] **Kusiak, A., Wenyan, L.** (2011). "The prediction and diagnosis of wind turbine faults. "*Renewable Energy Vol.36, pp.16-23*.
- [66] **Garlick, W. G., Dixon, R., Watson, S. J.** (2009). "A model-based approach to wind turbine condition monitoring using SCADA data. "*20th International Conference on Systems Engineering, 8-10 September, Coventry University, Coventry, UK*.
- [67] **Chandrashekhara, K., Chukwujekwu Okafor, A., Jiang, Y. P.** (1998). "Estimation of contact force on composite plates using impact-induced strain and neural networks. "*Composite Part B 29B, pp.363-370*.
- [68] **Ross, T. J.** (2010). "Fuzzy logic with engineering applications. "Wiley. 3rd.
- [69] **Kim, J., Kasabov, N.** (1999). "HyFIS: adaptive neuro-fuzzy interference systems and their application to nonlinear dynamical systems. "*Neural Networks Vol.12, pp.1301-1319*.
- [70] **Yen, G.G., Meesas, P.** (2001). "An effective neuro-fuzzy paradigm for machinery condition health monitoring. "*IEEE Transactions on Systems, Man and Cybernetics - Part B Vol.31 No.4, pp.523-536*.
- [71] **Mamdani, E. H.** (1977). "Application of fuzzy logic to approximate reasoning using linguistic synthesis. "*Transactions on Computers Vol.23 No.12, pp.1182-1191*.
- [72] **Sugeno, M.** (1985). "Industrial applications of fuzzy control . "*Elsevier Science Publishers*.
- [73] **Leardi, R., González, A. L.** (1998). "Genetic algorithms applied to feature selection in PLS regression: how and when to use them. "*Chemometrics and Intelligent Laboratory Systems Vol.41, pp.195-207*.
- [74] **Leardi, R.** The pls-genetic algorithm toolbox for Matlab. [Online] [Cited: 2011-05-24.] <http://www.models.life.ku.dk/GAPLS>.
- [75] **Wiggelinkhuizen, E. J., Rademakers, L. W. M. M., Verbruggen, T. W., Watson, S. J., Xiang, J., Giebel, G., Norton, E. J., Tiplucia, M. C., Christensen, A. J., Becker, E.** (2007). "Conmow final report. "*ECN-E--07-044*.
- [76] **Chandola, V., Banerjee, A., Kumar, V.** (2009). "Anomaly detection: A survey. "*ACM computing Surveys Vol. 41, No. 3, pp.15:1-15:58*.
- [77] **Hyers, R. W., McGrowan, J. G., Sullivan, K. L., Manwell, J. F., Syrett, B. C.** (2006). "Condition monitoring and prognosis of utility scale wind turbines. "*Energy Materials Vol.1 No.3, pp.187-203*.
- [78] **Rodríguez, R.V. J., Arkkio, A.** (2008). "Detection of stator winding fault in induction motor using fuzzy logic. "*Applied Soft Computing Vol.8, pp.1112-1120*.
- [79] **Wächter, M., Milan, P., Mücke, T., Peinke, J.** (2011). "Power performance of wind energy converters characterized as stochastic process: applications of the Langevin power curve. "*Wind Energy Vol. 14 pp.711-717*. Wind Energy Vol. 14 pp. 711-717.

-
- [80] **Üstüntaş, T., Duran Şahin, A.** (2008). "Wind turbine power curve estimation based on cluster center fuzzy logic modeling. *"Journal of Wind Engineering and Industrial Aerodynamics Vol. 96 pp.611-620.*
 - [81] **Stephen, B., Galloway, S.J., McMillan, D., Hill, D.C., Infield, D.G.** (2011). "A Copula model of wind turbine performance. *"IEEE Transactions on Power Systems Vol. 26 pp.965-966.*
 - [82] **Frank, E., Wang, Y., Inglis, S., Holmes, G., Witten, I.H.** (1998). "Using model trees for classification. *"Maschine Learning Vol. 32 No.1, pp. 63-76.*
 - [83] **Witten, I.H., Frank, E.** (1999). "Data mining: practical machine learning tools and techniques. *"ISBN: 978-1-558-60552-7.*
 - [84] **Kusiak, A., Wei, X.** (2011). "Prediction of methane production in wastewater treatment facility: a data-mining approach. *"Annals of Operations Research.*
 - [85] **Hansen, M.O.L.** (2008). "Aerodynamics of wind turbines. *"ISBN: 978-1-84407-438-9.*
 - [86] **International Electrotechnical Commission.** (2005). "Wind turbines - Part 12-1: Power performance measurements of electricity producing wind turbines. *"International Standard, IEC 61400-12-1.*
 - [87] **Chiu, S.L.** (1994). "Fuzzy model identification based on cluster estimation. *"Journal of Intelligent Fuzzy Systems Vol. 2 pp.267-278.*
 - [88] **Bishop, C.M.** (2009). "Pattern recognition and machine learning. *"ISBN: 978-0-387-31073.*
 - [89] **Schlechtingen, M., Santos, I.F., Achiche, S.** (2013). "Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description. *"Journal of Applied Soft Computing Vol.13, pp.259-270.*
 - [90] **Crabtree, C.J., Tavner, P.J.** (2011). "Condition monitoring algorithm suitable for wind turbine use. *"Renewable Power Generation (RPG 2011), Edinburgh.*
 - [91] **Sawalhi, N., Randall, R. B.** (2007). "Semi-automated bearing diagnostics - three case studies. *"Proceedings of the Comadem Conference, Faro, Portugal.*
 - [92] **Sawalhi, N., Randall, R. B., Endo, H.** (2007). "The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. *"Mechanical Systems and Signal Processing Vol.21, pp.2616-2633.*
 - [93] **Barszcz, T.** (2009). "Decomposition of vibration signals into deterministic and nondeterministic components and its capabilities of fault detection and identification. *"International Journal of Applied Mathematical Computation Vol.19-2 pp.327-335.*
 - [94] **Barszcz, T., Sawalhi, N.** (2012). "Fault detection enhancement in rolling element bearings using the minimum entropy deconvolution. *"Archives of Acoustics, Vol. 37-2, pp.131-141.*
 - [95] **Antoni, J., Randall, R. B.** (2004). "Unsupervised noise cancellation for vibration signals: part II—a novel frequency-domain algorithm. *"Mechanical Systems and Signal Processing Vol.18, pp.103-117.*
 - [96] **Niaoqing, H.,** (2012). "Enhanced fault detection of rolling element bearing based on cepstrum and stochastic resonance. *"Journal of Physics: Conference Series Vol.364, pp.1-8.*

- [97] **Dyer, D., Stewart, R.M.** (1978). "Detection of rolling element bearing defect by statistical vibration analysis. "*Journal of Mechanical Design Vol.100*, pp.229-235.
- [98] **Bonnardot, F., Randall, R. B., Antoni, J.** (2004). "Enhanced unsupervised noise cancellation using angular resampling for planetary bearing fault diagnosis. "*International Journal of Acoustics and Vibration Vol.9-2* pp.51-60.
- [99] **Wiggins, R. A.** (1978). "Minimum entropy deconvolution. "*Geoexploration Vol.16*, pp.21-35.
- [100] **McDonald, G.** *Minimum Entropy Deconvolution*. [Online] 2012-04. <http://www.mathworks.com/matlabcentral/fileexchange/29151-minimum-entropy-deconvolution-med-1d-and-2d/content/med2d.m>.
- [101] **Dwyer, R. F.** (1983). "Detection of non-Gaussian signals by frequency domain Kurtosis estimation. "*Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '83*, pp.607-610.
- [102] **Antoni, J., Randall, R. B.** (2006). "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. "*Mechanical Systems and Signal Processing Vol.20*, pp.308-331.
- [103] **Antoni, J.** (2006). "Fast computation of the kurtogram for the detection of transient faults. "*Mechanical Systems and Signal Processing Vol.21*, pp.108-124.
- [104] **Antoni, J.** *Programmes*. [Online] 04 2012. <http://www.utc.fr/~antoni/programm.htm>.
- [105] **Howard, I.** (1994). "A Review of rolling element bearing vibration "detection, diagnosis and prognosis". "*Aeronautical and Maritime Research Laboratory, DSTO-RR-0013*.
- [106] **Palshikar, G. K.** (2009). "Simple algorithms for peak detection in time-series. " *Proceedings of 1st IIMA International Conference on Advanced Data Analysis, Business Analytics and Intelligence, Ahmedabad, India*.
- [107] **Sawalhi, N., Randall, R. B.** (2008). "Helicopter gearbox bearing blind fault identification using a range of analysis techniques. "*Journal of Mechanical engineering Vol.5*, pp.157-168.
- [108] **Sawalhi, N., Randall, R.B.** (2008). "Novel signal processing techniques to aid bearing prognosis. "*IEEE PHM Conference, Denver, USA, 6-9. Oktober*.

7.3. Standards

- [109] **ISO 13373-1:2002 (2002)**. "Condition monitoring and diagnostics of machines – Vibration condition monitoring – Part 1: General procedures"
- [110] **ISO 13373-2:2005 (2005)**. "Condition monitoring and diagnostics of machines – Vibration condition monitoring – Part 2: Processing, analysis and presentation of vibration data"
- [111] **ISO 17359:2003 (2005)**. "Condition monitoring and diagnostics of machines – General guidelines"
- [112] **ISO 17359 (2007)**. "Condition monitoring and diagnostics of machines – Supplement 1: Explanations of technical terms"