Tracht, Kirsten ; Goch, Gert ; Schuh, Peter ; Sorg, Michael ; Westerkamp, Jan F.

Failure probability prediction based on condition monitoring data of wind energy systems for spare parts supply

Journal Article as: peer-reviewed accepted version (Postprint)

DOI of this document*(secondary publication): https://doi.org/10.26092/elib/3196

Publication date of this document: 10/08/2024

* for better findability or for reliable citation

Recommended Citation (primary publication/Version of Record) incl. DOI:

Kirsten Tracht, Gert Goch, Peter Schuh, Michael Sorg, Jan F. Westerkamp,
Failure probability prediction based on condition monitoring data of wind energy systems for spare parts supply,
CIRP Annals, Volume 62, Issue 1, 2013, Pages 127-130, ISSN 0007-8506,
https://doi.org/10.1016/j.cirp.2013.03.130.

Failure probability prediction based on condition monitoring data of wind energy systems for spare parts supply

Kirsten Tracht [a], Gert Goch (1)[b,*], Peter Schuh [a], Michael Sorg [b], Jan F. Westerkamp [b]

[a] Bremen Institute for Mechanical Engineering (bime), University of Bremen, Badgasteiner Straße 1, 28359 Bremen, Germany
[b] Bremen Institute for Metrology, Automation and Quality Science (BIMAQ), University of Bremen, Linzer Straße 13, 28359 Bremen, Germany

Keywords:
Maintenance
Predictive model
Reliability

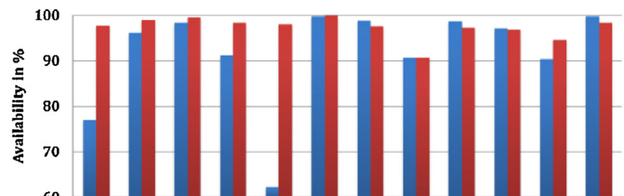
1. Introduction

Spare parts availability is essential for efficient maintenance repair and overhaul processes. These are necessary to ensure an economic machine operation. Long lead times of spare parts lead to the necessity of stock keeping, which ties a lot of capital because of high acquisition costs. Spare parts stocking is based on demand forecasting that possess high potential in reducing the amount of spare parts in stock. Even if online condition monitoring systems are installed in complex technical systems like wind energy systems (WES, Table 1), condition monitoring information is barely used to predict spare parts demand. The varying loads on wind energy system components and technically different system concepts of wind energy systems in general (e.g. rotational characteristics, load regulation, or generator type), varying operating conditions regarding the WES location, and components constructed the same way but from different suppliers, result in varying survival times of units and wide scattered results of the failure analysis [1]. In order to extract usable information about the failure probability for specific components, operational data, event, failure and damage descriptions have to be comprised and analyzed systematically. Therefore, an enhanced forecast model that considers condition information, has been developed. The example in Fig. 1 illustrates the availability of the research WES of the University of Bremen in its first year of operation.

For example, during May the elevator rope system failed and influenced the total availability of the wind turbine due to a repair cycle lasting 10 days. Fig. 1 highlights how this single event

Table 1
Acronyms and abbreviations.

Abbreviation	Augmentation
α	Shape parameter of Weibull distribution
β	Regression coefficient
CMS	Condition monitoring system
g	Density function of binomial distribution
$h_0(t)$	Baseline hazard function
$h_i(t)$	Hazard function
k	Number of spare part demands
λ	Scale parameter of Weibull distribution
n	Number of units
p	Failure probability
PHM	Proportional hazards model
$p(t)$	Density function of Weibull distribution
SCADA	Supervisory control and data acquisition
t	Time
T	Survival time of Weibull distribution
Temp	Temperature
WES	Wind energy systems
WT	Wind turbine
x	Covariate
CBM	Condition based maintenance
WONDER	Wind farm management system (brand name)
LWK	Chamber of agriculture Schleswig–Holstein
WMEP	Scientific measuring- and evaluation programme



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
availability	76,9	96,2	98,3	91,1	62,2	99,7	98,8	90,7	98,6	97,0	90,4	99,7
power/wind-availability	97,8	98,9	99,5	98,3	98,1	99,8	97,5	90,6	97,3	96,7	94,6	98,3

Fig. 1. Monthly availability (blue) and power output availability (red) of the research WES of the University of Bremen in (1st year of operation).

* Corresponding author.

reduced the availability of the research WES significantly in contrast to the "power/wind-availability".

2. State of the art

2.1. Preventive and corrective maintenance processes

The two most prevalently applied maintenance strategies are corrective and preventive maintenance processes. Corrective maintenance is carried out unscheduled in case of component failures or if faults are detected in WES components during the recurring inspection [2]. It is the most expensive strategy and operators strive for minimizing the number of these events, because of a high risk of unavailable spare parts and prolonged downtimes, caused by conditions that prohibit maintenance activities. By contrast preventive maintenance aims at repairing or replacing components before they fail. This can be achieved by scheduled maintenance activities, also known as time based (or planned) maintenance, which involves repair or unit replacements at regular time intervals, as recommended by the supplier, and regardless of its condition. Time based maintenance reveals the possibility of planning maintenance resources and the instant of maintenance [3], thus minimizing downtime. This advantage is contrary to the drawback that unnecessary frequent maintenance tasks increase the maintenance cost, because the lifespan of units is not entirely utilized. An alternative to preventive and corrective maintenance is the condition based maintenance strategy, in which specific components are monitored and maintenance tasks are determined ahead of failures [4]. Today, maintenance technicians manually perform failure detection with the help of condition monitoring systems (CMS).

2.2. Condition monitoring

Wind farm management or supervisory control and data acquisition (SCADA) systems acquire condition monitoring, as well as operation data. For example, the wind farm management system WONDER by Deutsche WindGuard records 10-min mean values of operation and condition data of a WES and transfers them to a data acquisition server. Data of that system have also been processed to validate the approach presented in this paper. Current and new emerging maintenance strategies for WESs depend on condition parameters and measurements. Those are either supplied by component specific CMS (designed e.g. for gearboxes or bearings) or manufacturer related, plant wide CMS (e.g. GE or Nordex).

For the maintenance of WES, emphasis is put on the gearbox and the main bearing as a direct consequence of the long machine downtimes caused by failures of these components (Fig. 2).

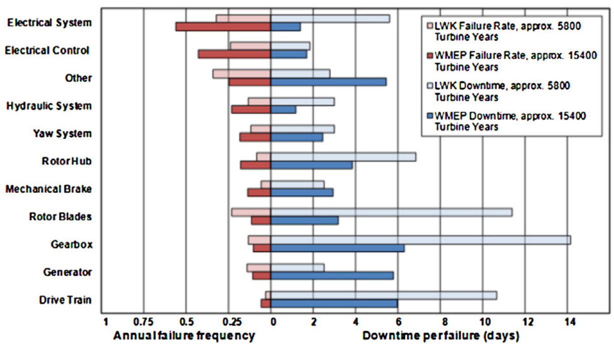


Fig. 2. Failure rate and downtime per failure of WES units for two surveys including over 20,000 turbine years of data as published in [5].

Today, parameters like oil, gearbox or bearing temperature, power output, wind speed, wind directions as well as vibrations are monitored online. The majority of available CMS focus on vibration characteristics [6] as defined by DIN ISO 10816. However, physical models are not available for failure forecasting because of complex interactions between WES components and the superposition of signals.

Despite the large amount of data, condition monitoring information is not used systematically to predict failures as well as spare parts demands and manual inspection of data becomes impractical with the increasing number of WES per operator.

SCADA data on the other hand are readily available [7] and systems like WONDER collect and store large amounts of data, which give indications about the WES status. At present, SCADA-systems are the most cost effective way of implementing a CMS [8]. Kusiak and Verma show that component failures can be predicted 5 to 60 min in advance [9,10]. This short period can be used to prevent further damages on the WES, but it is not suitable for a reasonable demand forecasting.

2.3. Demand forecasting of spare parts

Spare parts demand forecasting requires failure forecasting of units. It is either performed on the basis of historical data or based on hazard functions [11]. In case of demand prediction by means of historical data, time series analysis approaches, such as Crostons method is applied to predict intermittent and lumpy spare part demands. Crostons method has been modified by Syntetos, who hereby achieved the lowest forecast-error in demand prediction, compared to other well known time series analysis methods, like exponential smoothing or moving average [12]. These algorithms need a very large amount of historical data, which do not exist within the comparably young WES industry. Historical data are missing due to short innovation cycles of units, high WES growth rates and a lack of profound data recordings. Furthermore, condition monitoring data and characteristics of maintenance processes, applied for different machines, as well as wear or aging processes, are only considered indirectly in time series based approaches. Hence, observations of changing values of these parameters cannot be implemented in these methods.

In contrast to this, spare parts demand prediction with the help of hazard functions offers the opportunity of considering varying stress or machine loads. Lanza, for example, implemented a shape parameter into the Weibull distribution that varies with the machine load. Thereby, the author is able to consider different operating modes of machine tools [13]. More specific details of operating modes or condition monitoring data, like temperature values or oil conditions cannot be integrated into the approach.

Oil conditions have been implemented by Louit [14]. The model proposed there is a single unit system that investigates the impact of age and oil condition on the remaining useful lifetime of a unit with the help of a proportional hazards model (PHM). Consideration of external influences is not possible in his approach. The PHM has been proposed by Cox in 1972 and is capable of integrating factors influencing the survival time [15]. Originally the PHM has been applied in the field of biology to investigate the impact of various medical treatments. A comprehensive literature review about PHM applications is presented by Kumar and Klefsjö [16].

Ghodrati showed that the PHM can be used in technical applications, but neglected time dependent variables. The author applied the algorithm to predict spare part demands of mining machines by implementing external influences, like operating conditions and operator behaviour [17].

One single time dependent internal variable has been considered by Louit (oil condition), so it is not possible to utilize comprehensive operational and online condition monitoring data, which are available in most technical systems, today. Therefore, SCADA data are investigated in this paper in order to predict failures of critical units more accurately. In contrast to existing applications, internal as well as external time dependent influences will be integrated into an enhanced forecast model. Data are processed and applicability of time dependent SCADA data to a PHM is proven. The approach proposed is verified with SCADA data of WES. This ensures that failures of units are predicted depending on machine conditions, on stress states and on the operating environment, utilizing data that are monitored in nearly all technical systems, but not used for failure prediction, as yet.

3. Enhanced forecast model

3.1. Model definition

For integrating online condition monitoring and operation data into the enhanced failure forecast model data mining, the PHM and a binomial distribution have been applied in this paper. In general, the PHM is suitable if it is assumed that covariates (influencing parameters) induce a proportional change in the baseline hazard function. The forecast model and its coherences proposed in this paper are summarized in Fig. 3.

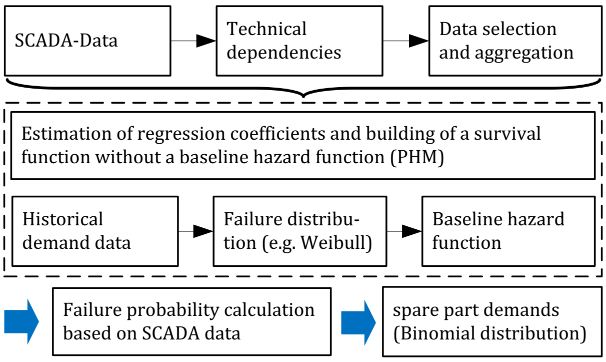


Fig. 3. Enhanced forecast model.

Within the enhanced forecast model, parameters influencing the degradation of the considered unit are derived from the SCADA-system. Parameter identification is realized by means of technical coherences as well as expert knowledge. Based on these parameters data mining is conducted with time dependent SCADA data for machine status monitoring. For example, in the past, maintenance technicians observed peaks in the value of the generator temperature sensor when the degradation of bearings advances. Hence, generator temperature is integrated as a condition parameter in the model.

Within the model, any baseline hazard function (h_0) can be combined with covariates x_i (Eq. (1)). Their importance is represented by regression coefficients (β). If all covariates equal zero, the baseline hazard function describes the risk of failure.

$$h_i(t) = h_0(t)\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in}) \tag{1}$$

Often, lifetime of units is not known, because the failure event has not happened at the instant of investigation. The reason for that could be an observation period that is too short. Lifetime data that are incomplete in terms of end of life are referred to as right censored data, which are implemented into the PHM with the help of the partial likelihood estimator (Eq. (2)).

For building the PHM (Eq. (1)), regression coefficients (β) are calculated for the covariates that influence survival time, whereas significance of covariates is tested later in the model. As proposed by Cox, calculation of regression coefficients is performed with the partial likelihood estimation, without assuming a baseline hazard function (h_0) [15]. For that purpose Eq. (2) is maximized by means of Newton's algorithm.

$$L(\beta) = \prod_{i=1}^{m} \frac{\exp(\beta x_i)}{\sum_{t(j) \geq t_i} \exp(\beta x_i)} \tag{2}$$

In the numerator of Eq. (2), only data from units that already failed are inserted. The denominator also includes operative units. For every covariate (e.g. wind speed), Eq. (2) calculates a regression coefficient.

$$\hat{G} = 2[\log(L(\hat{\beta})) - \log(L(0))] \tag{3}$$

Parameters identified for calculating failure probability of a unit in the first step of the model are tested regarding their statistical significance. This requires hazard functions, which are modelled with individual, all or combinations of covariates and computed with the likelihood ratio test [18]. The log-likelihood value of the null model, which is the model without covariates, is compared to the log-likelihood value of every other model (Eq. (3)). The result of Eq. (3) is inserted into the χ^2 distribution for evaluation of the p-value. The smaller the p-value is, the better the covariate or a set of covariates describe the risk of failure. Only the set of covariates with the smallest p-values are selected for failure probability prediction in further steps of the model.

After validating significance of covariates, the whole model is validated in terms of adequately describing the failure data. This validation process utilizes the graphical diagnosis proposed by Schoenfeld [19]. Schoenfeld residuals should vary randomly around zero, without shifting systematically. Otherwise, the model assumption is violated.

For obtaining the hazard rate of the unit considered, an appropriate failure distribution is chosen dependent on the failure process of the unit. Failure probability of electronic devices is calculated by means of the exponential distribution, for example. If the unit investigated is exposed to mechanical wear, the hazard function is calculated with the Weibull distribution [11]. Its density distribution is presented in Eq. (4). The scale parameter (λ) and the shape parameter (α) are calculated as proposed by the standard (DIN EN 61649:2008), whereas λ equals $1/T$ [20]. λ is calculated with historical demand data.

$$p(t) = \frac{\alpha}{T}\left(\frac{t}{T}\right)^{(\alpha-1)} e^{-(t/T)^{\alpha}} \tag{4}$$

The failure probability of a specific unit regarding its current condition and operation information is then inserted into a Binomial distribution. It calculates discrete demand events during lead time of spare parts. The density function of the binomial distribution in Eq. (5) estimates the number of exactly k events, if a failure probability of p exists, whereas n represents the amount of operating units.

$$g(k; n; p) = \binom{n}{k} p^k (1-p)^{n-k} \tag{5}$$

By combining new and established algorithms in the enhanced forecast model SCADA-data are utilized for calculating spare part demands depending on the units' condition.

3.2. Scenario description

For the validation of the enhanced forecast model, eleven units, twelve spare part demands (generator bearing) and SCADA data are analyzed in this paper. The data originate from 19 onshore WES, which have been put into operation in 2001. From 2006 to 2012 comprehensive SCADA data were recorded. All wind turbines are of the same machine type, but SCADA data are different in some machines of the wind farm. Hence, only data recorded completely were tested regarding their applicability for demand prediction of generator bearings (Table 2). Measurement of temperature within the WES nacelle is carried out at the spots highlighted in Fig. 4.

Table 2
Condition monitoring data of investigated wind turbines.

Measurement parameters	Completely recorded
Temperature of main bearing	
Temperature of stator	X
Temperature of generator	X
Wind speed	X
Outside temperature	
Wind direction	X
Nominal power output	X
Temperature of bearing A and B	

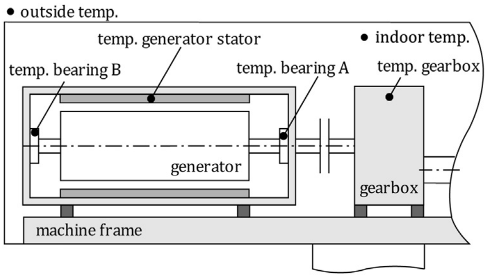


Fig. 4. Wind turbine and temperature measurement spots.

The bearing temperature has been investigated on the basis of the generator temperature, because the temperature of the bearing is not monitored in all wind turbines. This is reasonable, because comparisons of the two values measured showed that they vary proportionally according to operation conditions. For validation, this temperature comparison has been conducted for a time span of 5 years.

For the analysis of time dependent SCADA-data, exceedances of threshold values have been counted and summarized for a time span of one week. For example, exceedance of the temperature value threshold 100 °C of the generator has been counted. Thereby, no information about the machine status is lost. These prepared data are fed into the enhanced forecast model and experiments have been conducted with different reasonable parameter combinations.

4. Results

Computing operations with SCADA data and the statistical software package R showed that it is possible to process comprehensive SCADA and CMS data. The best result (lowest p-value of the χ^2-test) has been obtained with the PHM considering lifetime and temperature (Table 3). Investigations showed that the p value did not improve significantly, when more parameters were taken into account, like power output or wind force. Hence, it is not necessary to include more SCADA-data for failure prediction in the scenario investigated.

Table 3
Result of the investigations.

Parameter	Value
PHM with temperature (p-value)	0.058
PHM with temperature and power output (p-value)	0.251
PHM with temperature and wind force (p-value)	0.419
β of covariate temperature in PHM	0.023
Shape parameter of baseline hazard function (α)	1.339
Scale parameter of baseline hazard function (λ)	1/43.156

Though the sample size of demand events is low (12 demands), Schoenfeld residuals did not show any abnormality, which can be regarded as a first feasibility check of the proposed approach.

The hazard rate obtained with the enhanced forecast model (PHM(h)) is shifted in comparison to the baseline hazard rate (Fig. 5). This shifting is caused by operating conditions, recorded by the SCADA-System (80 temperature exceedances), influencing the predicted survival time and the number of spare part demands during lead time (10 weeks). Neglecting machine conditions, seven demands are calculated with the baseline hazard function, leading to understocking. In contrast to this, the enhanced forecast model considered SCADA-data, which represent the units' current condition, and predicted ten spare part demands. Thereby, the model allows for the more accurate estimation of inventory levels in spare parts planning without venturing machine availability and presents the basis for reducing maintenance costs.

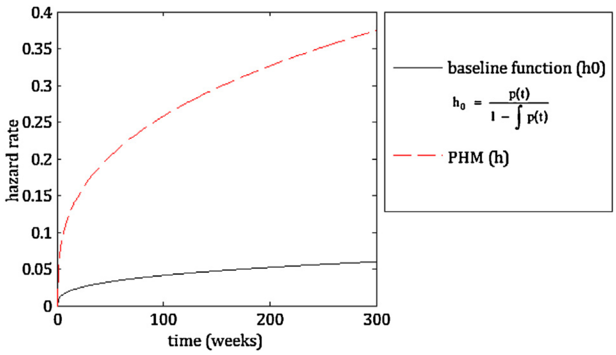


Fig. 5. PHM and baseline hazard rate of generator bearing.

5. Summary

Within this paper an approach utilizing SCADA data for spare parts prediction of technical systems has been presented. The method proposed has been tested with data of WES to predict failures of units that are exposed to mechanical wear. This is achieved by means of an enhanced forecast model, basing on a PHM, capable of considering time dependent covariates, and a binomial distribution. Information about the current temperature of a generator bearing and its age are implemented in the model. The model is also capable of considering external influences, but within the scenario no significant influence on survival time of the bearing has been observed. Even if 5,700,000 datasets of 6 years of operation have been investigated, only 12 units of the same type have been exchanged. Despite the small sample size, the model is able to calculated appropriate results. It can be used for any unit that has condition monitoring information available. Further investigations on other WES units will be conducted to show how, for example, failure prediction of electronic devices can profit from SCADA-information.

References

[1] Goch G, Knapp W, Härtig F (2012) Precision Engineering for Wind Energy Systems. *CIRP Annals – Manufacturing Technology* 61:611–634.
[2] García Márquez FP, Tobias AM, Pinar Pérez JM, Papaelias M (2012) Condition Monitoring of Wind Turbines: Techniques and Methods. *Renewable Energy* 46:169–178.
[3] Tracht K, Schuh P (2012) Multi Criteria Decision Making for Maintenance Planning of Wind Turbines. *1st International Conference on Through-life Engineering Services*, Shrivenham, 175–181.
[4] Pedregal DJ, García FP, Roberts C (2009) An Algorithmic Approach for Maintenance Management Based on Advanced State Space Systems and Harmonic Regressions. *Annals of Operations Research* 166(1):109–124.
[5] Feng Y, Qiu Y, Crabtree C, Tavner P (2010) Detecting WT Gearbox Failures. *European Wind Energy Conference 2010 (EWEC)*, Warsaw.
[6] Randall RB (2011) *Vibration-Based Condition Monitoring*, Wiley, Chichester, West Sussex, UK.
[7] Schlechtingen M, Santos IF, Achiche S (2013) Wind Turbine Condition Monitoring Based on SCADA Data Using Normal Behavior Models. Part 1: System Description. *Applied Soft Computing* 13(1):259–270.
[8] Yang W, Jiang J (2011) Wind Turbine Condition Monitoring and Reliability Analysis by SCADA Information. *IEEE* 1872–1875.
[9] Kusiak A, Wenyan L (2011) The Prediction and Diagnosis of Wind Turbine Faults. *Renewable Energy* 36:16–23.
[10] Kusiak A, Verma A (2011) A Data-Driven Approach for Monitoring Blade Pitch Faults in Wind Turbines. *IEEE Transactions on Sustainable Energy* 2(1):87–96.
[11] Abernethy RB (2006) *The New Weibull Handbook*, 5 ed. Robert B. Abernethy Publishing.
[12] Syntetos AA (2001) On the Bias of Intermittent Demand Estimates. *International Journal of Production Economics* 71(1–3):457–466.
[13] Lanza G, Niggeschmidt S, Werner P (2009) Optimization of Preventive Maintenance and Spare Part Provision for Machine Tools Based on Variable Operational Conditions. *CIRP Annals – Manufacturing Technology* 58: 429–432.
[14] Louit D, et al (2011) Condition-Based Spares Ordering for Critical Components. *Mechanical Systems and Signal Processing* 25:1837–1848.
[15] Cox DR (1972) Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B* 34(2):187–220.
[16] Kumar D, Klefsjö B (1994) Proportional Hazards Model: A Review. *Reliability Engineering and System Safety* 44(2):177–188.
[17] Ghodrati, B., 2005, Reliability and Operational Environment Based Spare Parts Planning. Lulea University of Technology, PhD. Thesis.
[18] Sachs L, Hedderich J (2009) *Angewandte Statistik, Methodensammlung mit R*, Springer Verlag, Berlin.
[19] Schoenfeld D (1982) Partial Residuals for the Proportional Hazards Regression Model. *Journal of Biometrika* 69(1):239–241.
[20] DIN EN 61649:2008 (2009) *Weibull Analyse*, Beuth Verlag GmbH, Berlin.