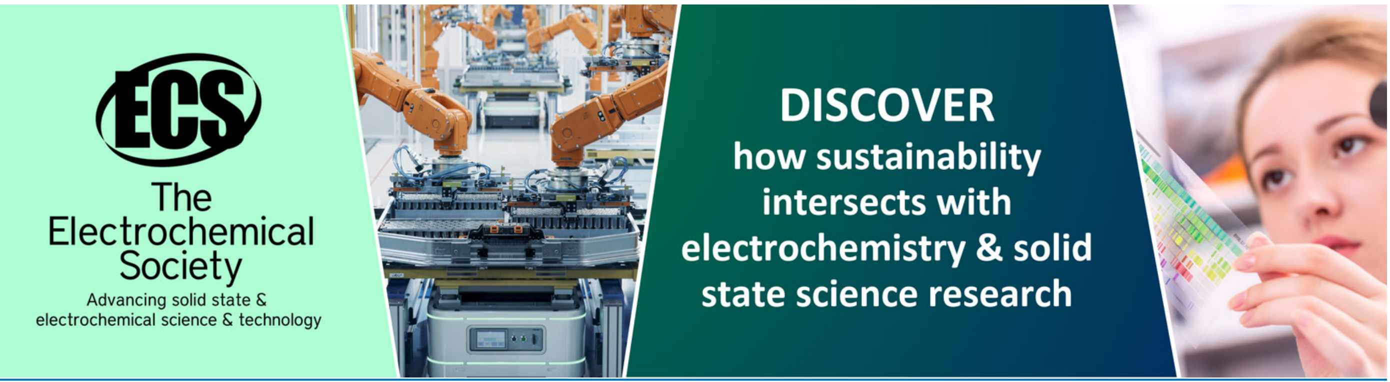
Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

A Novel Probabilistic Long-Term Fault Prediction Framework Beyond SCADA Data - With Applications in Main Bearing Failure

To cite this article: Jürgen Herp *et al* 2019 *J. Phys.: Conf. Ser.* **1222** 012043



IOP Conf. Series: Journal of Physics: Conf. Series **1222** (2019) 012043 doi:10.1088/1742-6596/1222/1/012043

A Novel Probabilistic Long-Term Fault Prediction Framework Beyond SCADA Data - With Applications in Main Bearing Failure

Jürgen Herp

Applied Statistical Processing Group (πSeG), Embodied Systems for Robotics and Learning (ESRL), Maersk Mc-Kinney Moeller Institute, University of Southern Denmark, 55 Campusvej, 5230 Odense, Denmark

E-mail: herp@mmmi.sdu.dk

Niels L. Pedersen

Diagnostics, Siemens Gamesa Renewable Energy, Borupvej 16, 7330 Brande, Denmark

E-mail: niels.l.pedersen@siemensgamesa.com

Esmaeil S. Nadimi

Applied Statistical Processing Group (πSeG), Embodied Systems for Robotics and Learning (ESRL), Maersk Mc-Kinney Moeller Institute, University of Southern Denmark, 55 Campusvej, 5230 Odense, Denmark

E-mail: esi@mmmi.sdu.dk

Abstract. Prognostic and health monitoring addresses the issue of detecting faults and monitoring the current state of a wind turbine. Details about the fault's progression, and from there, the remaining useful lifetime, are key features in monitoring and industrial operation and maintenance planning. In order to avoid increase in operation and maintenance cost, as well as subjective human involvement, we present an online and automated monitoring framework for prediction of the remaining useful lifetime based on deep learning models. This framework includes training and re-training of predictive models with minimal oversight by the operators.

Further, we explore the dependency of various models' predictive abilities based on the input variables available, such as SCADA and secondary measurements. Especially deep learning approaches, such as neural networks, benefit greatly from the volume of data that can be extracted from modern-day turbines. This work utilizes upon the volume of data to present a case study on main bearing failures for 108 turbines. In the presented setting, predictions of the remaining useful lifetime of more than 90 days can be expected on average, outperforming the closest state-of-the-art estimate by almost a factor of two on average.

1. Introduction

Accurate asset health assessment, e.g. Remaining Useful Lifetime (RUL) estimations, is an essential part of industrial operation and maintenance strategies (O&M), be it for increased productivity and/or reduced O&M costs. As wind power constitutes an substantial part of renewable energies, availability, reliability and lifetimes are taken more and more into consideration by business economics to carefully handle O&M costs. Blanco [1], Bussel et al. [2], Wilkinson et al. [3], and Walford [4] describe the impact of the available O&M cost during operation of a fleet of wind turbines. Further, the independent identify the need to minimize such cost as a business case. Nielsen et al. [5] and Petersen et al. [6, 7] have identified areas within O&M tasks that can improove the maintenance of offshore wind turbines. In this context, Petersen et al. [7] has proposed a Lean approach to modularization of O&M tasks and resources. Facilitating the need of having an oversight of a physical system, condition monitoring have been employed in many fields, and are still subject to research the day to day, including bearing monitoring, as many publications over the resent years bare witness: Singleton et al. [8] propose both an experimental and computational approach for the remaining lifetime of bearing by taking the relationship between current discharge events and the vibration signal into consideration in their model. Herp et al. [9] describes a Bayesian approach to handling bearing model residuals as generated by a model proposed by Bach-Andersen et al. [10], this model includes temperature measurements. Another temperature based approach is presented by Kusiak et al. [11]. Other statistical approaches are given by Loutas et al. [12] and Li et al. [13]. Neural network based models and frequency spectral models are covered by Deutsch et al. [14], Ali et al. [15], Saruhan et al. [16], Elasha et al. [17], and Teng et al. [18]. Later on, the contribution of Teng et al. [18], and Herp et al. [9] are highlighted further in Section 2.

While in early years a human operate was needed to make sense of recorded signals, this practice has gradually made way to more and more automated systems. This is also the case for wind turbines; especially nowadays where the volume of data is shear unlimited and new machine learning algorithms outperform the human ability of identifying patterns.

In the present *Big Data* era, new statistical methods are needed to describe and learn from these large dimensional data sets. This work utilizes upon the volume of data by applying *Deep Learning* and statistical driven models to present a framework in which wind turbine can be monitored in on online manner, with minimum oversight by an operator. Main bearing failure are presented as case studies to illustrate the performance of the proposed framework. It will be elucidated how the proposed framework outperforms existing long-term fault prediction frameworks (Si et al. [19, 20], Herp et al. [9, 21], Teng et al. [18]), while archiving prediction horizons beyond 2 month.

Even though, much work has been done on bearing vibrations to determine bearing failure, less work is found [10], to the knowledge of the author, on the use or combination of other sensory inputs. This work will go beyond SCADA (Supervisory Control and Data Acquisition) data to include other relevant features in the predictive considerations, namely band energies in vibration spectra.

The aim of this work is twofold: (i) showing what potential lies within large volumes of data, and (ii) provide a flexible framework for wind farm operators to use in their condition monitoring and O&M efforts. It shows insight in the working of *Big Data* analytics for system monitoring and enables the interpret of the results in the proposed framework, i.e. understand the connection between a recurrent neural network, including their sequential training, and its connection to probabilistic distributions of the RUL of a turbine. Ultimately pointing out the proposed framework as a generalization of fault prediction and RUL estimation beyond the presented main baring case studies.

The paper is organized as follows: The methodology is presented in Sections 2 and 3, providing the necessary definitions on the predictive models and framework. Section 4 shows the implementation of the proposed model in a use-case of wind turbines, evaluating the model performance and comparing it to the state-of-the-art. Finally the work is concluded in Section 5.

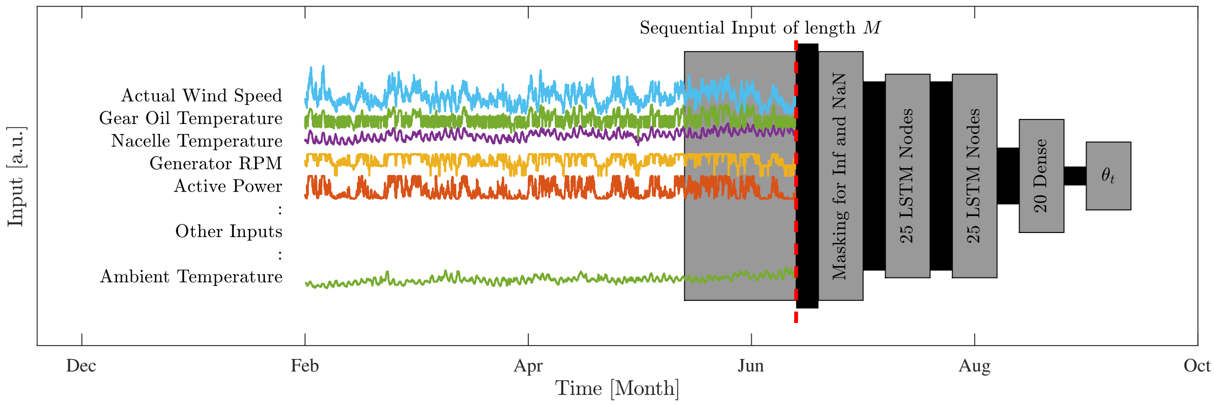


Figure 1: Illustration of the RNN used to predict the RUL. Selected timeseries shown as input reference.

2. Predictive Methodologies

Most wind turbine are equipped with SCADA systems (sampling 10 min averages), in addition thereto other means of measurements can also be available, one which this study is concerned with are the energy in selected bands of vibration Fourier spectra, other measures are preprocessed data, such as model temperature residuals as described by Bach-Andersen et al. [10]. These measurements make up a process $\{\mathbf{x}_t\}$, where t is a time instance, and $\mathbf{x}_t \in \mathbb{R}^m$ is a sample vector containing $m \leq 185$ features. This study comprises 108 turbines.

For the proposed framework models mapping from $\{\mathbf{x}_t\}$ to the RUL is denoted $M : \mathbf{X} \to \boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is a measure for the RUL. These models are:

W Weibull Model for RUL [22]: As part of the proposed framework, this model was developed to cater the need of a highly flexible model with regards to the input. As the model is based on a Recurrent Neural Network (RNN) the input is limited to the number of features, $m \leq 185$. The model predicts the RUL as parametrization of the Weibull probability distribution:

$$\mathbb{P}(t) = \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^{\alpha}}, \tag{1}$$

where the back-end of the model maximizes the likelihood of prediction by solving

$$\arg\max_{\boldsymbol{\omega}} \log \mathcal{L}(\boldsymbol{\omega}, \mathrm{RUL}, \Delta, \mathbf{x}_{[1,t]}) = \sum_{i=1}^{t}\left(\Delta_i\left[\alpha_i \log\left(\frac{\mathrm{RUL}}{\beta_i}\right) + \log(\alpha_i) - \log(\mathrm{RUL})\right] - \left(\frac{\mathrm{RUL}}{\beta_i}\right)^{\alpha_i}\right) \tag{2}$$

for the topology of the RNN with weights $\boldsymbol{\omega}$. The underlying network topology is illustrated in Figure 1.

BM Brownian Motion Model for RUL [19, 20]: The model will assume that the bearing temperature is a nonlinear drift model driven by an underlying Brownian motion:

$$X(t) = x_0 + \lambda \int_0^t \mu(\tau; \mathcal{V})d\tau + \sigma_{\mathrm{B}} B(t), \tag{3}$$

where $B(t)$ is the driving Brownian motion with nonlinear drift $\lambda\mu(t;\mathcal{V})$. Si et al. then compare Eq. (3) to a threshold, w, in order to estimate the remaining useful lifetime. The time for crossing a threshold w is be defined as:

$$T = \inf\{t : X(t) \geq w \mid X(0) < w\}, \tag{4}$$

As result, Si et al. derives a distribution for the remaining useful lifetime, referred to as Brownian Motion for Remaining Useful Lifetime (BM-RUL):

$$\mathbb{P}_{L_t|\mathbf{x}_{[0,t]}}(l_t \mid \mathbf{x}_{[0,t]}) \cong \frac{w_t\Lambda(l_t) - \alpha_t(l_t;\mathcal{V})\Delta(l_t)}{\sqrt{2\pi l_t^2 \Lambda^3(l_i)}} e^{-\frac{(w_t - \hat{\lambda}_t v(l_t))^2}{2\Lambda(l_t)}}, \tag{5}$$

where L_t is defined as the remaining useful lifetime, l_t is the residual measure corresponding to the remaining useful lifetime at t, $\omega_t = \omega - x_t$, $\hat{\lambda}_t = \mathbb{E}[\lambda_t \mid \mathbf{x}_{[0,t]}]$, $\Lambda(l_t)$ is associated with the variance of the process, $\alpha_t(l_t;\mathcal{V}) = v(l_t) - l_t\mu(l_t + t;\mathcal{V})$, $\Delta(l_t)$ is associated with the variance of the process and its drift, and $v(l_t) = \int_t^{l_t+t} \mu(\tau;\mathcal{V})d\tau$. See Si et al. [19, 20] for further details.

VS Vibration Spectral Model for RUL [18]: An artificial neural network to predict short-term tendencies of vibration energy bands. By combining the predicted and training features, a polynomial curve reflecting the long-term degradation process of bearings is fitted. Through solving the intersection between the fitted curve and a threshold, the RUL can be obtained.

GP Gaussian Process Model for RUL [9]: This is a statistical approach to abstract and predict turbine states in an online manner. The approach is based on the separability of the sufficient statistics and a hidden variable. By assuming that the feature space can be described by a multivariate Gaussian distribution, the prediction of the RUL is treated as a Gaussian process over the feature space. Any input can be provided to the model (but scales poorly with increasing input size), returning a non tractable probability distribution of RUL.

A list of the possible combination of inputs to each model is provided in Table 1, together with the model output. As seen in Table 1 the Weibull Model for RUL all necessary properties for all single input features and combinations thereof, as well as it provides full properties for the output space. In contrast, the Brownian Motion Model only provides properties for the single input feature space, but in practice is limited to selected SCADA input, mainly temperature, and requires careful initialization and threshold selection for vibration data. The output range is by construction limited to tractable probability distributions. For the Vibration Spectra Model the point-wise output is highly dependent on the chosen threshold. Due to poor scalability of the algorithm the Gaussian Process Model is limited to a small number of input features.

3. Online Monitoring and Training Framework

In this work an online monitoring framework is developed and illustrated by the flowchart in Figure 2 with **W** as an example of a model. In detail, the proposed approach starts with detecting a change in the turbine prior to its failure, this is referred to as opening a case. The turbine credential are then stored in a Library and mored as unprocessed. Following system will then acquire the relevant data for the monitoring based on **W**. For each t, since the detection, the RUL is estimated. As long as the turbine still is operating, for each consecutive times-stamp, $t = t + 1$, the RUL can still be estimated as a probability distribution. This is illustrated in Figure 3a. The monitoring system provides the operator with a graphical representation of the RUL distribution at selected monitoring points. Further, for each t the first moment, median,

Table 1: Model comparison including model features.

Study	Input						Output			Avg. max(RUL) [days] over all turbines rounded to nearest days					
	SCADA	Vibration	Residuals	SCADA+Vibration	SCADA+Residuals	SCADA+Residuals+Vibration	Point-wise	Probability Distribution	Tractable	SCADA	Vibration	Residuals	SCADA+Vibration	SCADA+Residuals	SCADA+Residuals+Vibration
Weibull Model for RUL (**W**)*	●	◐	●	◐	●	◐	●	●	●	94	25	-	86	93	32
+Dense Layer*	●	◐	●	◐	●	◐	●	●	●	99	21	-	82	80	-
+LSTM Layer*	●	◐	●	◐	●	◐	●	●	●	95	19	-	94	98	-
+LSTM Layer and Dense Layer*	●	◐	●	◐	●	◐	●	●	●	64	22	-	82	70	-
Brownian Motion Model for RUL (**BM**)†	◐	◐	●	-	-	-	-	●	●	20	42	-	-	-	-
+Study State*	◐	◐	●	-	-	-	-	●	●	32	12	-	-	-	-
Vibration Spectra Model for RUL (**VS**)†	-	●	-	-	-	-	◐	-	-	-	54	-	-	-	-
Gaussian Process Model for RUL (**GP**)†	◐	◐	◐	◐	◐	◐	-	◐	-	15	8	33	-	-	-

● = provides property; ◐ = partially provides property; - = does not provide property;
†model has been published; *model has not been published yet

and maximum likelihood is provided for the operator to provide point-wise measures of the RUL, since case opening. When the turbines operation is terminated, either by the operator or a fault, Figure 3c is generated, comparing the RUL estimations against the real remaining lifetime.

If the error of the prediction is small compared to a threshold E_0, the case is closed. Else, the predictive model is updated, and the error re-checked. At any point the models residual error, over a distribution of the RUL given empirical evidence, $\mathbb{P}(\text{RUL} \mid \mathbf{x}_{[1,t]})$, is defined as:

$$E(t \mid \mathbf{x}_{[1,t]}) \propto \int_t^\infty (\text{RUL} - \hat{\text{RUL}})^2 \mathbb{P}(\text{RUL} \mid \mathbf{x}_{[1,t]}) d\text{RUL}. \tag{6}$$

The threshold E_0 is then defined as E such that the confidence of prediction is 90%, i.e. with respect to a normal distribution with mean, $\mu = \hat{\text{RUL}}$ and spread, $\sigma = 0.05\hat{\text{RUL}}$. If the error is lower than the threshold, the model is updated, else, user input is required to evaluate the case. The model update steps requires all unprocessed turbines and acquires their data. In the case of the Weibull model, the framework would would load the current RNN topology and weights, and starts the training process with the new data. At the end of the update a new model is formed and send forward to be used in the monitoring efforts.

4. Case Studies

For selected turbines the online predictions are shown in Figure 4a and 4c, including the real remaining lifetime for reference. These turbine's prediction likewise follow the real remaining lifetime.

In order to summarize the outcome, the proposed frame work will be evaluated in comparison to the above presented state-of-the-art RUL prediction approaches, **BM**, **VS**, and **GP**. For the

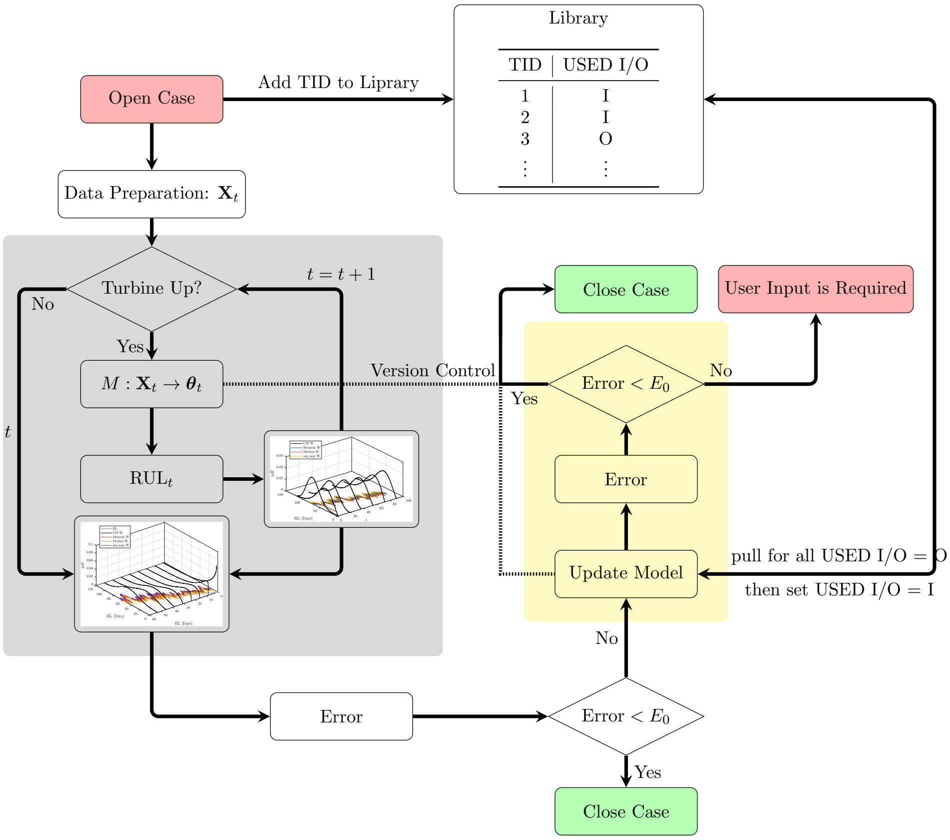


Figure 2: Flowchart of the proposed online monitoring framework.

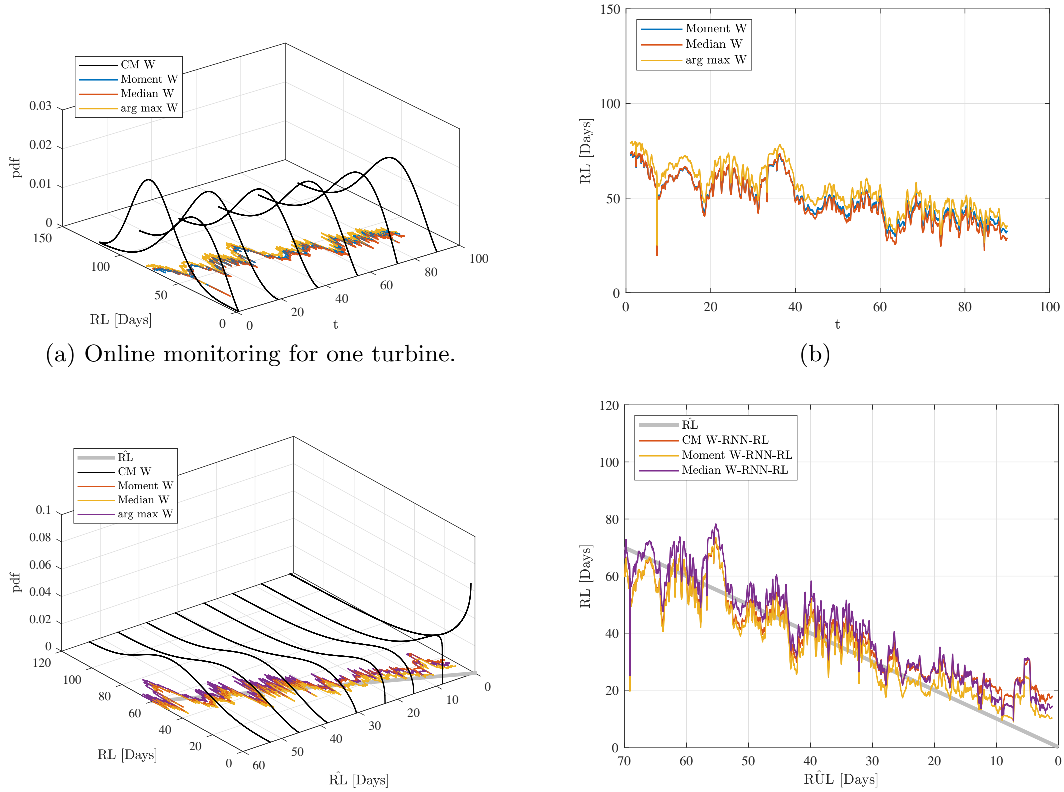
study, the maximum RUL, rounded to the nearest day, is defined as the similarity with 90% confidence, between the true RUL with in a ±5% margin and the models prediction.

The studies are outlined in Table 1. Breaking down Table 1 by Input, Output, and max(RUL), it can be seen that **W**, and variation thereof, can be used on a wide set of input spaces, and is flexible in its output. For **BM** the input space is drastically reduced to $\mathcal{O}(m = 1)$, while maintaining a tractable probability distribution. **VS** is the least flexible model with respect to the input and output space, however, it performs reasonably well for the niche of vibration data with a prediction horizon of 54 days on average. The Gaussian process model (**GP**) is flexible in its input space, but the output space is not tractable and the algorithm scales poorly with increasing m. Thus, no results could be obtained for combinations SCADA, Vibration, and Residuals data.

In terms of the predictive capabilities of each model, ranking from best to worse, the **W** scores RUL above 90 days in most of the presented cases. One may wonder why for **W** and the combination of all three inputs leads to a dramatic drop in the predictive capability. At the current state of this work this must remain as a question to be answered in the future. **VS** achieves prediction up to 53 days, while **BM** and **GP** estimate the average RUL from 42 days to as low as 8 days.

5. Conclusion

This work has proposed a framework for predicting the RUL of wind turbines embedded in on online monitoring framework - combining concepts of statistics and the Turing properties of a neural network. The implementation of the W model and evaluation on wind turbine bearing



(a) Online monitoring for one turbine.

(b)

(c) RUL curve for the same turbine as in Figure 3a.

(d)

Figure 3: Prediction of RUL. Left column: three-dimensional representation of the predicted probability distribution at selected CM points. Right column: top-down view.

data has shown that the proposed model can predict the remaining useful life time beyond 60 days. Besides that, the proposed model shows better long-term prediction capabilities compared to the proposed models by Si et al. [19, 20] (**B**) and Herp et al. [9] (**GP**) and Teng et al. [18] (**VS**).

This study was limited to monitor wind turbine main bearings. In that context it is believed that the proposed model is generic enough to be trained and used on other wind turbine components, as well as in other applications than wind turbines.

References

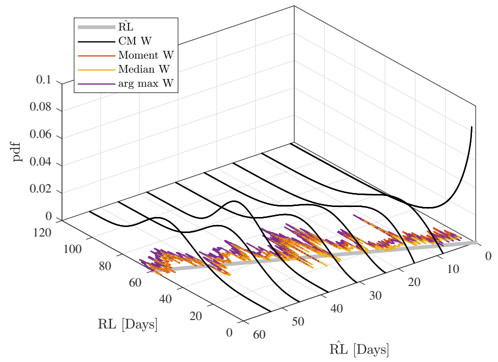
[1] María Isabel Blanco. The economics of wind energy. *Renewable and Sustainable Energy Reviews*, 13(6–7):1372–1382, 2009.

[2] G. J. W. van Bussel and M. B. Zaaijer. Reliability, Availability and Maintenance Aspects of Large-Scale Offshore Wind Farms, a Concepts Study. In *MAREC 2001 Marine Renewable Energies Conference*, volume 113, pages 119–126, March 2001.

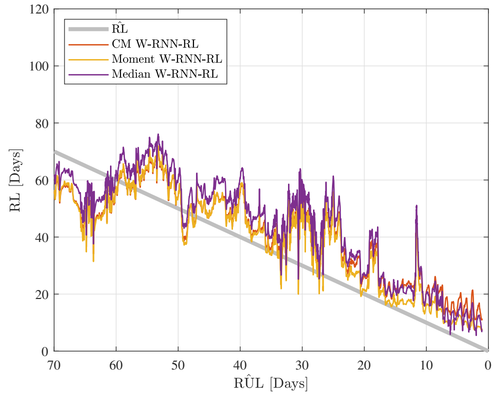
[3] M. Wilkinson, F. Spinato, and M. Knowles. Towards the zero maintenance wind turbine. In *Proceedings of the 41st International Universities Power Engineering Conference, 2006 (UPEC '06)*, volume 1, pages 74–78, 2006.

[4] Christopher A. Walford. Wind turbine reliability: Understanding and minimizing wind turbine operation and maintenance costs. Technical Report SAND2006-1100, Sandia National Laboratories, Sandia National Laboratories, Albuquerque, New Mexico 87185 and Livermore, California 94550, 2006.

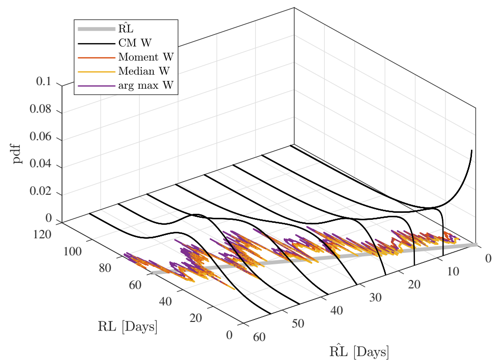
[5] Jannie Jessen Nielsen and John Dalsgaard Sørensen. On risk-based operation and maintenance of offshore wind turbine components. *Reliability Engineering & System Safety*, 96(1):218–229, 2011.



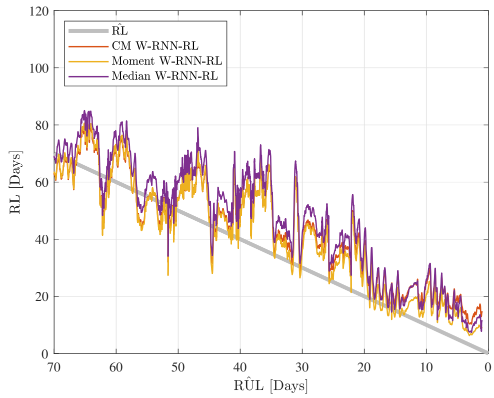
(a) Another turbine's RUL.



(b)



(c) A turbine with more variability in the prediction.



(d)

Figure 4: Prediction of RUL. Left column: three-dimensional representation of the predicted probability distribution at selected CM points. Right column: top-down view.

[6] Kristian Rasmus Petersen, Erik Skov Madsen, and Arne Bilberg. Offshore wind power at rough sea: The need for new maintenance models. In *20th EurOMA Conference*, 2013.

[7] Kristian R. Petersen, Erik Skov Madsen, and Arne Bilberg. First lean, then modularization: improving the maintenance of offshore wind turbines. *International Journal of Energy Sector Management*, 10(2):221–244, 2016.

[8] R. K. Singleton, E. G. Strangas, and S. Aviyente. The use of bearing currents and vibrations in lifetime estimation of bearings. *IEEE Transactions on Industrial Informatics*, 13(3):1301–1309, June 2017.

[9] Jürgen Herp, Mohammad H. Ramezani, Martin Bach-Andersen, Niels L. Pedersen, and Esmaeil S. Nadimi. Bayesian state prediction of wind turbine bearing failure. *Renewable Energy*, 2017.

[10] Martin Bach-Andersen, Bo Rømer-Odgaard, and Ole Winther. Flexible non-linear predictive models for large-scale wind turbine diagnostics. *Wind Energy*, 2016.

[11] Andrew Kusiak and Anoop Verma. Analyzing bearing faults in wind turbines: A data-mining approach. *Renewable Energy*, 48(0):110–116, 2012.

[12] T. H. Loutas, D. Roulias, and G. Georgoulas. Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic e-support vectors regression. *IEEE Transactions on Reliability*, 62(4):821–832, Dec 2013.

[13] N. Li, Y. Lei, J. Lin, and S. X. Ding. An improved exponential model for predicting remaining useful life of rolling element bearings. *IEEE Transactions on Industrial Electronics*, 62(12):7762–7773, Dec 2015.

[14] J. Deutsch and D. He. Using deep learning-based approach to predict remaining useful life of rotating components. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 48(1):11–20, Jan 2018.

[15] Jaouher Ben Ali, Brigitte Chebel-Morello, Lotfi Saidi, Simon Malinowski, and Farhat Fnaiech. Accurate

bearing remaining useful life prediction based on weibull distribution and artificial neural network. *Mechanical Systems and Signal Processing*, 56-57:150 – 172, 2015.

[16] H. Saruhan, S. Sandemir, A. Çiçek, and I. Uygur. Vibration analysis of rolling element bearings defects. *Journal of Applied Research and Technology*, 12(3):384 – 395, 2014.

[17] Faris Elasha, Matthew Greaves, David Mba, and Abdulmajid Addali. Application of acoustic emission in diagnostic of bearing faults within a helicopter gearbox. *Procedia CIRP*, 38:30 – 36, 2015.

[18] Wei Teng, Xiaolong Zhang, Yibing Liu, Andrew Kusiak, and Zhiyong Ma. Prognosis of the remaining useful life of bearings in a wind turbine gearbox. *Energies*, 10(1), 2017.

[19] X. S. Si. An adaptive prognostic approach via nonlinear degradation modeling: Application to battery data. *IEEE Transactions on Industrial Electronics*, 62(8):5082–5096, Aug 2015.

[20] X.S. Si, Z.X. Zhang, and C.H. Hu. *Data-Driven Remaining Useful Life Prognosis Techniques: Stochastic Models, Methods and Applications*. Springer Series in Reliability Engineering. Springer Berlin Heidelberg, 2017.

[21] Jürgen Herp and Esmaeil S. Nadimi. Dimensionality reduction by bayesian eigenvalue-analysis for state prediction in large sensor systems: with application in wind turbines. In *Proceedings of the 2018 Conference on Research in Adaptive and Convergent Systems, RACS 2018, Honolulu, HI, USA, October 09-12, 2018*, pages 1–5, 2018.

[22] Jürgen Herp, Niels L. Pedersen, and Esmaeil S. Nadimi. Using data driven learning for remaining probabilistic lifetime prognosis of wind turbine bearings. Unpublished manuscript, 2018.