Mechanical Systems and Signal Processing 162 (2022) 108087

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Online condition monitoring of floating wind turbines drivetrain by means of digital twin

Farid K. Moghadam *, Amir R. Nejad

Department of Marine Technology, Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Communicated by Paolo Pennacchi

Keywords:
Digital twin model
Modal estimation
System identification
Torsional measurements
Remaining useful life

ABSTRACT

This paper presents a digital twin (DT) condition monitoring approach for drivetrains on floating offshore wind turbines. Digital twin in this context consists of torsional dynamic model, online measurements and fatigue damage estimation which is used for remaining useful life (RUL) estimation. At first, methods for system parameter estimation are presented. The digital twin model provides sufficient inputs for the load observers designed in specific points of the drivetrain to estimate the online load and subsequently stress in the different components. The estimated real-time stress values feed the degradation model of the components. The stochastic degradation model proposed for estimation of real-time fatigue damage in the components is based on a proven model-based approach which is tested under different drivetrain operations, namely normal, faulty and overload conditions. The uncertainties in model, measurements and material properties are addressed, and confidence interval for the estimations is provided by a detailed analysis on the signal behavior and using Monte Carlo simulations. A test case, using 10 MW drivetrain, has been demonstrated.

1. Introduction

In order to realize EU's goal of climate neutrality by 2050, the EU strategy is that 22% of electricity demand in Europe ($300GW/1361GW = 0.22$) should be generated by offshore wind by 2050, [1,2]. A recent report confirms the feasibility of this plan by showing that the levelized cost of energy (LCOE) of both onshore and offshore wind power has dropped considerably in 2019, so that they along with solar photovoltaic (PV) power have been the cheapest form of power generation in Europe, where yet there is a big gap between offshore and onshore [3] wind LCOE. The results of an investigation performed al. [4] emphasizes on the operating expenditures (OPEX) as the main contributor to this gap, so that OPEX in bottom-fixed offshore wind turbines is based on another study [5] in average twice higher than the land-based turbines. There are yet limited experiences with floating offshore wind turbines (FWT) to estimate the actual OPEX.

The motivation of this research is reducing OPEX and subsequently LCOE in FWTs by increasing the wind turbine availability by means of performing predictive maintenance of the turbine critical components and the subsequent reduction of unexpected maintenance and expensive offshore transport/operation cost. The latter is realized through online monitoring based on computationally inexpensive digital twin (DT) models and the subsequent dynamic optimization of the turbine overhaul plan and scheduled maintenance intervals. Predictive maintenance accounts for the components of the system with the highest risk of loss of turbine availability. Power train system including rotor, main bearings, gearbox, generator and power converter accounts for 57% of turbine total failures

* Corresponding author.

https://doi.org/10.1016/j.ymssp.2021.108087
Received 30 December 2020; Received in revised form 13 April 2021; Accepted 25 May 2021

and 65% of turbine total downtime [6]. The overall consequences are expected to be higher in offshore and more specifically FWT which apply to higher power ranges, contain larger components and are encountered with wider range of excitation sources. The drivetrain system in this study comprises rotor, gearbox, generator, main and high-speed shafts, and main bearings, which together cause the majority of the total turbine downtime. The real-time lifetime monitoring of the critical drivetrain component for large FWT is expected to happen in near future [7].

According to the classification provided by Rausand et al. [8], predictive maintenance is a subcategory of condition-based maintenance, which is based on the assessment of remaining useful lifetime (RUL). The latter can be implemented for the turbine critical components to set alarms based on the severity and deviation from the nominal lifetime to inform the operator to take proper actions. The action can be integrated with the scheduled clock-based maintenance for further investigations. The other possibility is to integrate the predictive maintenance outputs to the farm-level decision making support to set the operating point of units based on the turbines condition. Dependent on the risk analysis phase, the action can also be integrated to the protection system.

DT models are proposed in the recent literature for both predictive and condition-based maintenance in different application domains, *e.g.* in aerospace and aviation, oil & gas structures and marine transports industries. The computationally fast equivalent models of the system components, which update themselves based on the operational measurements, facilitate the inexpensive real-time lifetime monitoring of the critical system components. These models are expected to be able to properly capture the physical variations in system. DT is a proven technology used by Siemens Gamesa for prediction of drivetrain loads and subsequently improving the drivetrain design in 11 MW offshore wind turbine technology manufactured by this company [9]. DT modelling is proposed in the industrial revolution Industry 4.0 [10] as the core of development process which can offer unlimited possibilities beyond the design process. The possibility of online estimation of loads enables another potential application domain for DT models which is for monitoring the drivetrain components residual life, where DT can be used for the real-time estimation of stress in the different components of drivetrain, which can provide sufficient input for physics-based/data-driven models for estimation of RUL or probability of failure (PoF) of the individual components and system.

The focus of this research is on a DT solution for RUL estimation of the drivetrain components based on the estimation of drivetrain real-time equivalent model and the subsequent monitoring of variations of stress concentration in different components, and applying the estimated stress to probabilistic and stochastic degradation models which can indicate the fatigue damage of the components. DT in the predictive maintenance context means the combination of *model*, *online measurements* and *degradation model* as suggested by [11]. A starting point in the realization of DT models for fault prognosis/diagnosis purposes is to specify the critical failure modes of the drivetrain system components and then to identify the related failure criterion/feature by analyzing one or a combination of different categories of measurements, namely vibration analysis, electrical signature (current and power signals), acoustic emissions analysis, thermography and temperature analysis, and analysis of oil particles. These measurements can also be used for the estimation of loads from the load effects (vibration, temperature, pressure, etc.), which can later be used by failure functions and degradation models. Nejad et al. [12] discloses a vulnerability map for the drivetrain gearbox which sorts the drivetrain components from the highest to lowest PoF, and discusses about the critical failure modes of each component. Sethuraman *et al.* [13] reports the critical failure modes of the main bearings, whereas Liu *et al.* [14] investigates the critical failure modes of the generator.

There are two different approaches to realize DT models: *In the first approach*, high fidelity physical models which are able to represent the full dynamics with respect to the under consideration failure mode are created. Then the order of model is reduced to maintain the same results while the computational speed is improved. The reduced models are preferred to be implementable in microcontrollers and integrated with turbine fully automated control and monitoring systems, where the operational measurements are also available. This model is then supported by the operational measurements aimed to estimate the equivalent real-time value of resultant load which can unfavorably affect the component with respect to the under consideration failure mode. The latter then plays the role as the input for failure functions which can be a function of different failure modes of different components, and then model-based reliability analysis which is able to capture the degradation and estimate the RUL of a component or system by comparing the estimated stress with the resistance of the components material due to the under consideration failure modes. Pedersen *et al.* [15] applies a similar approach to predict RUL in offshore platform structures, where DT is a finite element model (FEM) which is updated by the online estimation of modal parameters obtained by operational modal analysis (OMA). FEMs are computationally expensive models, which makes these models not to be considered as a feasible solution for condition monitoring of drivetrain as a complex system with a wide range of components. On the other side, drivetrain is a rotational system with various components, failure modes and excitation frequencies, which cause OMA not to be a reliable tool for estimation of drivetrain dynamic properties. *In the second approach*, signal processing approaches including machine learning techniques (*e.g.* artificial neural networks (ANN) and statistical learning approaches) are leveraged to extract the features which are can be connected to the failures. In this category, which is pointed with signal denoising and feature selection, RUL models are more based on empirical models and mostly limited to historical data and based on the available observations of the system operation. Challenges are problem in implementation on new systems with limited observations in different operational conditions, difficulty in interpretation of results, dependency on training data and manually set thresholds which can result in undetected changes or frequent false alarms. As an instance of this category, Herp et al. [16] uses a date-driven approach based on ANN to estimate the target variable related to the residual life of main bearing (temperature) from the turbine operational measurements. Then the monitoring of RUL is realized based on estimating and monitoring the variations of the statistical properties of the distribution fitting the PoF of the bearing based on the empirically defined hazard function.

The aforedescribed DT approaches are based on linking computation models with stochastic models. The resultant is able to relate the variations in the computation model to the expected value of degradation and fatigue damage, and the PoF of the system components [17]. If the stress time series is available, physics-based degradation models can be engaged to estimate RUL, where stochastic models and signal processing techniques (*e.g.* based on statistical learning and ANN) can be engaged to address uncertainties in the

degradation model [18]. For the data-driven degradation models which work based on penalizing the deviation from the distribution of the feature space defined for the normal operation, for instance, statistical learning approaches based on Bayesian method [19], particle filtering [20] and likelihood function [16] are used in the literature for estimating the residual life of wind turbine drivetrain components. Different classes of ANN are also proposed in the literature to support both the feature extraction and the probabilistic models proposed for RUL estimation, [16,21]. In this paper, DT is considered as a physics-based analytical model that its complexity can be adjusted based on the objective of the model. The latter can reduce the computational complexity of DT model and facilitate the implementation in real-time. For the modal estimation, the approach based on analysis of the torsional response proposed by Moghadam *et al.* [22] is employed.

Drivetrain is a rotary system with torsional dynamics and excitation sources. In order to monitor the operation of drivetrain in system level, *e.g.* for detecting the faults in an early stage, the possibility of using drivetrain torsional response is discussed in the recent literature, [22,23]. A drivetrain torsional ROM identified based on the operational torsional measurements, can support both fault diagnosis and prognosis in the drivetrain. It can support fault diagnosis since the drivetrain faults in the system-level show themselves by variations in ROM parameters (stiffness and moment of inertia) and dynamic properties (torsional natural frequencies and normal modes), so that the physics-based threshold can be defined based on the deviation from the reference values of these parameter estimated based on the system normal operations. Torsional models can also support drivetrain fault prognosis. Johansen *et al.* [11] performs a preliminary study about the capacity of different simulation models with different levels of complexity to play the role as drivetrain DT model for condition monitoring purposes. The criterion emphasized in that study is to achieve the same dynamic response by the equivalent model compared to the response from actual system. The possibility of using a 14-DOF equivalent torsional model of drivetrain for estimating the residual life of the gears in the drivetrain gearbox is investigated in [24].

The three steps of realization of DT models for estimation of residual life of the drivetrain components are online equivalent model identification, designing load observers to estimate load and stress, and degradation model to estimate damage. *The first step of* realizing DT platform is the specification and subsequently estimation of the effective real-time equivalent model, which is selected by the system expert based on the selected component and the under investigation failure mode. In this paper, the failure modes which are mainly connected to torsional dynamics are focused, and the equivalent torsional models are proposed in such a way to sufficiently capture the internal dynamics of the component associated to failure mode while the computational burden is minimized. The drivetrain ROM parameters are estimated by using the torsional measurements. The model will then be integrated with the real-time drivetrain measurements in a DT platform for updating the ROM parameters and dynamic properties. The drivetrain real-time ROM estimated/identified from the torsional measurements can be used for *two purposes*: *First*, for fault diagnosis and condition-based maintenance based on torsional response (see *e.g.* [22,23]). In a direct way, by having access to the real-time values of system parameters, it is possible to define different fault states of the different classes of progressive faults in terms of variations in the ROM parameters. Since these parameters are directly representing the physical nature of the system and components, defining the threshold for different fault states is straightforward. In an indirect way, the proposed algorithm helps to access a full knowledge on the real-time values of drivetrain dynamic properties. The deviation from the dynamic properties can be used by the method proposed by [25] to detect the drivetrain faults. The parameters of ROM are gradually updated based on the new data blocks, which help to attain the updated values of dynamic properties. *Second*, for monitoring the drivetrain components residual life and predictive maintenance, which is emphasized in this work, where the estimated ROM along with the real-time measurements are feeding the second step of realization of DT model which is the estimation of load and stress due to the under consideration failure mode. The required input for the proposed ROM estimation method are the torsional measurements including response and drivetrain applied loads. The torsional response can be provided by encoders or strain gauges. It is assumed that the different types of torsional response can be interchangeably used by performing derivation and integration operations. The main loads applied on the drivetrain are the aerodynamic and generator torques. A good estimation of the generator torque is available due to the generator control purposes. However, the measurement of aerodynamic torque is conventionally unavailable, but a good estimation can be attained by using the available measurements. The error in estimation of ROM parameters in both cases of using the actual value and the estimated value of aerodynamic torques as the input for the system identification approach is reported. Based on the theoretical study in [22,25], a 3-DOF torsional model of the drivetrain is sufficient for detection of the drivetrain system level faults. System-level faults are categorized into the faults which change the torsional stiffness the most (*e.g.* crack in the shafts and bearing wear specially in gearbox), and faults which influence mostly the inertia of the drivetrain bodies (changes in mass balance/distribution which can be due to *e.g.* loss of mass, wear and unbalance; and also change in the axis of rotation which can be due to *e.g.* misalignment and looseness), which all can be observed by using 3-DOF ROM. This model is also the reference model for the degradation studies in the system-level as will be discussed further. This model is used in this work for estimation of degradation in the drivetrain components. As the test case for the proposed algorithm in this work, the estimated system parameters along with the angular velocity measurements are used to design two observers for estimation of load in main- and high-speed shafts. The estimated torques will be then applied as inputs to the shaft degradation models which are able to estimate the accumulated damage in the drivetrain shafts. The shafts are the components that can directly contribute in the turbine downtime. They can also indirectly contribute by causing the expedited degradation of the other drivetrain components [26]. Motivations for detection of shaft faults for different applications *e.g.* ship propulsion and wind turbine drivetrain is discussed in the literature, (see *e.g.* [27,28]), which is possible to arise from various conditions such as inappropriate use, stress concentration and unanticipated loading conditions, improper prior fabrication, improper or inadequate design, inadequate maintenance or a combination of them, all of which reduce the fatigue strength of the shaft [29]. Shaft carries all the load and is a core mechanical component in the drivetrain, especially for FWTs which are encountered with high turbulent wind and larger variations of drivetrain loads. In direct-drive technologies the main shaft is considered as a more critical component. A detailed investigation on the wind turbine main shaft failure modes, failure rate, consequences, and economic justifications for monitoring the main shaft lifetime

are discussed in [27]. The consequences of shaft failures are usually high. Damaged shafts can also cause excessive vibrations in the other turbine-generator components, so that monitoring of their operation is important. In the literature, shaft fatigue damage due to different failure modes is discussed. Rauert *et al.* [30] focuses on fretting-fatigue-damage which is described by the product of the frictional shear stress and tangential tensile stress. Zhang *et al.* [31] presents the fracture analysis of the wind turbine main shaft. That study considers torsion, transverse moment and axial force, and estimates equivalent stress. All these studies agree on shear stress as the main role player in the shaft degradation while the bending stress can exacerbate the damage. Based on those studies, the critical step for the shaft fatigue damage estimation is monitoring the shear stress concentration. The main shaft is generally supported by two main bearings in high power drivetrain technologies. The main shaft is designed to meet the requirements of deflections and rigidity [32]. In the detailed design, the main shaft model includes a rotational shaft transmitting the torsional torque due to the rotor main torque component and a fixed shaft supporting the bending moment due to the shaft weight [33]. The equivalent stress can then be estimated by applying the von Mises theory. Safety factors are selected and applied according to the specific material type in EN10083 standard [34]. The same procedure is followed, so that the fatigue damage due to cyclic torsional-bending load is focused in estimations of shaft residual life in this study.

As discussed earlier, *in the second step of* DT model realization, the load in different parts of this system are estimated by using the updated DT model parameters and designed load observers. The load observers are designed by using the analytical model of the system and the ROM parameters fed by the operational measurements. The uncertainties in the torsional response, input load and ROM parameters are taken into account, and the states/unnoisy response are estimated by using Kalman filter, where the output provides the sufficient inputs for the designed load observers.

The third step of the attainment of DT model for residual life monitoring of drivetrain components is to establish degradation models. In general, degradation models comprise two components, namely failure parameter and RUL estimation units. The failure parameter can be directly measured by physical sensors or estimated by employment of either data-driven or physics-based observers. Data-driven approaches depend on analysis of historical data and application of artificial intelligence in different operational and environmental condition and estimation of a parameter which is directly related to the components lifetime heuristically and by analysis of the patterns observed based on the available data. More specifically, the artificial intelligence techniques preform regression or correlation analysis with respect to the data of a sensor network and look out for similarities, variations and deviations from a pattern in the dataset over the time. It is usually difficult to rely on the results obtained from analyzing limited datasets, and not easy to extend the results to different operating conditions and applications. However, physics-based approaches are based on the system physical rules of flow of energy in the components and the resulted vibrations and temperature as the responses. Finding an efficient way to address complex dynamics and transients, and various sources of uncertainties of the real problems by physical models is sometimes challenging. However, owe to advances in analytical models and stochastic modelling techniques, improved and reliable physical models are the preferred choice for engineering applications. Adjustability and adaptability of model complexity based on the application, independency of results of historical data and comfortability in defining meaningful thresholds are the other motivations for sticking to physical models. This paper is based on a proven physics-based approach for degradation estimation, which is supported by statistical approaches and stochastic models to address uncertainties to improve the accuracy. This method which relies on real-time measurements, is computationally fast and can apply to different dirvetrain components and is not restricted by the operational conditions.

The main purpose of this work is the proof of concept for the proposed DT model for residual life monitoring of the drivetrain components. The input measurements come from the high fidelity models. The validation of the estimated model parameters is performed by comparing the estimated parameters with the actual values. The RUL model is based on the proven physical rules of stress concentration and degradation of the components material. As discussed earlier, DT models rely on real-time operational measurements to update themselves. Due to the limited historical data of FWTs and the test facilities for high power applications, the validation is limited to simulation studies. The proposed method combines high fidelity models to generate the torsional response, reduced order models (ROM) to represent as the DT of the system, signal processing approaches to estimate ROM parameters, and stochastic frameworks to address uncertainties in the proposed DT model. High fidelity models are used to estimate the drivetrain loads from the global simulations, and then to calculate the drivetrain response in the different bodies. The degradation model works based on the updated DT parameters and online response measurements. The input torque is applied to the model once from the global simulations. Since in a real case, the input torque is not among the accessible measurements, in another simulation, applied input torque is estimated from the accessible turbine and blade parameters to account for the uncertainties in the drivetrain input loads. The sources of uncertainty in the proposed DT model mainly arise from the load estimation approach, and the estimation of fatigue damage by relying on the material properties obtained from the S-N curve, which are sufficiently addressed by using statistical approaches and stochastic models supported by Kalman filtering and Monte Carlo simulations.

An innovative drivetrain health monitoring approach based on estimation of residual life of the components is proposed by using the torsional measurements and digital twin modeling. Drivetrain model identification by using torsional measurements, and application of the real-time estimated DT model for estimation of load and residual life of the components is the main goal of this work. The method is designed in the general form for n-DOF torsional model of the drivetrain, and then developed by using 3-DOF torsional model as the drivetrain DT for monitoring the residual life of main and high-speed shafts. An algorithm is proposed for the model identification, which receives the torsional response and the estimated rotor and generator torques, and estimates the drivetrain ROM parameters and dynamic properties. The proposed method is computationally fast, and can be implemented in the farm level. The method can be adapted for higher DOF models to monitor the lifetime of different drivetrain components. For demonstration purposes, the algorithm is designed and tested for monitoring the lifetime of shafts in the drivetrain system. This approach is in particular useful for operators with limited knowledge of the drivetrain parameters and dynamic properties, as we propose a method for model

estimation first - in other words the approach is not limited to have all the drivetrain design data. In addition, in order to overcome the practical implementation challenges, the different sources of uncertainty in all the three steps of DT model realization are modelled and the impacts are mitigated. On this basis, *the main contributions* of this work are:

1. Establishing the digital twin approach and proposing a computationally fast digital twin model of the drivetrain system based on the torsional measurements, aimed at monitoring the RUL of the components, and providing confidence interval for damage by using stochastic models and statistical approaches,
2. Designing a robust and computationally fast method for estimating parameters of drivetrain equivalent models of different DOFs by torsional measurements, for fault diagnosis and prognosis purposes, and providing confidence interval for estimation error by using statistical approaches,
3. Proposing a stochastic physics-based degradation model for estimation of RUL in the drivetrain main shaft by using the online estimated ROM, real-time operational measurements, designed load observer and equivalent stress estimation approach, and taking into consideration the various sources of uncertainty by using statistical approaches and stochastic modelling of damage.

The rest of paper is organized as follows: In Section II, the drivetrain DT model and the online estimation of ROM parameters by using the torsional response and aerodynamic torque observer is discussed. Then, in the rest of this Section, the design of real-time load observers for the drivetrain components by using DT model and operational measurements, the estimation of equivalent stress and then the degradation model is explained. The proposed DT model estimation approach and its application for monitoring of drivetrain components residual life is evaluated by the simulation studies in Section III. The work is closed with the final remarks in Section IV.

2. Methodology

The three steps of realizing the proposed DT framework for estimating the residual life of drivetrain components (*e.g.* for the gears of gearbox) is summarized by the flowchart in Fig. 1. In the continued part of this section, the employed simulation-based drivetrain analysis approach and the steps of realization of DT model are explained.

2.1. Global simulation and drivetrain loads

DTU 10 MW reference wind turbine [35] with a spar floating support substructure is selected for this study. The wind turbine specification and the overall characteristics of the floating platform is obtained from [35,36]. This model is able to capture the global dynamics of spar FWT from the interactions with the environmental loads. The drivetrain system is a medium-speed permanent magnet synchronous generator technology based on the gearbox and generator design specifications reported by Moghadam *et al.* [37]. The decoupled simulation approach is used for the drivetrain studies in this work, which consists of two steps. In the first step, global simulation analysis for different environmental conditions is performed. In the global simulation, the blades and hub assembly, the structural module including the flexible multi-body systems for tower and platform including the floating support substructure and the nacelle are modelled. This model is coping with combined aerodynamic and hydrodynamic loading by using numerical and probabilistic models of wind, waves and current in the global simulation software to capture the integrated effect of the loads and the wind turbine control system on the turbine components. The results of the global simulation are the load effects on different parts of turbine, which contain the information about the interactions between the turbine subsystems. The latter includes the loads transmitted to the drivetrain by the rotor and structure specified by the time series of the resultant moments and forces on the rotor, tower top

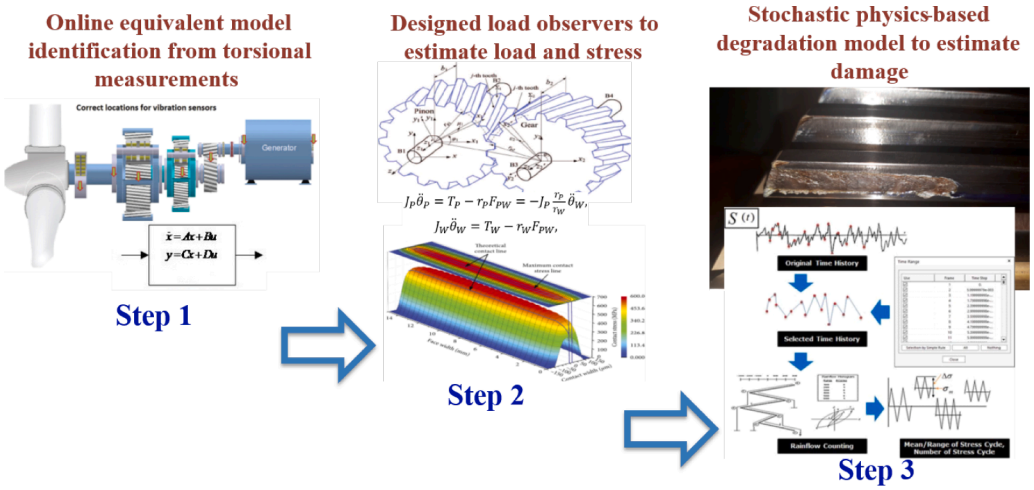


Fig. 1. Proposed digital twin framework for estimating the residual life of drivetrain components.

accelerations and other responses of interest. The second step of the decoupled analysis is that the calculated rotor aerodynamic torque and other responses of interest estimated from global simulation are applied as the input to an offline drivetrain model in mechanical/electrical machinery simulation software to calculate and analyse the drivetrain components local dynamic responses. Then the drivetrain local load effects are obtained for further post processing aimed at health monitoring of the drivetrain system based on monitoring the critical components residual life. The wind model in the global simulations is turbulent based on Kaimal distribution. The turbulence intensity at hub height $I\,(-)$ is assumed to be 0.14. The wave is modelled stochastic by two parameters, namely significant wave height $H_s\ (m)$ and peak period $T_p\ (s)$ in global analysis.

2.2. Estimation of the drivetrain model parameters and dynamic properties

The proposed algorithm for estimation of the drivetrain equivalent ROM and dynamic properties is summarized by the flowchart illustrated in Fig. 2. This algorithm consists of *four* main components, namely estimation of natural frequencies from torsional

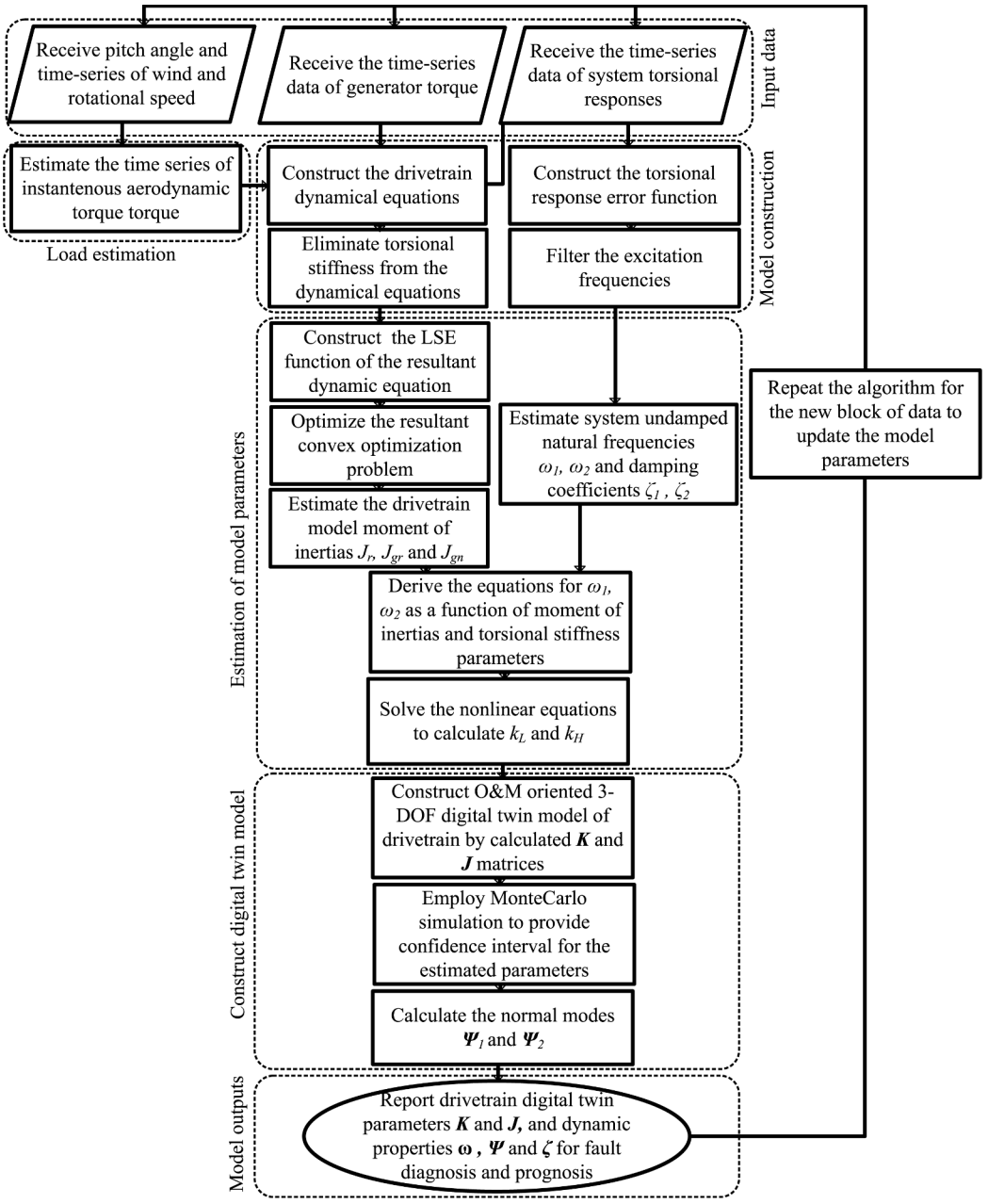


Fig. 2. Proposed algorithm for estimation of drivetrain equivalent model parameters.

response; estimation of ROM parameters by using estimated natural frequencies, torsional response and drivetrain input loads; estimation of drivetrain input rotor and generator torques by using available turbine measurements; and analytically calculating ROM parameters estimation error and confidence interval for estimated values. In the following, the different components of the algorithm are described.

2.2.1. Drivetrain modal analysis

The free vibration time-domain torsional response of n-DOF damped rotational system in terms of the system dynamic properties can be represented by

$$\boldsymbol{\theta}\left(t\right) = \sum_{i=1}^{n} \boldsymbol{a}^{i} e^{-\boldsymbol{\zeta}^{i}\boldsymbol{\omega}^{i}t} \boldsymbol{\Psi}^{i} \sin\left(\boldsymbol{\omega}_{d}^{i} t + \boldsymbol{\phi}^{i}\right), \tag{1}$$

where $\boldsymbol{\theta}$ is a vector representing response in different DOFs. The amplitude coefficients $\boldsymbol{a}^{i}$ and phase shifts $\boldsymbol{\phi}^{i}$ are determined by the initial conditions of angular displacements and velocities. The complex mode shape vectors $\boldsymbol{\Psi}^{i}$, the undamped natural frequencies $\boldsymbol{\omega}^{i}$, and the damping coefficients $\boldsymbol{\zeta}^{i}$ are all representing the dynamic properties of the system which all are functions of physical parameters of the system and not the loading or initial conditions. The damped natural frequencies $\boldsymbol{\omega}_{d}^{i}$ can also be related to the undamped frequencies by using the damping coefficients as

$$\boldsymbol{\omega}_{d}^{i} = \boldsymbol{\omega}^{i}\sqrt{1-(\boldsymbol{\zeta}^{i})^{2}}. \tag{2}$$

As mentioned earlier, the drivetrain dynamic properties can each be represented as a function of system parameters. In [25], the closed form equations of the reference 3-DOF model dynamic properties as a function of model parameters are derived. By this method, it is possible to achieve a full dynamic characterization of the drivetrain system by specification of natural frequencies, mode shapes and damping in the system. By estimation of the model parameters by using the torsional measurements and the theory developed in the continued parts of this section, it is possible to access a full knowledge on the system dynamic properties, which can be used as input for the drivetrain system fault detection method proposed by [22,25]. Among the dynamic properties, the natural frequencies can also be directly estimated by using the torsional response [22], which are later used as inputs for the proposed system identification approach.

The torsional response residual function between the inertias j_l and j_m from the point l is defined as [25]

$$e_{l,m}^{\Omega}\left(t\right) \triangleq \boldsymbol{\Omega}_{l}\left(t\right) - u_{l,m}\boldsymbol{\Omega}_{m}\left(t\right), \ \ for\ l\ and\ m\ \in \left\{1,\ldots,n\right\}, \tag{3}$$

with $\boldsymbol{\Omega}$ as the time series of angular velocity, and $u_{l,m}$ as the relative gear ratio between j_l and j_m to make them in the same coordinate. Gear ratio $u_{l,m}$ as per definition is $\frac{N_l}{N_m}$, where N_l and N_m are the speeds at l^{th} and m^{th} inertias. The analytical proof of estimation of natural frequencies from the frequency spectrum of $e_{l,m}^{\Omega}$ is provided by [25].

2.2.2. Estimation of moment of inertia matrix

As the first step, the general damped n-DOF torsional dynamical model of drivetrain is constructed. The summation of the moments on each inertia in the lumped parameter model yields n equations of the form

$$j_i\ddot{\boldsymbol{\theta}}_i(t) + C_i(\dot{\boldsymbol{\theta}}_i(t) - \dot{\boldsymbol{\theta}}_{i-1}(t)) - C_{i+1}(\dot{\boldsymbol{\theta}}_{i+1}(t) - \dot{\boldsymbol{\theta}}_i(t)) + k_i(\boldsymbol{\theta}_i(t) - \boldsymbol{\theta}_{i-1}(t)) - k_{i+1}(\boldsymbol{\theta}_{i+1}(t) - \boldsymbol{\theta}_i(t)) = T_i(t) \ \ for \ \ i = (1,\ldots,n) \tag{4}$$

where $\boldsymbol{\theta}_i$ is the angular displacement at i^{th} body. j_i is the inertia of i^{th} body. k_i is the equivalent stiffness between $(i-1)^{th}$ and i^{th} bodies. k_{i+1} is the equivalent stiffness between i^{th} and $(i+1)^{th}$ bodies. T_i is the external excitation applied to the i^{th} body. In the matrix form, these set of equations can be written as

$$\mathbf{J}\ddot{\boldsymbol{\Theta}}(t) + \mathbf{C}\dot{\boldsymbol{\Theta}}(t) + \mathbf{K}\boldsymbol{\Theta}(t) = \mathbf{T}(t). \tag{5}$$

where $\mathbf{J}, \mathbf{C}$ and $\mathbf{K}$ are the moment of inertia, damping and stiffness matrices with the size $n \times n$. $\boldsymbol{\Theta}$ and T are the response and load vectors with the size $n \times 1$, where each element of these two vectors represents a time series data. This model alongside the torsional measurements provide the inputs for the drivetrain model parameter estimation approach. The optimization variables are $\mathbf{J}, \mathbf{C}$ and $\mathbf{K}$ matrices which are the drivetrain equivalent lumped model parameters. The sparsity of the matrix variables $\mathbf{J}, \mathbf{C}$ and $\mathbf{K}$ are specified based on the drivetrain topology and is imposed to the optimization problem. Assuming that the load and response time series are known, the parameter estimation turns to the minimization of the $L2$-norm of error. The error function is defined by

$$E(t) \triangleq \widehat{\mathbf{J}}\ddot{\boldsymbol{\Theta}}(t) + \widehat{\mathbf{C}}\dot{\boldsymbol{\Theta}}(t) + \widehat{\mathbf{K}}\boldsymbol{\Theta}(t) - \mathbf{T}\left(\mathrm{t}\right). \tag{6}$$

The least square estimator is defined by

$$\begin{gathered}\widehat{\mathbf{J}}^{\mathbf{LS}}, \widehat{\mathbf{K}}^{\mathbf{LS}}, \widehat{\mathbf{C}}^{\mathbf{LS}} \in \operatorname{argmin}\left\{\|E\|_2\right\} \\ \mathbf{J}, \mathbf{K}, \mathbf{C} \succcurlyeq 0 \\ \mathbf{J}_{l,m} \in S^{\mathbf{J}} \\ \mathbf{K}_{l,m} \in S^{\mathbf{K}} \\ \mathbf{C}_{l,m} \in S^{\mathbf{C}},\end{gathered} \tag{7}$$

where $S^{\mathbf{J}}, S^{\mathbf{K}}$ and $S^{\mathbf{C}}$ are the sparsity sets of matrices $\mathbf{J}$, $\mathbf{K}$ and $\mathbf{C}$. The matrix $\mathbf{J}$ is diagonal. $\mathbf{K}$ and $\mathbf{C}$ are non-diagonal symmetric matrices, but are not full rank. The latter causes computational difficulty for the above quadratic matrix optimization problem, so that it may result in the divergence of the numerical solver. To be more specific, these terms can introduce larger perturbations into the calculations than its numerically stable counterpart; this can lead to larger errors in the final computed solution. In order to remove the coupling between the equations due to the $\mathbf{K}$ and $\mathbf{C}$ terms, to cope with the round-off error which arises from the ill-condition terms of model, and to reduce the computational complexity by reduction of the number of variables, the equivalent scalar optimization problem is constructed by the sum of the dynamic Eqs. (4) of each inertia, which leads to the elimination of stiffness and damping from the resultant scalar equation. The latter leads to the following error function in terms of the inertia variables j_i as the model scalar variables and torsional measurements time-series as input, with rotor as the reference of the rotary coordinate.

$$e(t) = j_1\dot{\mathbf{\Omega}}_1(t) + \ldots + u_{1,i}j_i\dot{\mathbf{\Omega}}_i(t) + \ldots + u_{1,n}j_n\dot{\mathbf{\Omega}}_n(t) - T_r(t) - u_{1,n}T_{gn}(t), \tag{8}$$

where $\dot{\mathbf{\Omega}}_i$ is the time series of angular acceleration and j_i is the moment of inertia of the i^{th} DOF. T_r and T_{gn} are the time series of the rotor and generator torques, respectively. The response used in this equation is the angular acceleration which can be obtained by applying a derivation operation on the angular velocity measurements, or the second derivation on the angular displacement measurements. The sign of $u_{1,i}$ is determined based on the direction of rotation of j_i. Since the real-time values of response is known, the inertia parameters can be estimated by minimizing the square error between the model and measurements. The average of the squares of the errors — that is, the average squared difference between the estimated values as the model outputs and the actual value as the sensor measurements is defined as least squares produces best linear unbiased estimators of the coefficients in a linear regression model. Therefore, a least-square-error (LSE) function is defined to minimize the error between the dual optimization problem and input operational measurements. The latter leads to the following quadratic scalar optimization problem as

$$\widehat{\mathbf{j}}^{\mathrm{LS}} = \operatorname{argmin}\left\{\|e\|_2\right\} j \succcurlyeq 0. \tag{9}$$

This estimator is robust to the measurement noises, and can provide a good approximation even with less than n input data samples (underdetermined case). For the case of more than n samples (overdetermined case), this estimator helps to obtain more accurate estimation than solving the linear equations, when the input measurements are subject to independent and identically distributed (IID) Gaussian noise. In other words, the total LS technique is able to correct the system with minimum perturbation [38]. The above convex multi-variate scalar optimization problem is numerically solved by Matlab CVX and the global optimizer $j^{LS} = \{j_1, \ldots, j_n\}$ is estimated. Since this $L2$-norm regression optimization problem is in a quadratic convex form, the results which are the estimated values of the drivetrain ROM inertia parameters are the global optimizers of the problem.

The maximum likelihood estimator (MLE) defined by the following maximization problem can also be used, which is mathematically equivalent to the LSE problem defined by Eq. (9), for the special case of Gaussian noise.

$$\widehat{j}^{ML} = L\left(e\left(1\right), \ldots, e\left(t\right); j_1, \ldots, j_n, \sigma^2\right) = \operatorname{argmax}\left\{\frac{1}{(2\pi)^{t/2}\sigma^t} exp\left(\frac{-1}{2\sigma^2}\|e\|_2\right)\right\} \equiv \arg\min\left\{\|e\|_2\right\} \tag{10}$$

where σ is the standard deviation of the Gaussian distribution fitting the noise. To use MLE to estimate those parameters, the method is restricted to the assumption that the form of the distribution of the random noise defined by the error function $\mathbf{e}(t)$ is known, so that the likelihood function can be obtained. Based on the extensive simulations, the error calculated from the torsional measurements in the under consideration application shows a near Gaussian distribution. Therefore, both LSE and MLE estimators lead to the similar set of results.

2.2.3. Estimation of stiffness matrix

Afterwards, the stiffness parameters of the model are estimated by using the estimated inertias and the drivetrain resonance frequencies estimated from the torsional measurements in the previous steps. The undamped torsional frequencies of the system are the nonlinear function of inertia and stiffness as

$$\boldsymbol{\omega}_i \left(for\ i = 1, \ldots, n\right) = \sqrt{eig\left(-\mathbf{J}^{-1}\mathbf{K}\right)}. \tag{11}$$

By using the estimated natural frequencies obtained from the modal estimation approach together with the estimated inertia matrix $\mathbf{J}$ from the LSE optimization problem, the stiffness matrix $\mathbf{K}$ is the root of g_i which is defined by the following nonlinear equation as

$$g_i = \boldsymbol{\omega}_i^2 - eig\left(-\mathbf{J}^{-1}\mathbf{K}\right), \quad (for\ i = 1, \ldots, n). \tag{12}$$

In general, there is not a unique matrix $\mathbf{K}$ from the above equation for the known set of eigenvalues $\lambda_i = \omega_1^2, \ldots, \omega_n^2$ of the matrix $-\mathbf{J}^{-1}\mathbf{K}$. However, by imposing the sparsity and symmetricity to matrix $\mathbf{K}$ from the lumped model, it is possible to calculate the unique matrix $\mathbf{K}$ numerically by using Matlab fsolve solver. The latter also helps to reduce commutation cost of this matrix algebraic equation by reducing the number of variables from n^2 to n. The matrix $\mathbf{J}^{-1}\mathbf{K}$ is not symmetric in the general case which may give the sense that there are multiple answers for $\mathbf{K}$ from this equation. However the fact that $-\mathbf{J}^{-1}\mathbf{K}$ always has positive eigenvalues (it is positive definite), brings us to the believe that this matrix is a small perturbation of a symmetric matrix with positive eigenvalues. Small perturbation keeps the eigenvalues positive [39].

The usual condition for the estimation problem is more restrictive. In other words, it is possible that only some of the eigenfrequencies of the drivetrain system can be estimated by employing the aforedescribed modal estimation approach, especially the higher eigenfrequencies which are excited with a lower energy of the input torque. In this case, the matrix $\mathbf{K}$ can still be estimated by using the following optimization problem in terms of the first i eigenfrequencies as defined by the following least square error estimator:

$$\widehat{\mathbf{k}}^{\mathrm{LS}} = \mathrm{argmin}\left\{ \|\lambda_i - eig(\boldsymbol{\Lambda},\ i)\|^2 \right\} \boldsymbol{\Lambda} \in \Lambda, \tag{13}$$

with $\boldsymbol{\Lambda}$ is the variable of this problem which is a function of the unknown variable $\mathbf{K}$ as $\boldsymbol{\Lambda} = -\mathbf{J}^{-1}\mathbf{K}$. Also $\widehat{k}^{LS}$ is the set of nonzero elements of matrix $\mathbf{K}$ which are estimated by the above nonlinear matrix optimization problem. The sign of the elements of k are forced in the optimization problem. λ_i is the set of i $(i \in \{1, \ldots, n\})$ smallest magnitude eigenvalues which are known from the modal estimation, $\lambda_i = \{\boldsymbol{\omega}_1^2, \ldots, \omega_i^2\}$. $eig(\boldsymbol{\Lambda},\ i)$ is the set of i $(i \in \{1, \ldots, n\})$ smallest magnitude eigenvalues defined in terms of matrix $\mathbf{J}$ and the unknown matrix $\mathbf{K}$. The feasible set Λ is also defined by

$$\Lambda = \left\{ \boldsymbol{\Lambda} : \boldsymbol{\Lambda} \in \mathbb{R}^{n \times n}, \quad \mathbf{K} \geqslant 0, \quad \boldsymbol{\Lambda}_{l,m} = 0, \ \forall\ \boldsymbol{\Lambda}_{l,m} \in S^{\boldsymbol{\Lambda}} \right\}, \tag{14}$$

where $S^{\boldsymbol{\Lambda}}$ is the sparsity set of matrix$\boldsymbol{\Lambda}$. The positive definiteness and sparsity of $\boldsymbol{\Lambda}$ are the nonlinear constraints which are imposed to this problem. For the set of positive semidefinite matrices, this problem is convex and the solution is the global optimizer. However, $\boldsymbol{\Lambda}$ is not symmetric in general so that the definition of the problem is nonconvex for the numerical solvers and convex optimization tools are not able to numerically solve the problem. For this purpose, Matlab {fmincon} solver as a powerful tool for the general class of nonlinear nonconvex problems is used.

The estimation of undamped natural frequencies and damping coefficients from the torsional measurements is discussed in the algorithm proposed by [22]. For the system with n-DOF, the system has $n-1$ non-rigid torsional modes. The latter leads to $n-1$ nonlinear equations which are the undamped natural frequencies as nonlinear functions of equivalent model inertia and stiffness links. These set of nonlinear equations are numerically solved, and n equivalent stiffness seen by each body are estimated. The estimated values of stiffness parameters of the main diagonal of the matrix, by considering both the natural frequencies and inertia parameters estimation errors. When the degree of the model increases, it is not easy to access the closed form of these n nonlinear equations, so that one may decide to solve the equations numerically in matrix form. The increase of the degree of model increases the algorithm computationally more expensive but can provide more detailed DT model for condition monitoring of the different drivetrain subcomponents. As discussed earlier, 3-DOF model can be used for the drivetrain system-level faults. This model can also be used for lifetime monitoring of the drivetrain main and high-speed shafts. The estimation of 3-DOF model parameters and dynamic properties by this approach needs an additional torsional measurement installed on the gearbox input or output shafts. fault prognosis *e.g.* for shaft crack may be performed by using more detailed models (*e.g.* finite element models) able to better model the stress concentration in the crack side in the nonuniform shape of shafts. In order to compromise between the complexity and accuracy in this application and for providing a rough estimation of RUL by using a method implementable by turbine on-board automation system, 3-DOF model is proposed. The summary of the proposed algorithm adjusted for estimation of 3-DOF model parameters and dynamic properties from the torsional measurements is presented by Fig. 2. In this case, the nonlinear matrix equation represented by Eq. (13) is reduced to the two nonlinear scalar equations for the two non-rigid modes as [25]

$$\mathrm{F1}\left(\mathrm{k_L}, \mathrm{k_H}\right) = \omega_1 - \sqrt{\frac{\mathrm{k_L}}{2\mathrm{J_r}} + \frac{\mathrm{k_L}+\mathrm{k_H}}{2\mathrm{J_{gr}}} + \frac{\mathrm{k_H}}{2\mathrm{J_{gn}}} - \sqrt{\left(\frac{-\mathrm{k_L}}{2\mathrm{J_r}} - \frac{\mathrm{k_L}-\mathrm{k_H}}{2\mathrm{J_{gr}}} + \frac{\mathrm{k_H}}{2\mathrm{J_{gn}}}\right)^2 + \frac{\mathrm{k_L k_H}}{\mathrm{J_{gr}^2}}}}, \tag{15a}$$

$$\mathrm{F2}\left(\mathrm{k_L}, \mathrm{k_H}\right) = \omega_2 - \sqrt{\frac{\mathrm{k_L}}{2\mathrm{J_r}} + \frac{\mathrm{k_L}+\mathrm{k_H}}{2\mathrm{J_{gr}}} + \frac{\mathrm{k_H}}{2\mathrm{J_{gn}}} + \sqrt{\left(\frac{-\mathrm{k_L}}{2\mathrm{J_r}} - \frac{\mathrm{k_L}-\mathrm{k_H}}{2\mathrm{J_{gr}}} + \frac{\mathrm{k_H}}{2\mathrm{J_{gn}}}\right)^2 + \frac{\mathrm{k_L k_H}}{\mathrm{J_{gr}^2}}}}. \tag{15b}$$

By solving these two nonlinear equations, the values of k_L and k_H which are respectively the stiffness of main and high-speed shafts are estimated. For all $k_L, k_H \in R^+$, there is a unique solution for the above set of nonlinear equations which make it easy to numerically solve the equations. In case of access to the preliminary values of the parameters, these two equations can be approximated with the affine functions by using the two first Taylor series terms

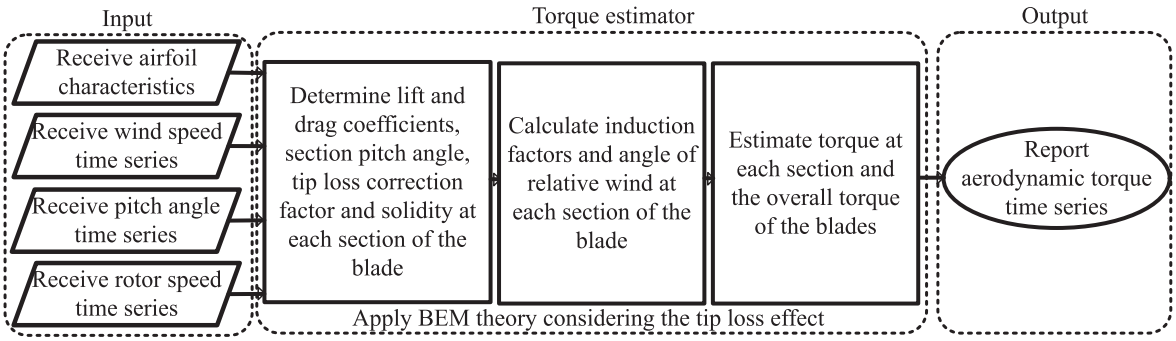


Fig. 3. Estimation of aerodynamic torque from the turbine measurements and airfoil characteristics.

$$\mathrm{F1}(\mathrm{k_L},\mathrm{k_H}) \approx \mathrm{F1}(\mathrm{k_L}^*,\mathrm{k_H}^*) + (\mathrm{k_L}-\mathrm{k_L}^*)\mathrm{F1}_{\mathrm{k_L}}(\mathrm{k_L}^*,\mathrm{k_H}^*) + (\mathrm{k_H}-\mathrm{k_H}^*)\mathrm{F1}_{\mathrm{k_H}}(\mathrm{k_L}^*,\mathrm{k_H}^*), \tag{16a}$$

$$\mathrm{F2}(\mathrm{k_L},\mathrm{k_H}) \approx \mathrm{F2}(\mathrm{k_L}^*,\mathrm{k_H}^*) + (\mathrm{k_L}-\mathrm{k_L}^*)\mathrm{F2}_{\mathrm{k_L}}(\mathrm{k_L}^*,\mathrm{k_H}^*) + (\mathrm{k_H}-\mathrm{k_H}^*)\mathrm{F2}_{\mathrm{k_H}}(\mathrm{k_L}^*,\mathrm{k_H}^*). \tag{16b}$$

Note that $\mathrm{F}_{\mathrm{k_L}}$ and $\mathrm{F}_{\mathrm{k_H}}$ are the partial derivatives of F with respect to k_L and k_H, which their values should be updated based on the new values of natural frequencies and inertia parameters.

2.2.4. Inputs of the proposed parameter estimation approach

The input data can be classified into two categories: first, the torsional response time series of the bodies, which are obtained from the drivetrain MBS model in Simpack. Second, the drivetrain loads consisting of the generator and rotor torques time series. The generator torque is a measurement available in the turbine for the generator control purposes. The generator torque is estimated from the generator electrical measurements, and available in turbine main control unit. The generator reference torque is calculated from the following equation

$$T_{gn}^* = \frac{P^*}{\Omega_{gn}}, \tag{17}$$

where P^* is the reference power and Ω_{gn} is the generator speed. T_{gn}^* is then used for the design of generator internal current control loop by using the equation

$$i_q^* = \frac{2T_{gn}^*}{3p\phi_f}. \tag{18}$$

where i_q^* is the reference q-axis current in rotating dq frame, used for the generator current control loop. p is the number of poles pairs, and ϕ_f is the flux linkage. The resulted electromagnetic torque on the shaft is calculated by using the measured current i_q as [40]

$$T_{gn} = \frac{3}{2}p\phi_f i_q. \tag{19}$$

The aerodynamic torque applied to the drivetrain MBS model in Simpack is obtained from the global simulations. This torque is considered as the applied torque on the drivetrain model. However, in the real case the real value of applied aerodynamic torque is not available, but a estimation can be available by using the turbine operational measurements and general information of the airfoil. The estimated torque is applied as the input to the drivetrain model identification approach. The algorithm for aerodynamic torque estimation approach is illustrated by flowchart shown in Fig. 3.

About the practicality issues of implementing this idea in current turbines, Supervisory Control and Data Acquisition (SCADA) existing system contains the encoder measurements, but implementation of the method needs higher resolution angular velocity measurements than what is available today from SCADA. SCADA data contains generator torque that is calculated from voltage and electric current. A good estimation of the rotor torque can also be available by using the available turbine measurements from SCADA.

2.2.5. Estimation error and confidence interval

The computational complexity of the proposed approach is proportional to the length t of the data block. In order to check the estimation accuracy versus the number of samples, the relative estimation error of each estimated parameter should be monitored versus different lengths of input data. To mitigate the influence of uncertainty in the input measurements in the estimated error, the crude Monte Carlo simulation is employed. The concept behind this is to make the calculated estimation error independent of the uncertainty in the input data. The model estimated parameter y is a random variable. Zero overlapping is allowed for the input data blocks which help to be able to assume that the estimated parameters are statistically independent. For the case that the data is captured during the similar operational conditions with respect to the operational speed, the response and subsequently the estimated

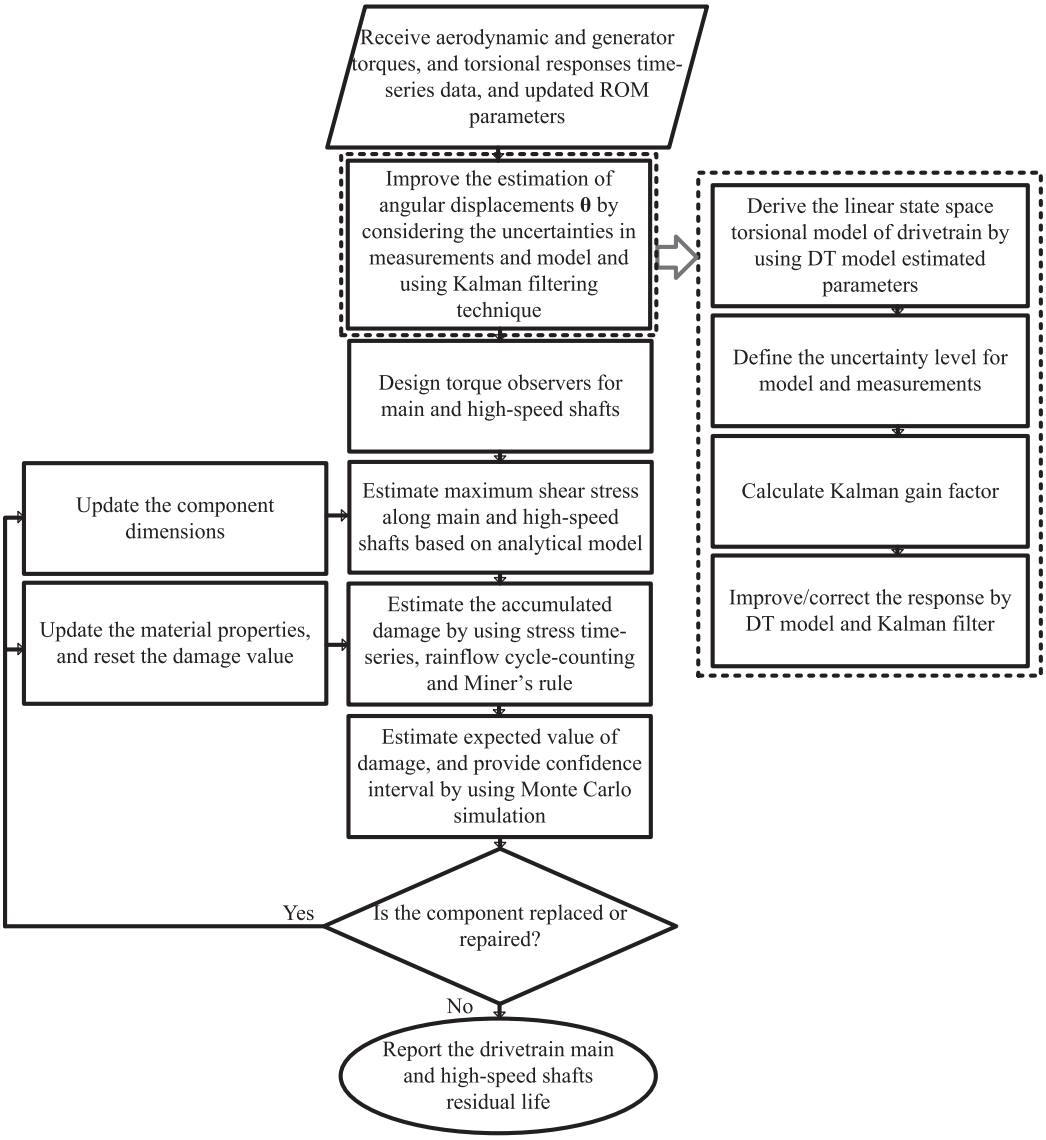


Fig. 4. Estimation of drivetrain components RUL by using DT model, torsional measurements and stress-life method.

parameters can be assumed to be identically distributed. Therefore, such rules in setting the test condition, will help to make the realistic assumption that y is IID. These conditions help to attain the confidence interval for the estimation error by using the results of IID central limit theorem (CLT). If $\varepsilon_1, \ldots, \varepsilon_r \in R$ are the errors in estimation of y based on r different blocks of input data each with the length t, with β% confidence, the estimation error locates in the interval [41]

$$\{\bar{\varepsilon}_r^t\}_y \in \left[\widehat{\mu}_l - \phi^{-1}\left(1-\beta\Big/2\right)\frac{s}{\sqrt{r}} \quad \widehat{\mu}_l + \phi^{-1}\left(1-\beta\Big/2\right)\frac{s}{\sqrt{r}}\right], \tag{20}$$

where $\{\bar{\varepsilon}_r^t\}_y$ is the average error associated with parameter y calculated based on t samples of data and r blocks of data. μ and s are the average and standard deviation of ε_i, $for\ i = (1, \ldots, r)$. Therefore to achieve a specific level of estimation error $\bar{\varepsilon}_r^t$ with a specific confidence β%, the number of data blocks can be analytically acquired. The size of each block can also influence on the covariance, so that the larger the block of data, the easier to attain a specific accuracy.

Finally by using the estimated parameters, the eigenvectors associated to each natural frequency can be estimated. The results are the drivetrain ROM parameters and dynamic properties. The DT model parameters and the drivetrain dynamic properties are updated over the time by using the online measurements, to be used for fault diagnosis and prognosis purposes. The estimated drivetrain dynamic properties supported with the model parameters can be used for drivetrain fault diagnosis algorithm as discussed in [25], where the estimated DT model is used first to suuport modal estimation, and then to establish analytical features and threshold for fault

detection by establishing the relationship between the dynamic properties and physical variations/faults in system. *In this paper*, the estimated DT model is used to provide inputs for the proposed drivetrain lifetime monitoring approach. Therefore, the estimated model feeds the load observers designed for estimation of load in main and high-speed shafts as elaborated in the following.

2.3. Drivetrain remaining useful lifetime monitoring

The possibility of estimating load in the different drivetrain components depends on the model complexity and its capacity in representing the internal dynamics. By using the 3-DOF model it is possible to estimate the real-time loading and equivalent stress on the drivetrain main and high-speed shafts, which helps to estimate the residual life of these shafts by physics-based estimation of fatigue damage. Stiffness-related faults and more specifically shaft cracks are among the prevalent and influential system-level faults (see [18]) which are selected in this study, and predictive maintenance algorithm is developed accordingly. The procedure used for online estimation of stress and the shafts degradation by using the model estimated in Section 2.2 is summarized in Fig. 4. The algorithm consists of three main components, namely the estimation of loads and equivalent stress by using real-time ROM model and measurements and statistical algorithms to mitigate the impact of uncertainties; analyzing the stress signal properties which helps to be able to estimate the expected value of damage and determine confidence interval; and stochastic physics-based degradation model in time-domain for online estimation of fatigue damage. These components are described in detail in the following.

2.3.1. Estimation of load and stress

The estimation of load and stress in the drivetrain components is based on real-time operational data and the online estimated ROM model to design the load observers. Then the load observers are used to calculate stress and subsequently damage. Dependent on the degree of the DT model, the algorithm can be adjusted for lifetime monitoring of different components of the drivetrain. The algorithm shown is based on 3-DOF DT model and is able to estimate the online load in the main and high-speed shafts and measures RUL based on the loading on the shafts. As can be seen, two torque observers are designed to estimate the main and high-speed shafts torques by using the torsional measurements and updated values of torsional stiffness parameters from the DT model. By using the estimated loads by the observers, the maximum equivalent stress throughout the two shafts is calculated. The latter provides sufficient inputs for estimation of the fatigue damage and residual life of the two shafts.

In order to account for the statistical uncertainties due to both the measurements noise and estimation error of model parameters, the torsional measurements as the inputs of the load observers are estimated by using Kalman filtering. This technique is used to preprocess the input torsional response applied to the degradation model, by cancellation of background noises and the model uncertainties. The assumed linear state-space model as the input numerical model applied to the Kalman filter is obtained by applying the generalized coordinate approach as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{w}, \quad (\textit{State equation}) \tag{21a}$$

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v}, \quad (\textit{Measurement equation}) \tag{21b}$$

$$\mathbf{x} = \begin{bmatrix} \dot{\boldsymbol{\theta}} \\ \boldsymbol{\theta} \end{bmatrix} \mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{J}^{-1}\mathbf{K} & -\mathbf{J}^{-1}\mathbf{C} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{\theta}} \\ \boldsymbol{\theta} \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{J}^{-1} \end{bmatrix}, \tag{21c}$$

where $\mathbf{x}$ and $\mathbf{z}$ represent states and measurements, respectively. $\mathbf{A}, \mathbf{B}$ and $\mathbf{H}$ are describing the expected relation between the measurements, states, and inputs. $\mathbf{H}$ is the identity matrix. $\mathbf{w}$ and $\mathbf{v}$ are Gaussian, uncorrelated noise sources, which correspond to the uncertainty on the model and measurements, respectively. The desired result is an improved estimate of the system state vector $\mathbf{x}$ by realization of discrete Kalman filter. At each discretized time step k, the following procedure is performed to estimate the torsional response by considering the uncertainties in the model and input measurements.

$$\mathbf{x}(k) = \mathbf{A}\mathbf{x}(k-1) + \mathbf{B}\mathbf{u}(k-1) s\mathbf{P}(k) = \mathbf{A}\mathbf{P}(k-1)\mathbf{A}' + \mathbf{Q} \tag{22}$$

where $\mathbf{x}(k)$ and $\mathbf{P}(k)$ are the state vector and its covariance matrix at k, respectively. $\mathbf{A}'$ is the transpose of matrix $\mathbf{A}$. Then the Kalman gain is calculated, and subsequently, the estimated states are corrected by the following equations and the calculated Kalman gain factors $\mathbf{G}$ as

$$\mathbf{G} = \mathbf{P}\mathbf{H}^{'}(\mathbf{H}\mathbf{P}\mathbf{H}^{'} + \mathbf{R})^{-1}, \ \mathbf{x} = \mathbf{x} + \mathbf{G}(\mathbf{z} - \mathbf{H}\mathbf{x}), \ \mathbf{P} = \mathbf{P} - \mathbf{G}\mathbf{H}\mathbf{P}. \tag{23}$$

In the above equations, $\mathbf{Q}$ is the covariance of process noise, and $\mathbf{R}$ is the covariance of measurements noise. The Kalman filtering operation is summarized in Fig. 4. The measurement and process noises are both modeled by Gaussian distributions. The Eq. 23 is used to calculate the proper value of Kalman filter gain which is then used to attain the improved estimation of the angular displacements by considering the system uncertainties.

The next step is the design of load observers for the main and high-speed shafts. The torsional moment on the low- and high-speed shafts is estimated based on the following equations

$$\mathrm{T_{LSS}} = \mathrm{k_{LS}}(\theta_{\mathrm{r}} - \theta_{\mathrm{gr}}) + \mathrm{c_{LS}}(\Omega_{\mathrm{r}} - \Omega_{\mathrm{gr}}), \tag{24a}$$

$$T_{HSS} = k_{HS}\left(\theta_{gr} - \theta_{gn}\right) + c_{HS}\left(\Omega_{gr} - \Omega_{gn}\right), \tag{24b}$$

where T_{LSS} and T_{HSS} are the equivalent torque on the low- and high-speed shafts, respectively. $\theta_r, \Omega_r, \theta_{gr}, \Omega_{gr}, \theta_{gn}$ and Ω_{gn} are the angular displacement and velocity on rotor, gearbox and generator, respectively. Therefore, the implementation of the above torque observers need an additional torsional measurement installed on the drivetrain system. However, the approximated values can be provided by approximating the torsional response of gear in Eq. 24a and 24b respectively with the scaled torsional response of generator and scaled torsional response of rotor. The estimated loads are used as the input to estimate the stress on the shafts. The employed stochastic model-based approach for online estimation of the expected value of the residual life is based on estimation of fatigue damage due to shear/torsional stress which plays the main role in the fatigue failure of the shafts in wind turbines and bending stress due to the shaft weight. In other words, deformation and fracture from simultaneous influence of shear and bending stress is emphasized, which is based on crack blunting and crack propagation on a plane of maximum shear stress on the shaft [29]. The shear stress of the shaft can be calculated by

$$\tau = \frac{Tc}{J}, \tag{25}$$

where T is the torsional moment which is estimated in real-time by Eq. (24), c is the radial distance from the shaft center line, and J is the polar moment of inertia around the shaft axis defined by

$$J = \frac{\pi}{2}\left(d_o^4 - d_i^4\right), \tag{26}$$

where d_i and d_o are respectively the inner and outer radius of the shaft. Therefore, the maximum shear stress happens on the outer surface of the shaft. Due to the large mass of main shaft especially in high-power applications, the bending stress can also take a significant value which can be calculated by

$$\sigma = \frac{Mc}{J}, \tag{27}$$

where M is the bending moment of the shaft due to the shaft weight. By using a distributed mass model, the maximum bending moment happens in the center point of the shaft on the surface area, as can be calculated by

$$M^{max} = \frac{WL^2}{8}, \tag{28}$$

where W is the mass per unit length of the shaft, and L is the length of the shaft. Therefore, the maximum bending and shear stresses happen simultaneously on the surface in the middle of the shaft as

$$\tau_{LSS}^{max} = \frac{2}{\pi}\frac{d_o^{LSS}\left(k_{LS}\left(\theta_r - \theta_{gr}\right) + c_{LS}\left(\Omega_r - \Omega_{gr}\right)\right)}{\left(d_o^{LSS}\right)^4 - \left(d_i^{LSS}\right)^4}, \tag{29a}$$

$$\tau_{HSS}^{max} = \frac{2}{\pi}\frac{d_o^{HSS}\left(k_{HS}\left(\theta_{gr} - \theta_{gn}\right) + c_{HS}\left(\Omega_{gr} - \Omega_{gn}\right)\right)}{\left(d_o^{HSS}\right)^4 - \left(d_i^{HSS}\right)^4}, \tag{29b}$$

$$\sigma^{max} = \frac{WL^2 d_o}{4\pi\left(d_o^4 - d_i^4\right)}. \tag{29c}$$

The equivalent stress is calculated by applying von Mises theory. The equivalent stress due to the combined bending and torsion is maximum in the middle of the shaft on the surface. Von Mises stress under combined bending and torsion loading is calculated by [44]

$$\sigma_d = \sqrt{\sigma_{max}^2 + 3\tau_{max}^2}, \tag{30}$$

where σ_d is von Mises stress, and τ_{max} and σ_{max} are the maximum torsional and bending stresses, respectively.

2.3.2. Classification of stress signal for estimation of average damage

Statistical analysis of damage consists of two main steps: First, classification of the sources of uncertainty and a proper way of modelling to address them in the degradation model. Second, analyzing the stress signal type and the specific properties of that type of signal, which helps to determine confidence interval for average and variance of damage. Based on Eq. (29), for determining the stress signal properties, one should look into the torsional response signal. In order to evaluate the type of torsional response signal, the spectral moments of the signal and the bandwidth parameters are evaluated. The spectral moments λ_m of the autospectral density function S both in discrete-frequency horizon are defined as

$$\lambda_m = \sum_{j=0}^{N/2}\left(j^m S\left(j\right)\right), \tag{31a}$$

$$S\left(j\right) = \frac{2(\sigma_d(j)\sigma_d(j)^*}{N},\ j = \left(0, \ldots, \frac{N}{2}\right), \tag{31b}$$

where S is the power spectral density of stress signal σ_d. The bandwidth parameters α_1 and α_2 have a critical role to specify the type of signal σ_d, which are defined by

$$\alpha_1 = \frac{\lambda_1}{\sqrt{\lambda_0\lambda_2}},\quad \alpha_2 = \frac{\lambda_2}{\sqrt{\lambda_0\lambda_4}}. \tag{32}$$

For a strictly narrow-band/harmonic signal, these two metrics tend to 1. However, for a strictly broad-band/independent signal, they tend to 0. Even though the drivetrain torsional response may seem as a narrow-band signal with the rotor revolution frequency as the characteristic frequency, our extensive observations on both simulation and operational data show that the two metrics tend to zero under different turbine operating conditions.

In the literature, the drivetain vibration responses are considered as both stationary and non-stationary signals, [45,46]. The possibility of fitting different distributions to the drivetrain response is investigated by researchers. Ghane et al.[47] proposes t-distribution as the best fit for the transnational vibration measurements captured from drivetrain. It is a common practice to model the drivetrain loads and responses as Rayleigh, Weibull and the generalized gamma distributions with reference to the distribution of mean wind speed. For data-driven fault detection of wind turbine drivetrain system, it is common to assume the operational vibration measurements as a stationary Gaussian process at each mean wind speed, which means each sensor measurement follows a Gaussian distribution [48]. Our observations on the drivetrain system torsional response show that the data at each operating condition follows a Gaussian distribution or in other words is identically distributed. The parameters of the distribution tend to vary as the operational speed changes. Therefore, in different operating zones/ranges of input wind speed, it can be assumed that the response follows a Gaussian distribution which its parameters vary in transition between different zones. The method can use an additional input from the operational speed to accordingly select the relevant parameters of the associated distribution. Based on this explanation, the torsional response is assumed to be IID. The metric α_2 is more commonly used. From mathematical perspectives, this metric for a stationary signal is equal to the cross correlation between the signal and its second derivative, which is 1 for a narrow-band and 0 for a broad-band signal. As discussed earlier, torsional response for each operating speed shows a stationary Gaussian behavior. Stress as a linear function of response shows the same pattern as expected. For a particular case of stationary-Gaussian process (our case), the metric α_2 is equivalent to the irregularity factor, *IF*, defined as

$$IF = \frac{{s_{\dot{\sigma}_d}}^2}{s_{\sigma_d} s_{\ddot{\sigma}_d}}, \tag{33}$$

where $s_X, s_{\dot{X}}$ and $s_{\ddot{X}}$ are the standard deviations of the signal and its first and second derivatives.

The response is IID and so does the stress signal as a linear function of torsional response. For such a signal, it is possible to apply crude Monte Carlo to obtain CLT-based approximate confidence interval [41] for the fatigue damage as is discussed in the next part. For some applications which the stress signal demonstrates narrow-band properties, the expected value and variance of damage can be estimated by using the theory explained by [18].

2.3.3. Degradation model for residual life estimation

The estimation of fatigue damage and residual life of the shaft is performed by using stress-life method. Even though the stress life method is not the most accurate approach, this method is able to represent high-cycle applications adequately. The criterion for high-cycle fatigue is $N > 10^3$ over the component lifetime, which is the case for a wide range of applications *e.g.* wind turbines and ship propulsion systems. The estimated online time series of maximum stress feeds the time-domain cycle counting approach based on rainflow approach. Therefore, the rainflow cycle counting and Goodman rule are used to calculate the effective stress and the number of cycles at each stress level. Then damage is estimated in real-time and the shaft RUL is calculated. For this purpose, Miner's rule is used to calculate the accumulated damage and subsequently residual life of the drivetrain components.

The degradation model for the main shaft is elaborated in the following. The number of stress cycles at different stress levels is counted by using the time-domain rainflow cycle counting approach [50]. Cycle counting is especially important for the broad-band stress signal to distinguish small cycles which are interruptions of larger ones. Rainflow method has shown proven performance in time domain analysis of stress signal to count the stress cycles. The outputs are the amplitude stress levels σ_s , and the number of stress cycles at σ_s *for* $s = (1, \ldots, S)$. In order to consider the influence of nonzero mean stress level, Goodman rule is employed to calculate the effective stress (the equivalent zero mean alternating stress) by the equation [51]

$$\sigma_s^e = \frac{\sigma_s}{1 - \frac{\sigma_m}{\sigma_u}},\ \forall\ s \in \left\{1, \ldots, S\right\}, \tag{34}$$

where σ_m and σ_u are the mean stress and material yield strength, respectively. The accumulated damage for the data block t with S different stress levels $\sigma_s^e(s \in \{1, \ldots, S\})$ is calculated by using Miner's rule as

$$d^t = \sum_{s=1}^{S} \frac{n_s}{N_s}, \tag{35}$$

where n_s is the number of cycles at the stress level σ_s^e and N_s is the number of cycles to yield at stress level σ_s^e, where the relationship is defined by S-N curve characteristic as

$$\sigma_s^e = a(2N_s)^b. \tag{36}$$

The absolute total online accumulated damage can then be calculated by

$$D = \sum_{t=1}^{Time} d^t, \tag{37}$$

where *Time* stands for the last data block which represents the current time. Note that d^t represents the "short-term" damage, since is related to only a specific environmental condition. However, damage D characterizes the long-term damage provided having measurements from all possible environmental conditions [42]. The described method can also be used for estimation of relative damage between different operational periods over the time, to give an insight on variations in degradation between different operational periods. It can also be used to estimate the relative damage between the different drivetrain components at the time, to give the operator a sense about the most vulnerable parts of the system at different operational periods.

The described deterministic approach does not necessarily provide precise results. Stochastic models which can adequately address the uncertainties in the aforedescribed stress-life degradation estimation approach can be used in conjunction with the above described deterministic model-based approach to improve the accuracy. According to Eq. (35), the two sources of uncertainty are evolved out of the stress calculation and the material property captured by S-N curve. To address the stress calculation uncertainty, a significant part of uncertainty arises from the estimation of load. In order to account for uncertainties in load estimation approach, the influence of measurement noise and uncertainties in estimated model parameters were mitigated by employing Kalman filtering. However, the other significant source of uncertainty arises from stress life method and the procedure employed to obtain S-N curve parameters a and b which represent the best-fitting estimates of experimental fatigue data [43]. In this work, the statistical uncertainty in fatigue calculation due to material uncertainties accounted for by assuming the damage at each stress level as a random variable. In other words, in order to estimate damage for the t^{th} data block of the stress timeseries, d^t, the damage is calculated by assuming the S-N curve parameters a and b as random variables, so that for each stress level σ_i, the number of cycles to failure are estimated by randomly selecting a and b in the intervals ±5% of their nominal values.

Benasciutti et al. [52] suggests closed form approximations for the expected value of damage for both Gaussian narrow-band and broad-band processes. In this research, for the case that stress time series are IID, which is a practical assumption based on the observations and measurements of the under consideration test case and the analysis of stress signal spectral moments in Section 2.3.2, in order to obtain confidence interval for average damage, Monte Carlo simulation is employed and the results of IID CLT are applied, [49,41]. The above procedure for stochastic estimation of fatigue damage is repeated k different times with the same data block to realize k different cases of $d_i^t \in \{d_1^t, \ldots, d_k^t\}$ to attain the confidence interval for the average damage. As a result, with the confidence 95%, the average damage is placed in the following interval

$$\left[\widehat{\mu}_{d^t} - 1.96 s_{d^t} \Big/ \sqrt{(k)} \quad \widehat{\mu}_{d^t} + 1.96 s_{d^t} \Big/ \sqrt{(k)}\right], \tag{38a}$$

$$\widehat{\mu}_{d^t} = \frac{1}{k} \sum_{i=1}^{k} d_i^t, \tag{38b}$$

$$s_{d^t} = \sqrt{\frac{1}{k-1} \sum_{i=1}^{k} \left(d_i^t - \widehat{\mu}_{d^t}\right)^2}, \tag{38c}$$

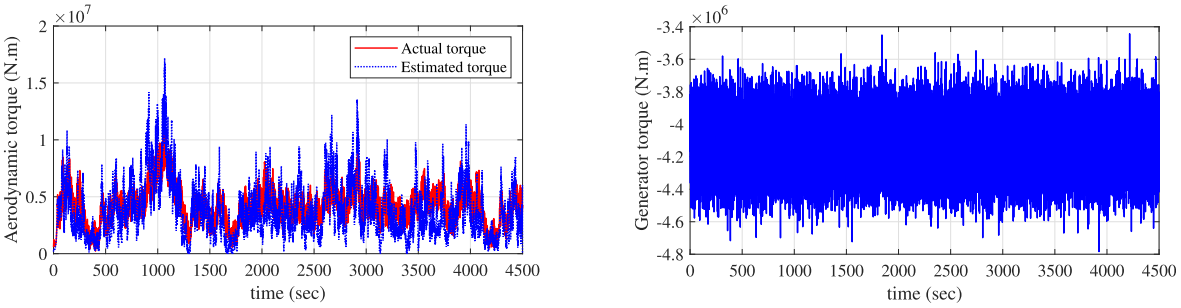
where the value of k is selected to realize an interval ±5% around the mean value of d_i^t. $\widehat{\mu}_{d^t}$ and s_{d^t} are the average and variance estimates.

In order to account for the maintenance operations in the proposed digital twin approach, a feedback from the maintenance about the component replacement/repairment is applied to the stress and damage estimation algorithms as shown in Fig. 4. It is worth noting that improper maintenance, manufacturing errors, and the influences of corrosion or other oil/grease contamination can be seen by this approach, because they have all a potential to influence the drivetrain online vibration response. The real-time vibration response is the main input of the proposed digital twin algorithm which influences the estimated stress and subsequently the estimated degradation.

Using the available physical models, integration of them in a unified framework, proposing signal processing techniques for estimating these models and model inputs in real-time from available measurements, optimizing the data streaming between models, continuous processing architectures, and using statistical approaches and stochastic modeling techniques first to model and then to mitigate the impact of uncertainties is the functionality of digital twin in this paper.

Table 1
Environmental conditions for drivetrain analysis.

Environmental condition	EC1	EC2	EC3
$U_w\ (m/s)$	7	9	11
$H_s\ (m)$	2.5	3.5	3.5
$T_p\ (s)$	6.5	7.5	7.5



(a) T_a vs. time (b) T_g vs. time

Fig. 5. Drivetrain model input loads. (a) Aerodynamic torque. (b) Generator torque.

3. Simulation studies

In this section, the main test case and different environmental conditions for simulation studies; the results of applying the proposed drivetrain equivalent model identification algorithm in different test scenarios, namely normal, faulty and overload conditions by using both the actual and estimated values of drivetrain input loads; the results of estimating load and equivalent stress in the main shaft by using the designed online load observers in the different test scenarios; and the results of estimating the expected value of fatigue damage of the main shaft by using the proposed degradation model in the different test scenarios are presented.

3.1. Test case

The methodology proposed for the residual life monitoring of the drivetrain components is simulated and tested to monitor the RUL of the main shaft of the drivetrain system in DTU 10 MW wind turbine. The shaft dimensions and weight are taken based on the work performed by Wang et al. [53]. The shaft material properties are obtained from [33]. Environmental conditions for the global simulations are summarized in the Table 1. The system identification algorithm is tested for different operational conditions, namely below rated, rated and over rated conditions.

3.2. Drivetrain equivalent model estimation

The drivetrain model input load calculated from global simulations, the estimated value of aerodynamic torque from the turbine measurements by using the algorithm in Fig. 3, and the results of applying the proposed drivetrain equivalent ROM parameters estimation by using the algorithm in Fig. 2 in different test scenarios, namely normal, faulty and overload conditions by using both the actual and estimated values of drivetrain input loads are presented in the following.

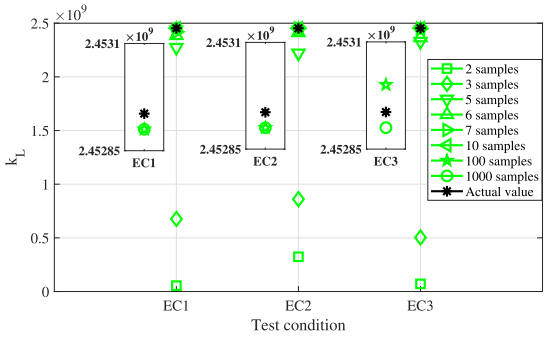
3.2.1. Estimation of the drivetrain model loads

The loads applied to the model are the aerodynamic and generator torques. A comparison between the actual and estimated aerodynamic torque is shown in Fig. 5a. The actual torque is assumed to be the torque obtained from the global simulations, and the estimated torque is the value estimated by using the turbine online measurements as explained in Section 2.2.4. The applied generator torque is also shown in Fig. 5b. For demonstration purposes, the generator torque is up-scaled by applying the gear ratio.

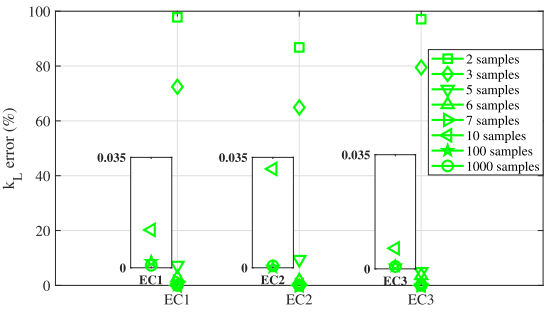
3.2.2. ROM Parameter estimation by using torsional measurements

Here, the possibility of using the proposed algorithm for identifying the model parameters of 3-DOF torsional DT by using the real-time torsional measurements is investigated. This model will then be used for monitoring the variations in stiffness and subsequently feed the load observers designed in the drivetrain shafts. The actual values of the drivetrain 3-DOF model parameters and the estimated non-rigid natural frequencies by using the angular velocity error function are reported in [22]. The validation criterion for the estimated model parameters is the relative error percentage to be less than 5%.

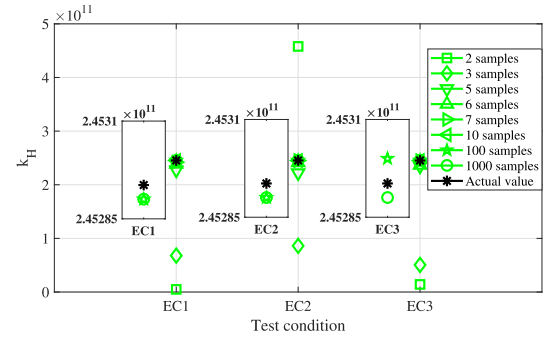
In order to identify this model, five parameters, namely J_r, J_{gr}, J_{gn}, k_L and k_H which are respectively the equivalent rotor inertia,



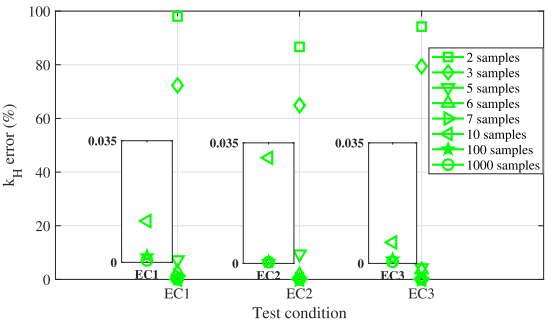
(a) Estimated main shaft stiffness vs. EC_i



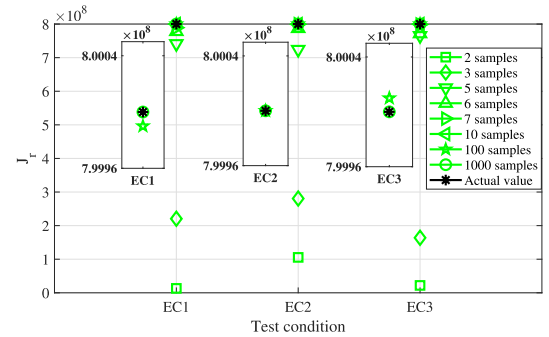
(b) Error in estimated main shaft stiffness vs. EC_i



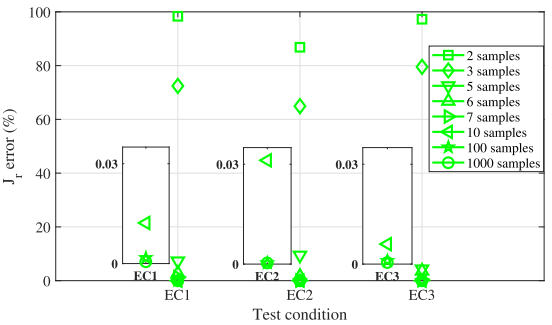
(c) Estimated high-speed shaft stiffness vs. EC_i



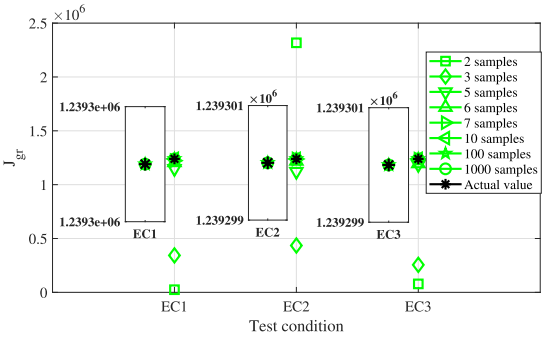
(d) Error in estimated high-speed shaft stiffness vs. EC_i



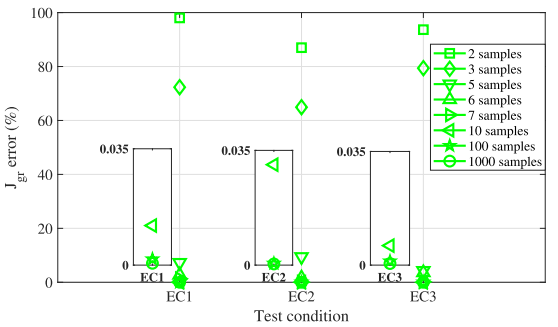
(e) Estimated rotor inertia vs. EC_i



(f) Error in estimated rotor inertia vs. EC_i

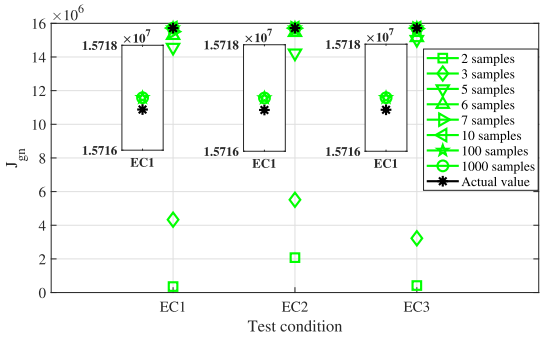


(g) Estimated gearbox inertia vs. EC_i

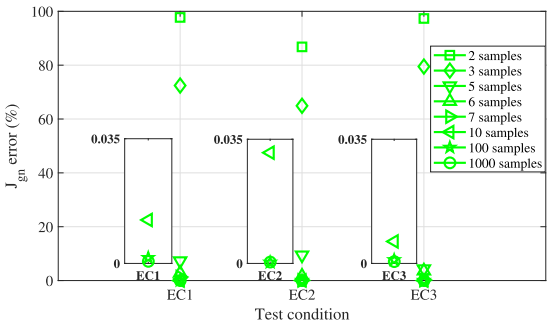


(h) Error in estimated gearbox inertia vs. EC_i

Fig. 6. Estimated ROM model parameters, by using the actual aerodynamic torque. (a) k_L . (b) k_L relative error. (c) k_H. (d) k_H relative error. (e) J_r. (f) J_r relative error. (g) J_{gr}. (h) J_{gr} relative error. (i) J_{gn}. (j) J_{gn} relative error.



(i) Estimated generator inertia vs. EC_i



(j) Error in estimated generator inertia vs. EC_i

Fig. 6. (*continued*).

gearbox inertia, generator inertia, main shaft stiffness and high-speed shaft stiffness should be specified. The error of these parameters estimation versus the number of response samples (2, 3, 5, 7, 10, 100, 1000) for the three different drivetrain operational conditions (EC1, EC2 and EC3) by using both the actual input torque and the estimated torque is investigated in here. The procedure explained in Section 2.2.5 is followed to provide an analytical criterion/margin of the error for the proposed parameter estimation approach. With respect to the fact that the error time series at each test case meets the conditions of IID CLT, in order to attain a reliable value of error in each test case, the number of data blocks for each test case is selected to ensure that with the confidence 99%, the error places in the interval $\widehat{\mu}_l \pm 2.58\frac{\hat{s}}{\sqrt{l}}$. Therefore, the reported estimated parameters and the associated errors are the most expected values with the specified confidence interval. For the case of fewer number of samples, the variance of errors is higher, which need higher number of data blocks to realize the specified confidence level. In order to realize the confidence 99% at the interval ±5% around the average estimate of error, the minimum required number of data blocks at each test is calculated by $l > 2662.56\left(\frac{\hat{s}}{\hat{\mu}}\right)^2$.

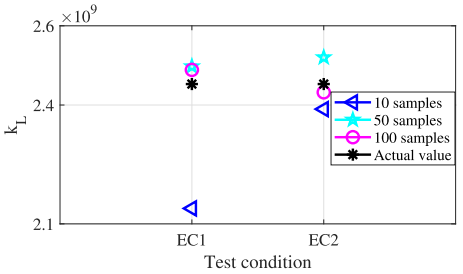
The Fig. 6 shows the expected value of estimated parameters and estimation errors versus different environmental conditions and different numbers of input samples. As it can be seen, the expected value of error is less than 1% when the number of algorithm input samples are more than 5. By increase of the number of samples, the error tends to zero. In addition, the method is not sensitive to the turbine operational conditions, and demonstrates a similar performance under different environmental conditions. By reduction of the number of input measurements samples from 1000 to 10, the computational time reduces by about 70%. When the number of variables (the order of ROM) increases, the improvements in the computation time by reduction of the number of samples will be more significant. The use of 10 samples with the sampling frequency 300 Hz, leads to the estimation of parameters in only fraction of a second, which shows that the algorithm can be executed in real-time.

The Fig. 7 illustrates more realistic values for estimation error vs. number of samples, by using the estimated torque instead of the actual torque as the input of parameters estimation approach. The latter is to address the uncertainty in the input torque for the proposed system identification algorithm. The estimated torque is calculated in real-time by an aerodynamic torque observer based on the turbine and blades online information by using the theory explained in Section 2. The values of error shown in Fig. 7 are again the expected values of error. As it can be seen, for the case of using the estimated torque, the method needs more input data to maintain the 5% threshold set for the estimation error. The reason is that in case of few samples, the LS estimator is more sensitive to the error of torque estimation of the individual samples. The error tends mostly to reduce by increase of the input samples.

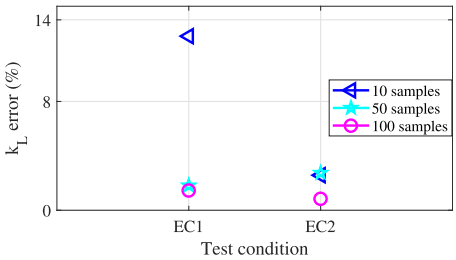
The Fig. 7 estimation of drivetrain parameters in monitoring the variations of the model parameters which can model five different categories of faults, namely main shaft stiffness, high-speed shaft stiffness, rotor inertia, gear inertia and generator inertia sensitive faults. In Fig. 8a, the main shaft stiffness is reduced by 20%, in Fig. 8c, the generator shaft stiffness is reduced by 20%, in Fig. 8e, the rotor inertia is increased by 10%, in Fig. 8g, the gearbox inertia is increased by 10%, and in Fig. 8i, the generator inertia is increased by 10%. These simulation cases are designed to evaluate the capability of the proposed algorithm in tracking the variations of the system parameters which are representing different system-level fault cases. Variations in each equivalent model parameter can represent a specific class of the drivetrain faults. The results show that the proposed DT model parameter estimation algorithm can track/measure the variations in system perfectly. Our observations also show that the algorithm is also not sensitive to the sampling frequency. However, the sufficient sampling frequency will ensure the observability of higher frequency modes as the input of model estimation approach to be able to realize higher DOF models.

3.3. Estimation of stress and damage

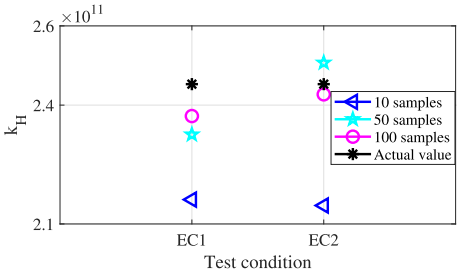
The results of applying the online load observer and equivalent stress estimation of the main shaft based on the theory developed in Section 2.3.1 in different test scenarios, namely normal, faulty and overload conditions by using both the actual and estimated values of drivetrain input loads; and the results of estimating the expected value of fatigue damage of the main shaft by using the proposed degradation model by using the algorithm in Fig. 4, in different test scenarios are presented in the following.



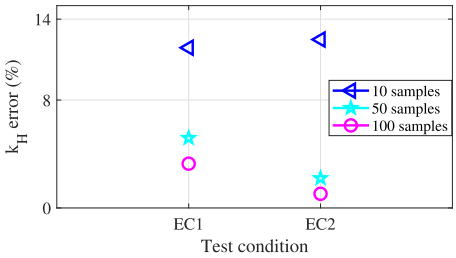
(a) Estimated main shaft stiffness vs. EC_i



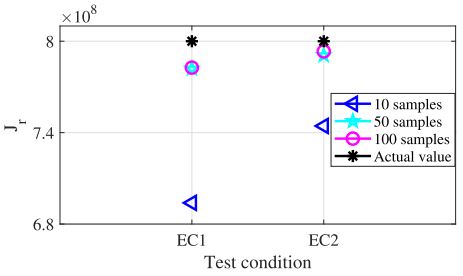
(b) Error in estimated main shaft stiffness



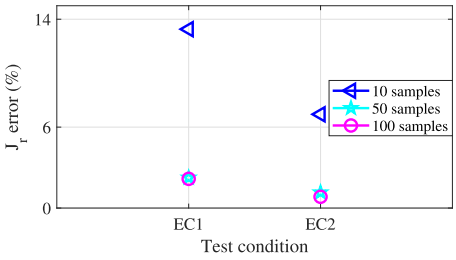
(c) Estimated high-speed shaft stiffness vs. EC_i



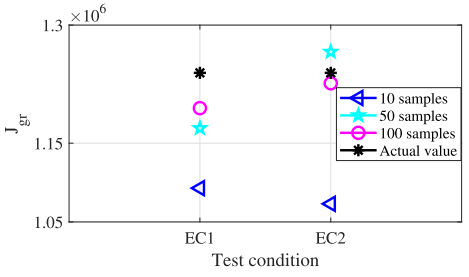
(d) Error in estimated high-speed shaft stiffness



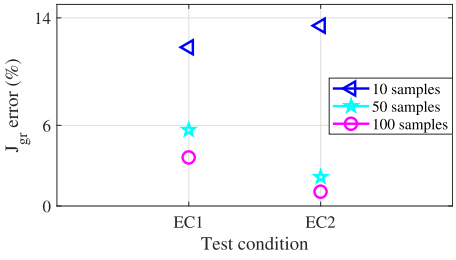
(e) Estimated rotor inertia vs. EC_i



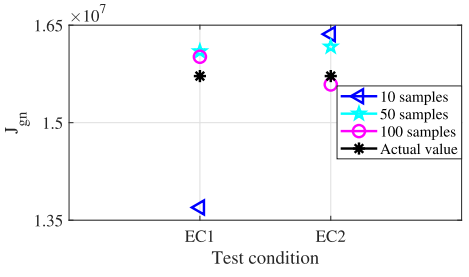
(f) Error in estimated rotor inertia vs. EC_i



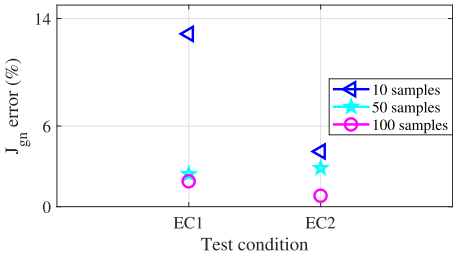
(g) Estimated gearbox inertia vs. EC_i



(h) Error in estimated gearbox inertia vs. EC_i

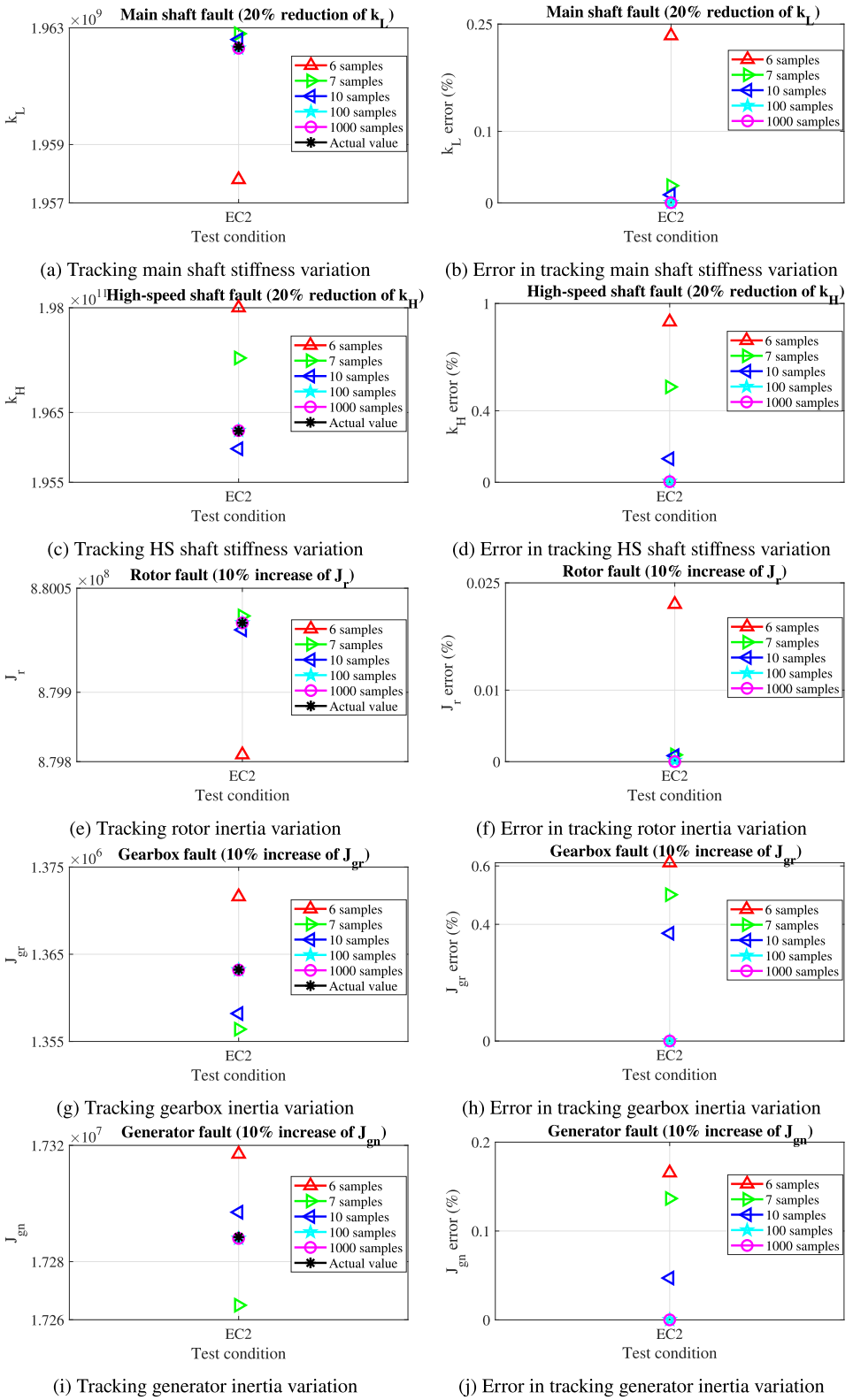


(i) Estimated generator inertia vs. EC_i



(j) Error in estimated generator inertia vs. EC_i

Fig. 7. Estimated ROM model parameters, by using the actual aerodynamic torque. (a) k_L . (b) k_L relative error. (c) k_H. (d) k_H relative error. (e) J_r. (f) J_r relative error. (g) J_{gr}. (h) J_{gr} relative error. (i) J_{gn}. (j) J_{gn} relative error.



(a) Tracking main shaft stiffness variation

(b) Error in tracking main shaft stiffness variation

(c) Tracking HS shaft stiffness variation

(d) Error in tracking HS shaft stiffness variation

(e) Tracking rotor inertia variation

(f) Error in tracking rotor inertia variation

(g) Tracking gearbox inertia variation

(h) Error in tracking gearbox inertia variation

(i) Tracking generator inertia variation

(j) Error in tracking generator inertia variation

Fig. 8. Estimated ROM model parameters in different fault scenarios (Test condition: EC_2). (a) k_L . (b) k_L relative error. (c) k_H. (d) k_H relative error. (e) J_r. (f) J_r relative error. (g) J_{gr}. (h) J_{gr} relative error. (i) J_{gn}. (j) J_{gn} relative error.

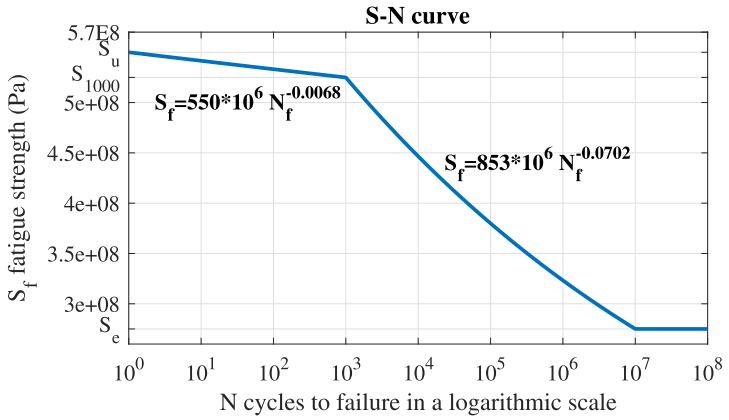


Fig. 9. Main shaft steel S-N curve.

3.3.1. *Shaft material S-N curve*

According to EN10083 [34], the S-N curve of 42CrMo4 which is commonly used as the main shaft material, with the specified finite-life region consisting of low-cycle and high-cycle fatigue regions, is shown in Fig. 8. The infinite-life region of the shaft is characterized by 10^{11} number of cycles [29].

3.3.2. *Estimation of main shaft load and stress in normal operations*

The time series of the estimated main shaft load and the equivalent von Mises stress by using the proposed algorithm, for 200 s of normal operations, are shown in the Figs. 9 and 10. Uncertainties in input real-time measurements and estimated model parameters are accounted for by using zero mean Gaussian distributions and added to the system of equations. Figs. 10a and 11a are the actual values of load and stress on the main shaft. Figs. 10b and 11b are the load and stress when the actual value of aerodynamic torque is applied as input to the DT algorithm. Figs. 10d and 11d are the errors calculated for this case. As it can be seen, the estimation error is negligible when the exact value of aerodynamic torque is available.

Figs. 10c and 11e are the load and stress when the aerodynamic torque applied as input to the DT algorithm is not available, but the estimated value of real-time aerodynamic torque is calculated and applied to DT model. Figs. 10e and 11e are the errors calculated for this case. As it can be seen, even though the error increases, it is still within the 5% interval considered as limit for the relative error of the DT model estimations in the work.

3.3.3. *Estimation of load and stress in overload condition*

The overload is modeled with 20% increase in the drivetrain torque. The time series of the estimated main shaft load and the equivalent stress by using the proposed algorithm, are shown in the Figs. 9 and 10.

Figs. 10a and 11a are the actual values of load and stress on the main shaft. Figs. 10b and 11b are the estimated load and stress when the actual value of aerodynamic torque is applied as input to the DT algorithm. Figs. 10d and 11d are the errors calculated for this case. As it can be seen, the estimation error is negligible when the exact value of aerodynamic torque is available. Figs. 10c and 11c are the estimated load and stress when the aerodynamic torque applied as input to the DT algorithm is not available, but the estimated value of real-time aerodynamic torque is calculated and applied to DT model. Figs. 10e and 11e are the errors calculated for this case. As it can be seen, the proposed DT error in estimation of both load and stress in overload conditions is lower than normal and faulty operations.

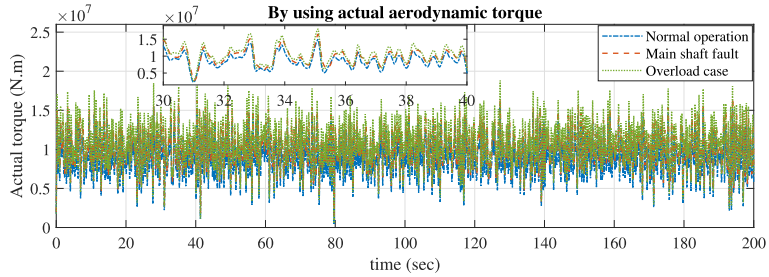
3.3.4. *Estimation of load and stress in main shaft fault condition*

The early stage fault in the main shaft is modeled with 10% reduction in the shaft stiffness. The time series of the estimated main shaft load and the equivalent stress by using the proposed algorithm, are shown in the Figs. 9 and 10.

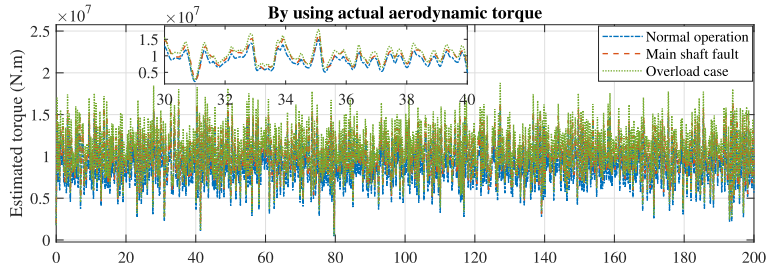
Figs. 10a and 11a are the actual values of load and stress on the main shaft. Figs. 10b and 11b are the estimated load and stress when the actual value of aerodynamic torque is applied as input to the DT algorithm. Figs. 10d and 11d are the errors calculated for this case. As it can be seen, the estimation error is negligible when the exact value of aerodynamic torque is available. Figs. 10c and 11c are the estimated load and stress when the aerodynamic torque applied as input to the DT algorithm is not available, but the estimated value of real-time aerodynamic torque is calculated and applied to DT model. Figs. 10e and 11e are the estimation errors of load and stress in this test scenario. The estimation error in case of main shaft fault shows a higher value compared to both the normal operation and overload condition.

3.3.5. *Estimation of damage in different operational conditions*

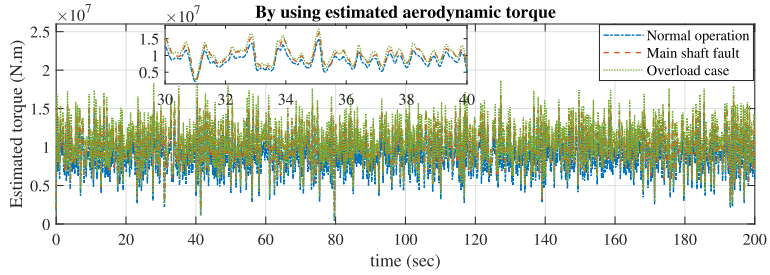
In order to address the material uncertainty in damage calculations, for each stress data block, k different pairs of the two coefficients of S-N curve are randomly generated in an interval $\pm 5\%$ around their nominal values. Then the confidence interval for the damage d^t can be provided by assuming that the k different cases of d^t are IID. These k cases are generated from k independent cases, and our observations show a Gaussian pattern as the number of samples increases, which shows the consistency of the results of IID



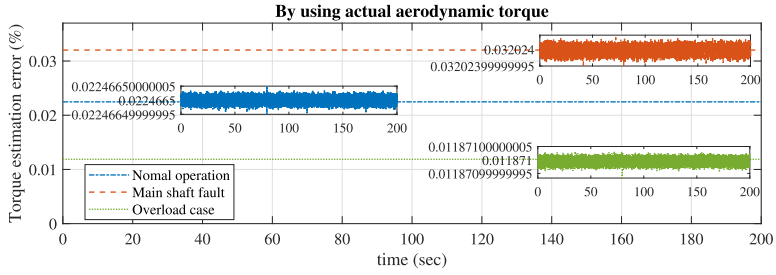
(a) Main shaft actual torque



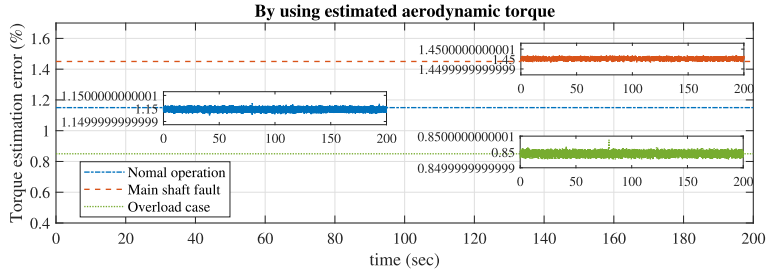
(b) Main shaft estimated torque



(c) Main shaft estimated torque

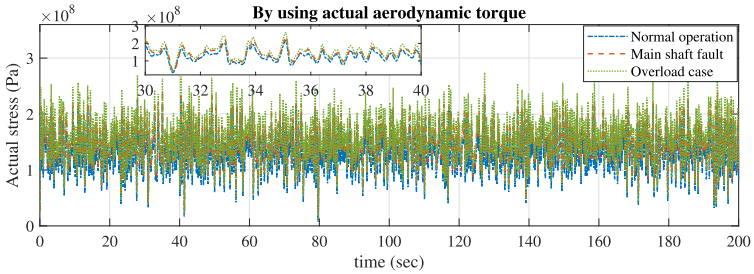


(d) Main shaft estimated torque error

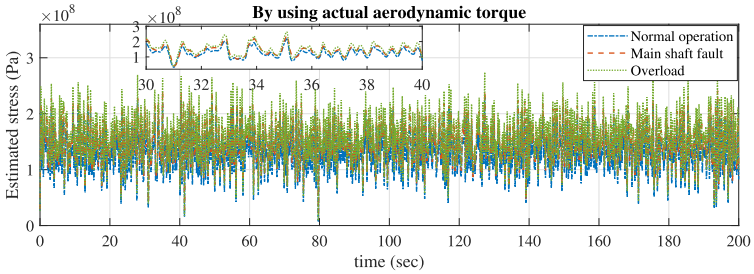


(e) Main shaft estimated torque error

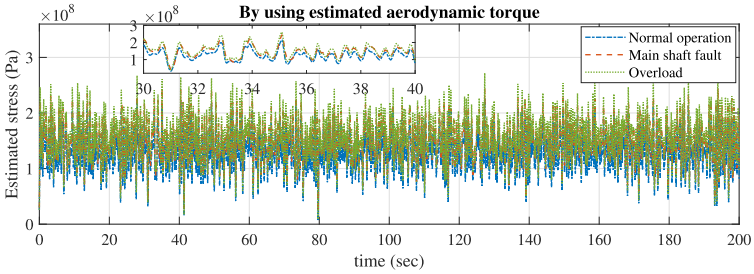
Fig. 10. Real-time torque on the main shaft. (a) Actual torque, (b) estimated torque in case of using actual aerodynamic torque, (c) estimated torque in case of using estimated aerodynamic torque, (d) torque estimation relative error in case of using actual aerodynamic torque, (e) torque estimation relative error in case of using estimated aerodynamic torque.



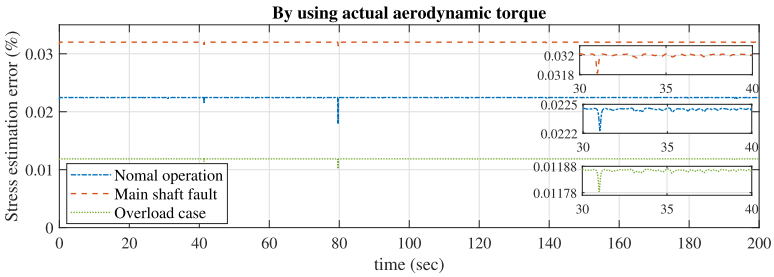
(a) Main shaft actual maximum stress



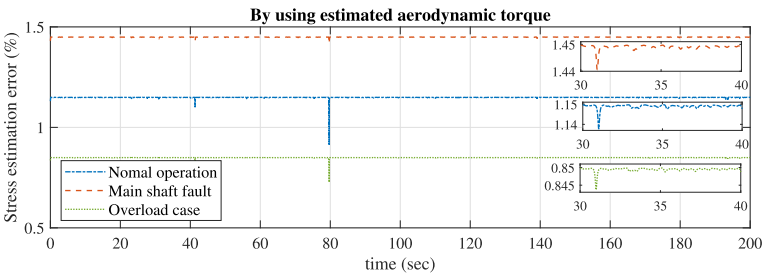
(b) Main shaft estimated maximum stress



(c) Main shaft estimated maximum stress



(d) Main shaft estimated stress error



(e) Main shaft estimated stress error

Fig. 11. Real-time equivalent von Mises stress on the main shaft. (a) Actual stress, (b) estimated stress in case of using actual aerodynamic torque, (c) estimated stress in case of using estimated aerodynamic torque, (d) stress estimation relative error in case of using actual aerodynamic torque, (e) stress estimation relative error in case of using estimated aerodynamic torque.

Table 2
Expected value of main shaft accumulated damage for different test scenarios for 3600 seconds of operation.

Test scenario	Actual damage	Estimated damage$_1$	Estimated damage$_2$
Normal operation	$1.8e^{-7}$	$1.8e^{-7}$	$1.8e^{-7}$
Overload (20%)	$1.1e^{-5}$	$1.1e^{-5}$	$1.1e^{-5}$
Fault (10%)	$8.3e^{-7}$	$8.3e^{-7}$	$8.3e^{-7}$
Overload and fault	$3.4e^{-5}$	$3.4e^{-5}$	$3.3e^{-5}$

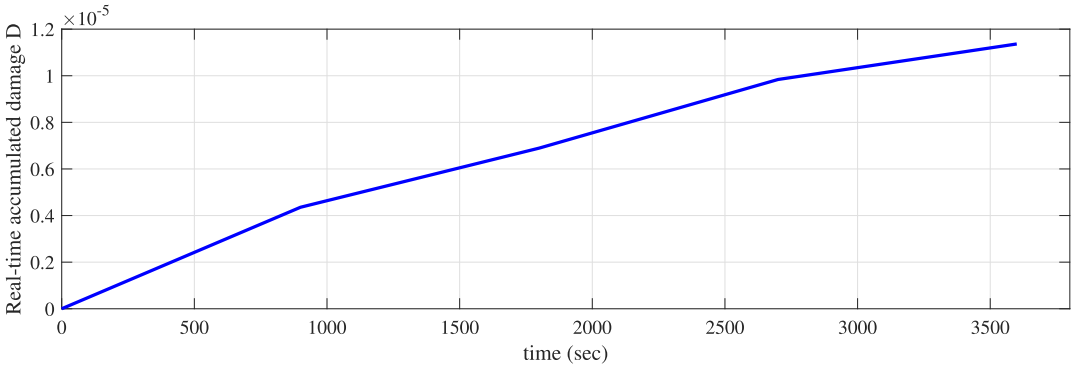


Fig. 12. Expected accumulated damage of the main shaft over one hour of operation (overload condition: 20%).

central limit theorem for d_i^t, *for* $i \in (1, \ldots, k)$. The average real-time accumulated damage by using the proposed DT approach for different operating conditions, namely normal operation, overload, main shaft fault and the combination of overload and shaft fault is listed in the Table 2. The early stage fault in the main shaft is modeled by 10% reduction of the main shaft stiffness. The overload is modeled by 20% increase of the drivetrain torque. *Actual damage* is the accumulated damage when the actual value of drivetrain parameters and input loads are accessed. *Estimated damage*$_1$ is the damage estimated by using the proposed DT model and the actual value of drivetrain input loads. *Estimated damage*$_2$ is the damage estimated when the DT model and estimated values of drivetrain input loads are used. As it can be seen, the estimated damage by the proposed approach in all the different test scenarios matches with the actual damage, with an exception for the test case which represents a combination of fault and overload, where the proposed approach slightly underestimates the damage. The real-time estimation of the accumulated damage D of the main shaft for one hour operation during the drivetrain overload is shown in the Fig. 12.

It should be noted that this work does not mean that the faults in the main shaft are the most prevalent faults in the wind turbine drivetrain systems, though it can contribute to a wide range of other secondary faults. The main purpose of focusing on the main shaft in the simulation studies is the proof of concept for the possibility of using an innovative RUL monitoring approach based on torsional vibration measurements for monitoring the residual life of the drivetrain components by using a simple model. For other components, *e.g.* the gears and bearings of the gearbox, it is still possible to use the proposed approach but by using more detailed torsional models, and probably a combination of torsional and translational models which is considered as the future work.

4. Conclusion

This work provided a basis for the development of preventive maintenance in FWT drivetrain systems based on monitoring the residual life of the components, by means of digital twin and employment of a stochastic physics-based model for determining the drivetrain components RUL. As a critical part of DT, a toolbox was proposed which receives the drivetrain torsional response and estimated input loads, and calculates the system dynamic properties (torsional natural frequencies, damping and mode shapes) and the equivalent torsional dynamic model parameters (torsional stiffnesses and moment of inertias). The application of this model in estimation of the drivetrain components residual life in real-time, and more specifically the RUL of the drivetrain main shaft was demonstrated. In order to evaluate the proposed DT approach, it was shown that by using only more than 10 data samples with the sampling frequency 300 Hz for the real-time measurements, the estimation error of DT model parameters and the estimated values of load and stress in the drivetrain components in all simulation cases, namely normal operations, fault on the main shaft and overload is always less than 5 percent.

The application of higher DOF torsional models as more detailed DT models which can capture real-time variations in mesh stiffness and inertia of individual gears and intermediate shafts, which can help to estimate the real-time load of the individual subcomponents by taking into account the components internal dynamics, and subsequently to calculate stress and fatigue damage due to different failure modes for a wide range of drivetrain subcomponents is looked as the future work.

CRediT authorship contribution statement

Farid K. Moghadam: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing - original draft. **Amir R. Nejad:** Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank John Marius Hegseth of Norwegian University of Science and Technology, Trondheim, Norway, for providing the 10 MW spar FWT global analysis simulation data, and for his support in estimation of aerodynamic torque in general.

References

[1] European Commission – Press release. Boosting Offshore Renewable Energy for a Climate Neutral Europe, https://ec.europa.eu/commission/presscorner/detail/en/IP_20_2096, (accessed 19 November 2020).
[2] P. Feldhaus, T. Vahlenkamp, Transformation of Europe's Power System until 2050 Including Specific Considerations for Germany Electric Power and Natural Gas Practice, 2010.
[3] IEA, World Energy Outlook 2020, IEA, Paris, 2020. URL: https://www.iea.org/reports/world-energy-outlook-2020.
[4] Philipp Beiter, Walter Musial, Patrick Duffy, Aubryn Cooperman, Matthew Shields, Donna Heimiller, Michael Optis, The Cost of Floating Offshore Wind Energy in California Between 2019 and 2032, Technical report, NREL, United States, 2020.
[5] A. Ioannou, A. Angus, F. Brennan, Parametric CAPEX, OPEX, and LCOE expressions for offshore wind farms based on global deployment parameters, Energy Sour. B 13 (5) (2018) 281–290.
[6] S. Pfaffel, S. Faulstich, K. Rohrig, Performance and reliability of wind turbines: a review, Energies 10 (11) (2017) 1904.
[7] P. Beiter, W. Musial, P. Duffy, A. Cooperman, M. Shields, D. Heimiller, M. Optis, The Cost of Floating Offshore Wind Energy in California Between 2019 and 2032 (No. NREL/TP-5000-77384; BOEM-2020-48). National Renewable Energy Lab. (NREL), Golden, CO (United States), 2020.
[8] M. Rausand, A. Barros, A. Hoyland, System Reliability Theory: Models, Statistical Methods, and Applications, John Wiley & Sons, 2020.
[9] D. Snieckus, https://www.rechargenews.com/wind/siemens-gamesa-unveils-digitally-souped-up-11mw-offshore-turbine/2-1-711795, 2019.
[10] P. Goossens, Industry 4.0 and the power of the digital twin, Retrieved 5 (3) (2017) 2017.
[11] S.S. Johansen, A.R. Nejad, On digital twin condition monitoring approach for drivetrains in marine applications, in: International Conference on Offshore Mechanics and Arctic Engineering, vol. 58899, American Society of Mechanical Engineers, 2019. V010T09A013.
[12] A.R. Nejad, Z. Gao, T. Moan, Fatigue reliability-based inspection and maintenance planning of gearbox components in wind turbine drivetrains, Energy Procedia 53 (2014) 248–257.
[13] L. Sethuraman, Y. Guo, S. Sheng, Mitigation of Micropitting in Wind Turbine Main Shaft Bearings (No. NREL/JA-5000-65026). National Renewable Energy Lab. (NREL), Golden, CO (United States), 2015.
[14] Z. Liu, L. Zhang, A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings, Measurement 149 (2020), 107002.
[15] E.B. Pedersen, D. Jørgensen, H.J. Riber, J. Ballani, S. Vallaghé, B. Paccaud, True Fatigue Life Calculation Using Digital Twin Concept and Operational Modal Analysis. In The 29th International Ocean and Polar Engineering Conference. International Society of Offshore and Polar Engineers, 2019.
[16] J. Herp, N.L. Pedersen, E.S. Nadimi, Assessment of early stopping through statistical health prognostic models for empirical RUL estimation in wind turbine main bearing failure monitoring, Energies 13 (1) (2020) 83.
[17] Liu, H., Song, W., Niu, Y., & Zio, E.A generalized cauchy method for remaining useful life prediction of wind turbine gearboxes. Mechanical Systems and Signal Processing, 153, 107471.
[18] J.M.E. Marques, D. Benasciutti, R. Tovo, Variability of the fatigue damage due to the randomness of a stationary vibration load, Int. J. Fatigue 141 (2020), 105891.
[19] M. Rezamand, M. Kordestani, R. Carriveau, D.S.K. Ting, M. Saif, An integrated feature-based failure prognosis method for wind turbine bearings, IEEE/ASME Trans. Mechatron. (2020).
[20] F. Cheng, L. Qu, W. Qiao, L. Hao, Enhanced particle filtering for bearing remaining useful life prediction of wind turbine drivetrain gearboxes, IEEE Trans. Industr. Electron. 66 (6) (2018) 4738–4748.
[21] M. Kordestani, M.F. Samadi, M. Saif, A new hybrid fault prognosis method for MFS systems based on distributed neural networks and recursive Bayesian algorithm, IEEE Syst. J. (2020).
[22] F.K. Moghadam, A.R. Nejad, Natural frequency estimation by using torsional response, and applications for wind turbine drivetrain fault diagnosis, in Journal of Physics: Conference Series (Vol. 1618, No. 2, pp. 022019), 2020 IOP Publishing.
[23] F.K. Moghadam, R.A. Nejad, Experimental Validation of Angular Velocity Measurements for Wind Turbines Drivetrain Condition Monitoring, in: International Conference on Offshore Mechanics and Arctic Engineering, vol. 59353, American Society of Mechanical Engineers, 2019. V001T01A035.
[24] F.K. Moghadam, G.F.D.S. Rebouças, A.R. Nejad, Digital twin modeling for predictive maintenance of gearboxes in floating offshore wind turbine drivetrains, Forsch. Ingenieurwes. (2021) 1–14.
[25] F.K. Moghadam, A.R. Nejad, Theoretical and experimental study of wind turbine drivetrain fault diagnosis by using torsional vibrations and modal estimation, J. Sound Vib. 509 (2021) 116223.
[26] S. Sheng, H. Link, W. LaCava, J. van Dam, B. McNiff, P. Veers, F. Oyague, Wind turbine drivetrain condition monitoring during GRC phase 1 and phase 2 testing (No. NREL/TP-5000-52748). National Renewable Energy Lab. (NREL), Golden, CO (United States), 2011.
[27] J.A. Andrawus, Maintenance optimisation for wind turbines (Doctoral dissertation) (2008).
[28] G. Vizentin, G. Vukelić, M. Srok, Common failures of ship propulsion shafts, Pomorstvo 31 (2) (2017) 85–90.
[29] W.T. Becker, S. Lampman, Fracture appearance and mechanisms of deformation and fracture, ASM International, Materials Park, OH, 2002, pp. 559–586.
[30] T. Rauert, J. Herrmann, P. Dalhoff, M. Sander, Fretting fatigue induced surface cracks under shrink fitted main bearings in wind turbine rotor shafts, Proc. Struct. Integr. 2 (2016) 3601–3609.
[31] Z. Zhang, Z. Yin, T. Han, A.C. Tan, Fracture analysis of wind turbine main shaft, Eng. Fail. Anal. 34 (2013) 129–139.
[32] P. Bortolotti, H. Canet Tarrés, K. Dykes, K. Merz, L. Sethuraman, D. Verelst, F. Zahle, Systems engineering in wind energy-WP2. 1 reference wind turbines. Technical Report No. NREL/TP-5000-73492, National Renewable Energy Laboratory (NREL), 2019.
[33] J.H. Kang, H. Lee, Structural safety analysis of main shaft for wind power generators considering mass effect, Int. J. Appl. Eng. Res. 12 (17) (2017) 6862–6869.

[34] DIN EN 10083-3: 2007-01, Steels for Quenching and Tempering–Part 3: Technical Delivery Conditions for Alloy Steels, German Version EN 10083-3: 2006, 2007.
[35] C. Bak, F. Zahle, R. Bitsche, T. Kim, A. Yde, L.C. Henriksen, A. Natarajan, M.H. Hansen, Description of the DTU 10 MW reference wind turbine. DTU Wind Energy Report-I-0092 2013;5.
[36] J.M. Hegseth, E.E. Bachynski, A semi-analytical frequency domain model for efficient design evaluation of spar floating wind turbines, Mar. Struct. 64 (2019) 186–210.
[37] F.K. Moghadam, A.R. Nejad, Evaluation of PMSG-based drivetrain technologies for 10-MW floating offshore wind turbines: Pros and cons in a life cycle perspective, Wind Energy (2020).
[38] M. Pilanci, O. Arikan, M.C. Pinar, Structured least squares problems and robust estimators, IEEE Trans. Signal Process. 58 (5) (2010) 2453–2465.
[39] G.H. Golub, C. Van Loan, Unsymmetric positive definite linear systems, Linear Algebra Appl. 28 (1979) 85–97.
[40] C. Zhang, F. Plestan, Adaptive sliding mode control of floating offshore wind turbine equipped by permanent magnet synchronous generator, Wind Energy.
[41] A.B. Owen, Monte Carlo theory, methods and examples, Art Owen, 2013.
[42] A.R. Nejad, Z. Gao, T. Moan, On long-term fatigue damage and reliability analysis of gears under wind loads in offshore wind turbine drivetrains, Int. J. Fatigue 61 (2014) 116–128.
[43] B. Sudret, Z. Guédé, P. Hornet, J.M. Stéphan, M. Lemaire, Probabilistic assessment of fatigue life including statistical uncertainties in the SN curve, in: In Proceedings of the 17<Superscript>th</Superscript> International Conference on Structural Mechanics in Reactor Technology, 2003.
[44] B. Engel, S.S.H. Al-Maeeni, Failure analysis and fatigue life estimation of a shaft of a rotary draw bending machine, Constraints 3 (2017) 1785–1790.
[45] R.U. Maheswari, R. Umamaheswari, Trends in non-stationary signal processing techniques applied to vibration analysis of wind turbine drive train-A contemporary survey, Mech. Syst. Signal Process. 85 (2017) 296–311.
[46] W. Bartelmus, R. Zimroz, A new feature for monitoring the condition of gearboxes in non-stationary operating conditions, Mech. Syst. Signal Process. 23 (5) (2009) 1528–1534.
[47] M. Ghane, A. Rasekhi Nejad, M. Blanke, Z. Gao, T. Moan, Condition monitoring of spar-type floating wind turbine drivetrain using statistical fault diagnosis, Wind Energy 21 (7) (2018) 575–589.
[48] H.S. Toft, J.D. Sørensen, D. Veldkamp, Assessment of load extrapolation methods for wind turbines, J. Solar Energy Eng. 133 (2) (2011).
[49] D.C. Montgomery, G.C. Runger, Applied statistics and probability for engineers, sixth ed., John Wiley & Sons, Hoboken, NJ, USA, 2014.
[50] S.D. Downing, D.F. Socie, Simple rainflow counting algorithms, Int. J. Fatigue 4 (1) (1982) 31–40.
[51] J.F. Manwell, J.G. McGowan, A.L. Rogers, Wind energy explained: theory, design and application, John Wiley & Sons, 2010.
[52] D.E.N.I.S. Benasciutti, R.O.B.E.R.T.O. Tovo, Rainflow cycle distribution and fatigue damage in Gaussian random loadings. Internal Report No. 129, Department of Engineering, University of Ferrara, Italy, 2004.
[53] S. Wang, A.R. Nejad, T. Moan, On design, modelling, and analysis of a 10-MW medium-speed drivetrain for offshore wind turbines, Wind Energy 23 (4) (2020) 1099–1117.