Digital twin modeling for predictive maintenance of gearboxes in floating offshore wind turbine drivetrains

Farid K. Moghadam[1] · **Geraldo F. de S. Rebouças**[1] · **Amir R. Nejad**[1]

Abstract

This paper presents a multi-degree of freedom torsional model of drivetrain system as the digital twin model for monitoring the remaining useful lifetime of the drivetrain components. An algorithm is proposed for the model identification, which receives the torsional response and estimated values of rotor and generator torques, and calculates the drivetrain dynamic properties, e.g. eigenvalues, and torsional model parameters. The applications of this model in prediction of gearbox remaining useful lifetime is discussed. The proposed method is computationally fast, and can be implemented by integrating with the current turbine control and monitoring system without a need for a new system and sensors installation. A test case, using 5 MW reference drivetrain, has been demonstrated.

Modellierung eines digitalen Zwillings zur vorausschauenden Wartung von Getrieben in Antriebssträngen schwimmender Offshore-Windkraftanlagen

Zusammenfassung

In diesem Artikel wird ein multidimensionales Torsionsmodell eines Antriebsstrangs als digitaler Zwilling vorgestellt, der zum Monitoring der verbleibenden nutzbaren Lebensdauer von Antriebsstrangkomponenten verwendet wird. Für die Modelidentifizierung wird ein Algorithmus vorgeschlagen, der ausgehend von der Torsionsantwort und den Schätzwerten des Rotor- und Generatordrehmoments die dynamischen Eigenschaften des Antriebsstrangs, insbesondere Eigenwerte, und Torsionsmodelparameter berechnet. Die Anwendung dieses Models zur Vorhersage der verbleibenden nutzbaren Lebensdauer wird diskutiert. Die vorgeschlagene Methode ist berechnungseffizient und kann in das aktuelle Turbinenkontroll- und Überwachungssystem integriert werden, ohne dass die Installation von neuen Systemen und Sensoren notwendig wird. Schließlich wird ein Testfall am Beispiel eines 5 MW-Referenzantriebsstrangs demonstriert.

1 Introduction

Power train system including rotor, main bearings, gearbox, generator and power converter accounts for *57%* of turbine total failures and *65%* of turbine total downtime [1]. The consequent costs are expected to be higher in floating wind turbines (FWT) due to costly and slow-going marine operations. In FWTs, the power train system is exposed to a wider range of excitations than bottom-fixed turbines due to synergistic impacts of wind, wake, currents and structural motions-induced vibrations [2]. Predictive maintenance is proposed in the literature as a support tool for the time base maintenance to reduce the unexpected shutdowns and the consequent downtimes while it can also help to optimize the inspection intervals. The motivation of this research is increasing the wind turbine (WT) availability by performing predictive maintenance of the drivetrain system through the digital twin models.

Digital twins are highly accurate but computationally fast models of the system which can be run in-line with the real machine. The model should be able to update itself by the online operational measurements, to capture the physical variations in the system over the time [3]. In this paper, digital twin is proposed as a tool for lifetime monitoring of the drivetrain system, and more specifically the gearbox, by using the torsional vibration measurements, which

✉ Farid K. Moghadam
farid.k.moghadam@ntnu.no

1 Norwegian University of Science and Technology, Trondheim, Norway

can be included in the farm Supervisory Control and Data Acquisition (SCADA) existing system by using higher resolution encoders. The input data may also be provided by additional installations *e.g.* proximity sensors and angular accelerometers. Correa *et al.* [4] performs a review on the drivetrain components condition monitoring techniques and faults prediction in general and also by using the SCADA data. Digital twin is a proven technology which is used by Siemens Gamesa for prediction of drivetrain loads and subsequently improving the drivetrain design [5]. The application of digital twin for monitoring of the drivetrain lifetime may be looked as an expensive solution. However, in this work, it is shown that by using innovative approaches based on analysis of torsional measurements it is possible to offer computationally-fast solutions for monitoring the lifetime of the drivetrain critical components with a potential to be integrated with the currently available control and monitoring system with minimum additional cost. Digital twin in this context means the combination of model, online measurements and remaining useful lifetime (RUL) model as defined and suggested by [6].

Rebouças *et al*. [7] suggests the 14-DOF torsional model as a fair compromise between complexity and accuracy when the structural integrity of gearbox is taken into consideration. The structural integrity is mainly measured in terms of contact fatigue stresses of the gear pairs, so that this model can be used as the indicator for the gears contact loads. For this purpose, the applicability of a 14-DOF lumped-parameter torsional model of the drivetrain system as the digital twin model for monitoring the degradation in a high-speed drivetrain technology with a three-stages gearbox which consists of two planetary and one parallel stages is investigated. The algorithm for estimation of digital twin model parameters is based on a data-driven approach which operates based on the real-time drivetrain torsional measurements, and the estimated dynamic properties from the torsional response. The algorithm employed for estimation of dynamic properties, namely natural frequencies and damping from the torsional response is according to the modal estimation approach proposed by [8]. In order to estimate the moment of inertia of the components in the equivalent model, an optimization problem is defined based on the least square error between the digital twin model and the online measurements consisting of torsional response and estimated input torque. After all, the stiffness parameters of the model are estimated by using the equations of natural frequencies as a function of equivalent model parameters (stiffness and inertia). It is worth noting that the equivalent lumped parameter model of drivetrain in the proposed digital twin approach is linear. However, the eigenfrequencies of the drivetrain are nonlinear functions of equivalent model parameters. Since the assumption of accessibility of all the system dynamic states is opti-

mistic, the problem is extend to a more general class of problem with a partial knowledge about the system modes. The parameters of the digital twin model are updated by using the real-time operational data. The most prevalent failures of the large gears in WT drivetrain systems are due to gear tooth root bending and pitting fatigue damage [9]. For monitoring the RUL of the gears in this work, the degradation of the gears due to pitting fatigue damage and the gear pairs contact loads are taken into consideration. The contact loads on the gear pairs are estimated by using the load observers designed for the gear pairs based on the real-time estimated digital twin model and torsional measurements. The employed model-based degradation model is based on stress-life method. The online estimated contact stress feeds the time-domain cycle counting approach based on rainflow method for online estimation of fatigue cycles, and then for estimation of fatigue damage for each gear pair by using the Miner's rule. For all the simulation cases, the floating 5 MW NREL WT with a spar support substructure as a common way of realization of FWT is used. The reason for the latter is the higher motivation in FWT for predictive maintenance.

This paper is the proof of concept for the idea of near real-time estimation of the load and subsequently the residual life in the drivetrain components by using computationally efficient equivalent models estimated from torsional measurements supported by real-time values of torsional response. The proposed algorithm is able to estimate the equivalent model parameters by using only few measurement samples which feed a robust linear regression estimator designed to estimate the equivalent model parameters. On this basis, the main contributions of this work are:

- Proposing the 14-DOF equivalent lumped-parameter model as the digital twin of the drivetrain system for remaining useful lifetime monitoring of the gearbox,
- Proposing an algorithm for the near real-time estimation of the equivalent model parameters by using the torsional measurements,
- Designing computationally inexpensive load observers for the near real-time estimation of the contact load and stress on the gears in planetary and parallel stages of the drivetrain gearbox, by using equivalent model parameters and real-time torsional responses,
- Proposing a physics-based degradation model for near real-time estimation of the residual life of the gears in the drivetrain gearbox by using the online estimated equivalent model, real-time measurements, designed load observers and contact stress estimation approach,
- Simulation studies to evaluate the proposed digital twin approach for estimation of fatigue damage of the gears, and validation of the results by high-fidelity simulation models.

Fig. 1 Representation of the torsional model [7]
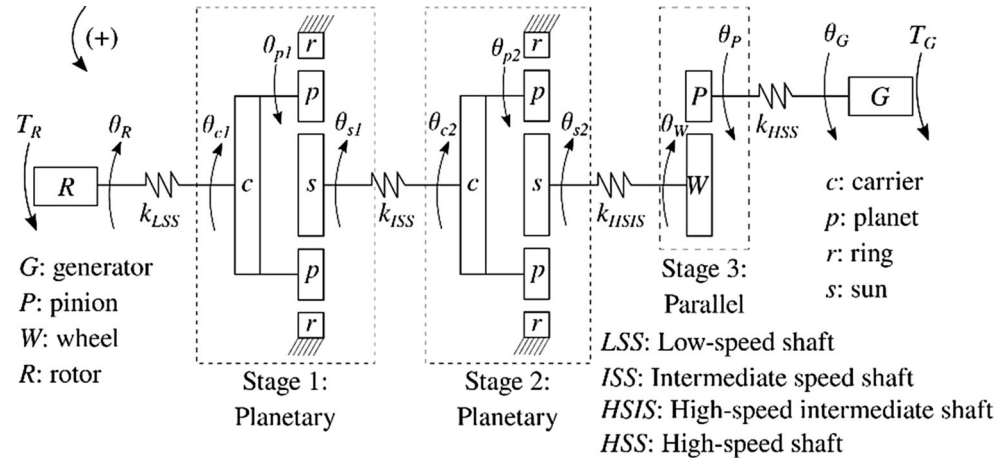
2 Methodology

2.1 Test case

NREL 5 MW reference wind turbine [10] 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 [11]. This model is able to capture the global dynamics of spar floating wind turbine from the interaction with the environmental loads. The 5 MW reference drivetrain [12] is employed in this study.

Fig. 2 Assembling the global matrices from their local equivalents [13]

2.2 Global simulation and estimation of drivetrain loads

The decoupled simulation approach is used for the drivetrain studies in this work. The latter means that the turbine global simulation results, namely the rotor aerodynamic load and the bedplate motions are applied to the detailed local drivetrain model in a secondary software. Then the drivetrain local load effects are obtained for further postprocessing aimed at monitoring the drivetrain remaining lifetime. More details on the utilized decoupled approach and the global simulation model is available in [11]. The wind is turbulent based on Kaimal distribution. The turbulence intensity at hub height I (–) is assumed to be 0.14, for all the wind speeds, according to IEC 61400-1 class B turbines. The wave is modelled stochastic by two parameters, namely significant wave height H_s (m) and peak period T_p (s) in global analysis.
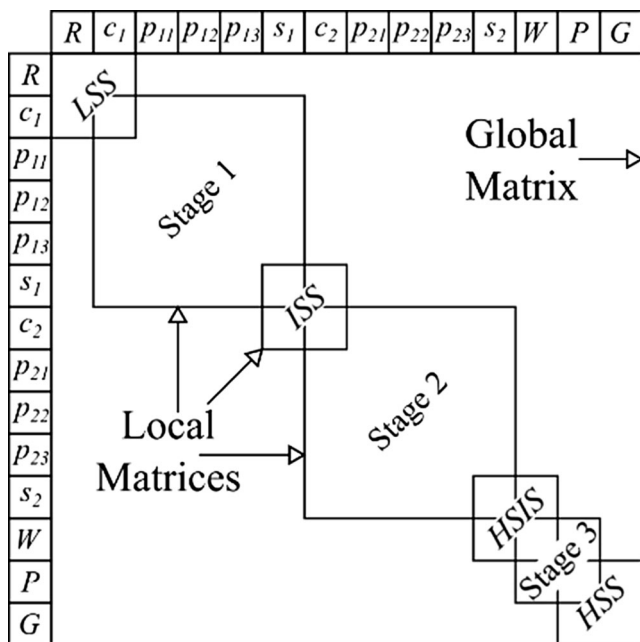
2.3 Drivetrain torsional model as the digital twin model for predictive maintenance

The NREL 5 MW reference drivetrain proposed in [12] is considered in this study, considering only its torsional behavior. Fig. 1 contains a schematic representation of the model used in this work. The gearbox is assumed to be clamped to the nacelle's bedplate by rigid torque arm supports, leading to no rotation of the ring gears. This model is based on [13], where the resonant behavior of a single-stage planetary gearbox is discussed. The equations of motion are shown in Eq. (1).

$$\mathbf{J}\ddot{\theta} + \mathbf{D}\dot{\theta} + \mathbf{K}\theta = \mathbf{T}, \tag{1}$$

$$\theta = \begin{bmatrix} \theta_R & \boldsymbol{\theta}_1^T & \boldsymbol{\theta}_2^T & \boldsymbol{\theta}_3^T & \theta_G \end{bmatrix}^T,$$

$$\mathbf{T} = \begin{bmatrix} T_R & 0 & \cdots & 0 & T_G \end{bmatrix}^T, \tag{2}$$

In these equations, θ_k and T_k ($k \in \{R, 1, 2, 3, G\}$) represent torsional coordinate and external torques, respectively. The indexes R and G are associated with rotor and generator variables, respectively. The sub-vectors $\boldsymbol{\theta}_i (i \in \{1, 2, 3\})$ refer to the gear stages, see Eqs. (3) and (6). The global inertia and stiffness matrices $\mathbf{J}$ and $\mathbf{K}$ are obtained from their local counterparts. The damping is defined using Rayleigh parameters, being proportional to the inertia and stiffness matrices $\mathbf{J}$ and $\mathbf{K}$, i.e $\mathbf{D} = \alpha\mathbf{J} + \beta\mathbf{K}$.

The local matrices for the planetary stages can be seen in Eqs. (3) to (5) and are based on the matrices used in [13], while the local matrices for the parallel stage, shown in Eq. (6), were obtained by isolating the sun-planet gear pair of the planetary stage. In these equations, θ, m, k, and r, represent the angular displacement, mass, stiffness and base radius respectively. The indices c, p, s, P, and W represent the carrier, planet and sun elements of the ith stage, respectively. The mesh stiffness between sun-planet, ring-planet and pinion-wheel gears is represented by k_{sp}, k_{rp}, and k_{PW}, respectively. In this work, the mesh stiffness between different gear pairs is calculated according to ISO 6336-2 [14], which takes a constant value. The center distance for planetary stages is represented by a_w [7]. Shafts are used to connect the adjacent inertia elements, such as the rotor, generator, and gear stages. The shafts were modelled using finite element theory leading to the shaft's inertia and stiffness matrices J_S and K_S in Eq. (7). Assembling the global matrices from the local matrices is illustrated in Fig. 2. In Eq. (7), D, L, and G represent the shaft's diameter, length, and the shear elasticity modulus of its material. The indices a and b on the shaft's displacement vector $\boldsymbol{\theta}_S$ represent DOFs adjacent to the shaft. For the high-speed intermediate shaft (HSIS), shown in Fig. 1, $a = s_2$ and $b = W$.

$$\boldsymbol{\theta}_i = \begin{bmatrix} \theta_{c_i} & \theta_{p_{1i}} & \theta_{p_{2i}} & \theta_{p_{3i}} & \theta_{s_i} \end{bmatrix}^T, i = 1,2 \tag{3}$$

$$\boldsymbol{J}_i = \text{diag}\left(\begin{bmatrix} m_c a_w^2 & m_p r_p^2 & m_p r_p^2 & m_p r_p^2 & m_s r_s^2 \end{bmatrix}\right), \tag{4}$$

$$\boldsymbol{K}_i =$$

$$\begin{bmatrix} 3a_w(k_{rp}+k_{sp}) & a_w r_p(k_{rp}-k_{sp}) & a_w r_p(k_{rp}-k_{sp}) & a_w r_p(k_{rp}-k_{sp}) & -3a_w r_s k_{sp} \\ a_w r_p(k_{rp}-k_{sp}) & r_p^2(k_{rp}+k_{sp}) & 0 & 0 & r_p r_s k_{sp} \\ a_w r_p(k_{rp}-k_{sp}) & 0 & r_p^2(k_{rp}+k_{sp}) & 0 & r_p r_s k_{sp} \\ a_w r_p(k_{rp}-k_{sp}) & 0 & 0 & r_p^2(k_{rp}+k_{sp}) & r_p r_s k_{sp} \\ -3a_w r_s k_{sp} & r_p r_s k_{sp} & r_p r_s k_{sp} & r_p r_s k_{sp} & 3r_s^2 k_{sp} \end{bmatrix} \tag{5}$$

$$\boldsymbol{\theta}_3 = \begin{bmatrix} \theta_W & \theta_P \end{bmatrix}^T,$$

$$\boldsymbol{J}_3 = \text{diag}\left(\begin{bmatrix} m_w r_W^2 & m_P r_P^2 \end{bmatrix}\right), \tag{6}$$

$$\boldsymbol{K}_3 = k_{PW} \begin{bmatrix} r_W^2 & r_P r_W \\ r_P r_W & r_P^2 \end{bmatrix}$$

$$\boldsymbol{\theta}_S = \begin{bmatrix} \theta_a & \theta_b \end{bmatrix}^T,$$

$$\boldsymbol{J}_S = \frac{mD^2}{48} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \tag{7}$$

$$\boldsymbol{K}_S = \frac{\pi G D^4}{32L} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

The generator torque is realized by a proportional-integral (PI) speed controller, defined as

$$T_G = K_P e + K_I \int_0^t e \, d\tau, e = \omega_G - \dot{\theta}_G, \tag{8}$$

where the error e is given as the difference between reference and measured generator speeds ω_G and $\dot{\theta}_G$, respectively. $K_P = 2200$ is the proportional gain and $K_I = 220$ is the integral gain.

2.4 Estimation of model parameters from torsional response

Drivetrain modal estimation The torsional response residual function between the inertia j_m and j_n from the point m is defined as

$$e_{m,n}^\alpha \triangleq \alpha_m - u_{m,n}\alpha_n, \text{ for } m, n \in \{1, ..., 14\}, \tag{9}$$

with α is the angular acceleration, and $u_{m,n}$ is the relative gear ratio between j_m and j_n to make them in the same coordinate. Gear ratio $u_{m,n}$ as per definition is $\frac{N_m}{N_n}$, where N_m and N_n are the speeds at m^{th} and n^{th} inertias. The theoretical explanation of estimation of natural frequencies from the frequency spectrum of $e_{m,n}^\alpha$ is provided by [8]. The different residual functions can be defined by using the response obtained from the different locations of drivetrain, which present different levels of visibility of the drivetrain torsional modes.

Estimation of moment of inertia matrix By summing the moments on each of the inertias of the lumped parameter model yields 14 equations of the form

$$j_i \ddot{\theta}_i + D_i \left(\dot{\theta}_i - \dot{\theta}_{i-1} \right) - D_{i+1} \left(\dot{\theta}_{i+1} - \dot{\theta}_i \right)$$
$$+ k_i \left(\theta_i - \theta_{i-1} \right) - k_{i+1} \left(\theta_{i+1} - \theta_i \right) = 0, \tag{10}$$
$$\text{for } i = (1, ..., 14).$$

In the matrix form, these set of equations can be written as

$$\mathbf{J}\ddot{\Theta}(t) + \mathbf{D}\dot{\Theta}(t) + \mathbf{K}\Theta(t) = \mathbf{T}. \tag{11}$$

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

$$E(t) \triangleq \widehat{\mathbf{J}}\ddot{\Theta}(t) + \widehat{\mathbf{D}}\dot{\Theta}(t) + \widehat{\mathbf{K}}\Theta(t) - \mathbf{T}(t). \tag{12}$$

The least square estimator is defined by

$$\widehat{\mathbf{J}}^{\mathbf{LS}}, \widehat{\mathbf{K}}^{\mathbf{LS}}, \widehat{\mathbf{D}}^{\mathbf{LS}} \in \operatorname{argmin} \left\{ \|E\|^2 \right\}, \\ \mathbf{J}, \mathbf{K}, \mathbf{D} \geq \mathbf{0} \tag{13}$$

where $\mathbf{J}$ is a diagonal matrix. $\mathbf{K}$ and $\mathbf{D}$ are nondiagonal but symmetric matrices. The matrices D and K are not full rank. The latter causes some computational difficulty for the above quadratic matrix optimization problem, so that there is no guarantee for the convergence. In order to remove the coupling between these equations due to the $\mathbf{K}$ and $\mathbf{D}$ terms, 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 individual equations of each inertia. The latter leads to the following error function in terms of the variables j_i with rotor as the reference of the rotary coordinate.

$$e(t) = j_1\alpha_1(t) + ... + a_{1,i}j_i\alpha_i(t) + ...+ \\ a_{1,14}j_{14}\alpha_{14}(t) - \tau^{\mathbf{Rot}}(t) - a_{1,14}\tau^{\mathbf{Gen}}(t), \tag{14}$$

where α_i is the time series of angular acceleration at the ith body. $\tau^{\mathbf{Rot}}$ and $\tau^{\mathbf{Gen}}$ are the time series of the rotor and generator torques respectively. The sign of $a_{1,i}$ is determined based on the direction of rotation of j_i. The latter leads to the following quadratic scalar optimization problem as

$$\widehat{j}^{LS} = \operatorname{argmin} \left\{ \|e\|^2 \right\}. \\ j \geq 0 \tag{15}$$

This estimator is robust to the measurement noises, and can provide a good approximation even with less than 14 input data samples (underdetermined case). For the case of more than 14 samples (overdetermined case), this estimator helps to obtain more accurate estimation than solving the linear equations, when the input measurements are subject to independent identically distributed (i.i.d.) Gaussian noise. In other words, the total LS technique is able to correct the system with minimum perturbation [15]. The above convex optimization problem is numerically solved by Matlab CVX and the global optimizer $j^{LS} = \{j_1, ..., j_{14}\}$ is estimated.

Estimation of stiffness matrix The undamped torsional frequencies of the system are the nonlinear function of inertia and stiffness as

$$\omega_i \text{ (for } i = 1, ..., 14) = \sqrt{\operatorname{eig}\left(-J^{-1}\mathbf{K}\right)}. \tag{16}$$

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

$$g_i = \omega_i{}^2 - \operatorname{eig}\left(-J^{-1}\mathbf{K}\right), \text{ (for } i = 1, ..., 14). \tag{17}$$

In general, there is not a unique matrix $\mathbf{K}$ from the above equation for the known set of eigenvalues $\lambda_i = \{\omega_1{}^2, ..., \omega_{14}{}^2\}$ of the matrix $-J^{-1}\mathbf{K}$. However, by imposing the sparsity and symmetricity to matrix 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 $J^{-1}\mathbf{K}$ 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 $-J^{-1}\mathbf{K}$ 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 [16].

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 n eigenfrequencies as defined by the following least square error estimator:

$$\widehat{k}^{LS} = \operatorname{argmin} \left\{ \|\lambda_n - eig(\mathbf{\Lambda}, n)\|^2 \right\}, \\ \mathbf{\Lambda} \in \Lambda \tag{18}$$

with $\mathbf{\Lambda}$ is the variable of this problem which is a function of the unknown variable $\mathbf{K}$ as $\mathbf{\Lambda} = -J^{-1}\mathbf{K}$. 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_n is the set of n ($n \in \{1, ..., 14\}$) smallest magnitude eigenvalues which are known from the modal estimation, $\lambda_n = \{\omega_1{}^2, ..., \omega_n{}^2\}$. $eig(\mathbf{\Lambda}, n)$ is the set of n (($n \in \{1, ..., 14\}$)) smallest magnitude eigenvalues de-

Fig. 3 Free-body diagrams for **a**–**c** planetary and **d**–**e** parallel gear stages
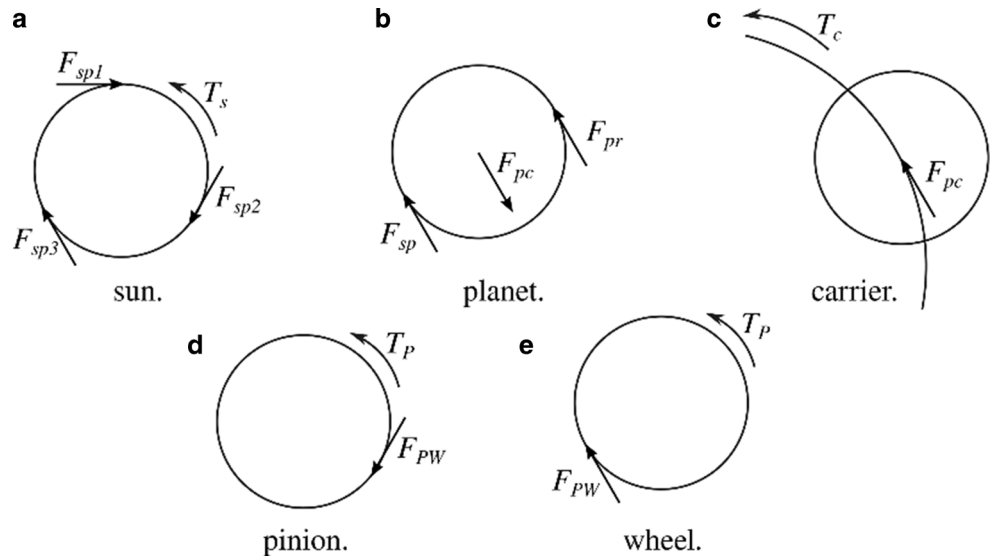
fined in terms of matrix $\mathbf{J}$ and the unknown matrix $\mathbf{K}$. The feasible set Λ is also defined by

$$\Lambda = \left\{ \mathbf{\Lambda} : \mathbf{\Lambda} \in R^{14 \times 14}, \mathbf{K} \geq \mathbf{0}, \Lambda_{l,m} = 0, \forall \Lambda_{l,m} \in S^{\Lambda} \right\}, \quad (19)$$

where S^{Λ} is the sparsity of matrix $\mathbf{\Lambda}$. The positive definiteness and sparsity of $\mathbf{\Lambda}$ 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, $\mathbf{\Lambda}$ 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.

2.5 Estimation of gearbox loads, and defining degradation model

Estimating contact stresses and loads at each stage The free-body diagram for the gears in planetary and parallel stages can be seen in Fig. 3. For wind turbine gearboxes, the input and output torques for planetary (parallel) stages are the carrier (pinion) torque T_C (T_P) and the sun (wheel) torque T_s (T_W), respectively. From Fig. 3, one obtains Eqs. (20) and (21) as

Planetary stage:

$$\begin{aligned} J_s \ddot{\theta}_s &= T_S - N_p r_s F_{sp}, \\ J_p \ddot{\theta}_p &= r_p \left(F_{\text{pr}} - F_{\text{sp}} \right) = -J_p \frac{r_s}{r_p} \ddot{\theta}_s, \\ m_p a_w \ddot{\theta}_c &= F_{\text{pr}} + F_{\text{sp}} - F_{\text{pc}} = m_p a_w \frac{r_s}{r_s + r_r} \ddot{\theta}_s, \\ J_c \ddot{\theta}_c &= T_c - N_p r_s F_{\text{sp}} = J_c \frac{r_s}{r_s + r_r} \ddot{\theta}_s, \end{aligned} \quad (20)$$

Parallel stage:

$$\begin{aligned} J_P \ddot{\theta}_P &= T_P - r_P F_{PW} = -J_P \frac{r_P}{r_W} \ddot{\theta}_W, \\ J_W \ddot{\theta}_W &= T_W - r_W F_{PW}, \end{aligned} \quad (21)$$

One can obtain the relationships between the input and output torques by eliminating the internal forces between elements. Speed relations are time differentiated which enables writing the contribution of inertial torques by using a single torsional acceleration and equivalent mass moment of inertia. For the planetary stage, one obtains

$$\begin{aligned} T_S &= J_{EQ} \ddot{\theta}_s - T_c \frac{r_s}{2a_w}, \quad \text{where} \\ J_{EQ} &= J_s + \frac{N_p r_s^2}{2} \left(\frac{J_p}{r_p^2} + m_p \frac{a_w}{r_s + r_r} \right) + J_c \frac{r_s^2}{2a_w(r_s + r_r)}, \end{aligned} \quad (22)$$

with J_{EQ} as the equivalent mass moment of inertia, having contributions from the sun, N_p planet gears and planet carrier. Similarly, for the parallel stage one has

$$\begin{aligned} T_W &= J_{EQ} \ddot{\theta}_W + T_P \frac{r_W}{r_P}, \\ J_{EQ} &= J_W + J_P \left(\frac{r_W}{r_P} \right)^2, \end{aligned} \quad (23)$$

with J_{EQ} having contributions from both pinion and wheel gears. The term multiplying the torque at the pinion is the gear ratio for parallel gears.

Torque transfer between stages is made via shafts, and can be estimated by

$$T_{\text{out}} = T_{in} - J_S \ddot{\theta}_S, \quad (24)$$

where $J_S \ddot{\theta}_S$ is the inertial torque from the shaft, [17]. Therefore, for the NREL 5MW drivetrain, one should have

$$
\begin{aligned}
T_{in1} &= T_{c1} = T_R - J_{\text{LSS}}\ddot{\theta}_R, & T_{out1} &= T_{s1} = J_{EQ1}\ddot{\theta}_{s1} - T_{in1}\tfrac{r_{s1}}{2d_{w1}}, \\
T_{in2} &= T_{c2} = T_{out1} - J_{\text{ISS}}\ddot{\theta}_{s1}, & T_{out2} &= T_{s2} = J_{EQ2}\ddot{\theta}_{s2} - T_{in2}\tfrac{r_{s2}}{2d_{w2}}, \\
T_{in3} &= T_P = T_G - J_{\text{HSS}}\ddot{\theta}_G, & T_{out3} &= T_W = J_{EQ3}\ddot{\theta}_W - T_{in3}\tfrac{r_W}{r_P},
\end{aligned} \tag{25}
$$

The input torques estimated above can be used to estimate the stresses at the different gear stages as described below.

Gear contact stresses are analyzed in this work following ISO 6336-2:2019 [14]. According to this standard, the contact stresses are defined as

$$
\begin{aligned}
\sigma_{Hi} &= Z_{BD}Z_H Z_E Z_\epsilon Z_\beta \\
&\cdot \sqrt{K_A K_{\gamma i} K_{vi} K_{H\beta i} K_{H\alpha i}} \sqrt{\frac{2000 T_i}{b d_1^2} \frac{u+1}{u}},
\end{aligned} \tag{26}
$$

where u is the gear ratio of the pair, d_1 and b are reference diameter and face width of the pinion, respectively, T_i is the input torque. The other parameters account for different aspects of the problem, such as contact relations Z_{BD} and ratios Z_ϵ, material properties Z_E, helix angle Z_β, mesh load K_γ, gear speed K_v, load distributions $K_{H\beta i}$ and $K_{H\alpha i}$. These factors are discussed in [15] and references therein. Eq. (26) can be rewritten in a compact form as

$$
\sigma_{Hi} = C\sqrt{T_i}, \tag{27}
$$

where C represents the K and Z design parameters mentioned above. This parameter, unknown to the maintenance engineer, can be roughly estimated from nominal conditions as $C = \sigma_{HN}/\sqrt{T_{iN}}$. Therefore, Eq. (27) turns to

$$
\sigma_{Hi} = \sigma_{HN}\sqrt{\frac{T_i}{T_{iN}}}, \tag{28}
$$

where σ_{Hi} is defined in terms of nominal contact stress σ_{HN} and torque T_{iN}, which have a clear physical meaning. Additionally, one can also expand Eq. (28) using Taylor series around the nominal torque T_{iN} as

$$
\begin{aligned}
\sigma_{Hi} - \sigma_{HN} &= (T_i - T_{iN})\frac{\sigma_{HN}}{2T_{iN}} - (T_i - T_{iN})^2\frac{\sigma_{HN}}{8T_{iN}^2} \\
&+ O\left((T_i - T_{iN})^3\right).
\end{aligned} \tag{29}
$$

This formula gives a polynomial relationship between stress and load deviations. The accuracy of this expression

for higher torque deviations can be increased by considering other terms of the Taylor series. The nominal stresses are of order 10^3 and the nominal torques vary between 10^4 and 10^6, thus unitary torque deviations lead to 0.05 to 5% stress deviations, if only the first term of the series in Eq. (29) is considered. A limitation of this derivation is that C depends on parameters that may change with T_i, but these variations should be negligible.

Remaining useful lifetime estimation The real-time accumulated damage is estimated as follows. First, the time-varying gear transmitted loads of the gear pairs of three gearbox stages are estimated by using the torque observers designed as explained in the previous part. The latter is based on the real-time aerodynamic and generator torques, and the time-varying digital twin model parameters. Then the gear tooth surface pitting stress is estimated for the different gears as explained on above. The number of gear tooth contact stress cycles at different stress levels is counted by using the time-domain rainflow cycle counting approach [18]. The outputs are the amplitude stress level σ_s, and the number of stress cycles at σ_s for $s = (1, ..., S)$. 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 [19],

$$
\sigma_s^e = \frac{\sigma_s}{1 - \frac{\sigma_m}{\sigma_u}}, \forall\, (s \in \{1, ..., S\}), \tag{30}
$$

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

$$
d^t = \sum_{s=1}^{S} \frac{n_s}{N_s}, \tag{31}
$$

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 $N_s = k(\sigma_s^e)^{-m}$. The total absolute online accumulated damage will then be calculated by

$$
D = \sum_{t=1}^{T} d^t, \tag{32}
$$

where T stands for the last data block which represents the current time. This method can also be used for estimation of relative damage between different operational periods over

Table 1 14-DOF drivetrain model natural frequencies (in Hz)

Mode	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Healthy^Actual	0	2	120	153	208	567	716	716	971	1224	1232	1232	1384	1943
Healthy^Estimated	0	2	118	153	208	567	706	706	970	1220	1220	1220	1381	1933

the time, to give an insight on variations in degradation between different operational periods.

3 Simulation results

3.1 Aerodynamic and generator torques, and drivetrain torsional response

The input aerodynamic torque obtained from SIMO-RIFLEX-AeroDyn and the performance of the designed controller in controlling the generator torque under variable input torque to set the speed on the shaft is shown in Fig. 4. The drivetrain model responses are the angular displacement, velocity and acceleration in the different bodies in the described 14-DOF model. As an example, the angular velocity responses of the rotor and generator are shown in Fig. 4. For demonstration purposes the torques

and response are scaled with the gear ratio. The turbine operation is assumed to be near the rated operation.

3.2 Estimation of dynamic properties and digital twin model parameters

The undamped frequency modes of the 14-DOF model for the healthy system are listed in the Table 1. The torsional response error function can be defined between different bodies in the drivetrain model. For example, the undamped natural frequencies estimated from the angular acceleration error function for the 2nd gearbox stage is shown in Fig. 5. The main feature of angular acceleration compared to angular velocity and displacement is the amplification of the higher natural frequencies in the response. By defining the angular acceleration error functions in terms of different pairs of bodies in the model, it is possible to estimate the different drivetrain modes. However, the significance
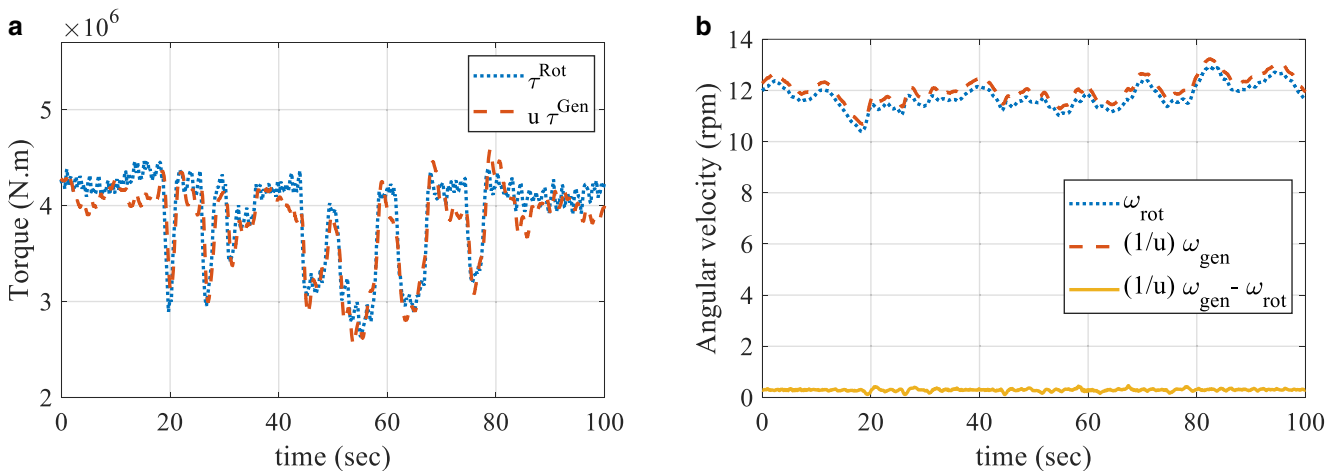
Fig. 4 Drivetrain model loads and responses **a** Rotor and generator torques, **b** Torsional response
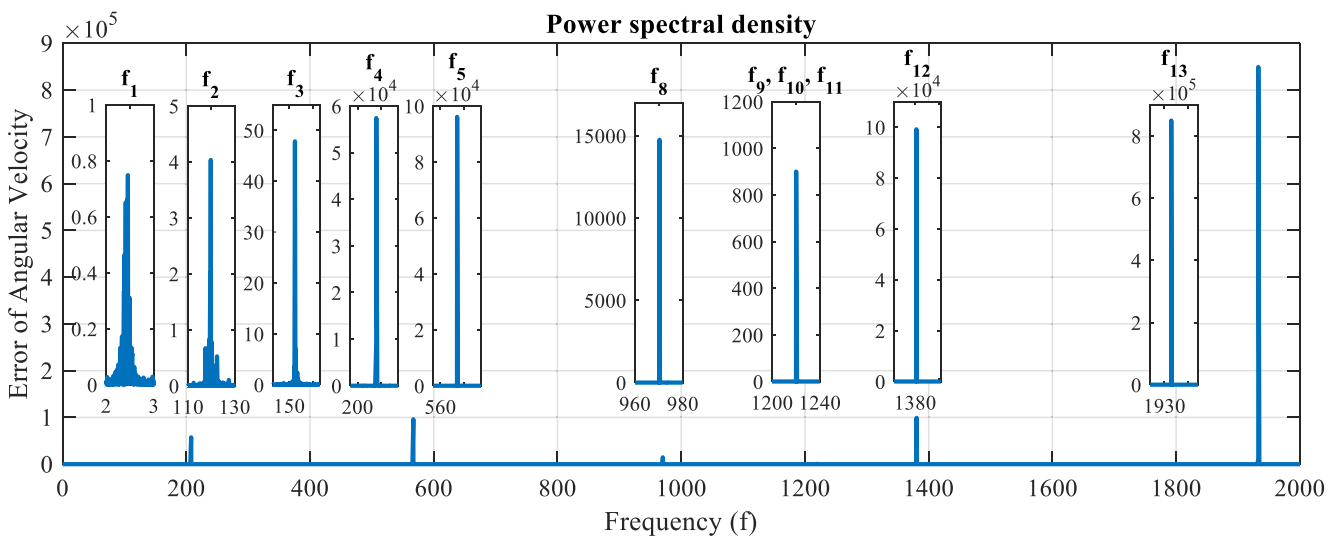
Fig. 5 Estimation of torsional modes from gearbox 2nd stage error function

Table 2 14-DOF drivetrain model natural frequencies in the different fault cases (in Hz)

Mode	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Healthy$^{\text{Actual}}$	0	2	120	153	208	567	716	716	971	1224	1232	1232	1384	1943
Fault$_{\text{stage1}}^{\text{sun}}$	0	2	120	153	208	565	700	700	929	1224	1232	1232	1384	1943
Fault$_{\text{stage2}}^{\text{sun}}$	0	2	119	153	208	567	716	716	971	1220	1220	1224	1383	1850
Fault$_{\text{stage1}}^{\text{ring}}$	0	2	120	153	208	541	696	696	968	1224	1232	1232	1384	1943
Fault$_{\text{stage2}}^{\text{ring}}$	0	2	120	153	208	567	716	716	971	1181	1181	1224	1314	1941

of modes is different in the error functions defined between different bodies. As it can be seen in Fig. 5, the 6th and 7th modes are not visible in the error function of the 2nd stage, but they can be observed in the error function defined between rotor and generator bodies. The estimated natural frequencies by using the employed modal estimation approach based on angular acceleration error function are listed in the Table 1. The maximum relative error of estimation is less than *1.5%*. In practice, the torsional response in commercially available turbines is only accessible by the encoders placed on rotor and possibly generator shafts, so that some of the modes may not be observable by only using the error function between rotor and generator.

The influence of pitting on parameters of drivetrain equivalent model is usually represented by a decrease on the mesh stiffness of the faulty gear pair [20, 21]. The NREL 5 MW drivetrain presented in Sect. 2.3 is used to illustrate this behavior, assuming that the early occurrence of pitting at the sun gear can lead to a 10% decrease in the sun-planet mesh stiffness. The 10% reduction of the faulty gear mesh stiffness assumed for modelling a very early stage fault might be high, but this assumption is made to demonstrate more clearly the influence of the fault on the variations of drivetrain dynamic properties. The results of the mentioned fault in the first and second gear stages can be seen in the second and third rows of Table 2, where one can see that pitting in the first stage sun gear affected mainly the 7th, 8th and 9th resonances, whereas pitting in the second stage sun gear affected mainly the 14th resonances. The results of fault simulation reported in Table 2 show that a 10% reduction in the sun-planet mesh stiffness of the first gear stage, results in around 2.5%, 2.5% and 4.5% reduction of 7th, 8th and 9th drivetrain natural frequencies, respectively, and a 10% reduction in the sun-planet mesh stiffness of the second gear stage, results in about 5% reduction of the highest natural frequency of drivetrain. The influence of pitting fault in the ring of first and second gear stages on the drivetrain natural frequencies are listed respectively in the fourth and fifth rows of Table 2. A 10% reduction in the ring-planet mesh stiffness of the first gear stage, results in around 5%, 3% and 3% reduction of 6th, 7th and 8th natural frequencies, and a 10% reduction in the ring-planet mesh stiffness of the second gear stage, results in the 3.5%, 4% and 5% reduction of 10th, 11th and

13th natural frequencies, respectively. Since the relative error of estimating the natural frequencies can be as high as 1.5%, variations of the natural frequencies which are less than 1.5% cannot be used as the indicator of fault. More detailed models of pitting fault can be engaged to capture more precisely the dynamics of fault [22] and consequent influence on the drivetrain dynamic properties, which is not the scope of this work. The influence of drivetrain faults at system-level on dynamic properties is analytically studied by [23]. Monitoring the variations in the drivetrain system natural frequencies can support the fault detection of the drivetrain at component-level, *e.g.* faults in the gears of the gearbox gear stages, which is also not the main scope of this work.

The actual values of moment of inertia of the different bodies in the 14-DOF model and the performance of the designed LS estimator in estimation of those parameters from the torsional measurements based on the theory elaborated in Sect. 2 can be seen in the Table 3. The accuracy increases as the number of input samples increases while the data outliers are filtered to improve the estimation. The number of input samples can be selected to reach a good tradeoff between the accuracy and computational speed. The actual values of the diagonal elements of stiffness matrix as the main stiffness parameters are listed in the Table 3. In the case that the exact values of eigenfrequencies and inertia parameters are accessed, the solution of Eq. (16) gives the exact values of stiffness parameters. The estimated values of stiffness parameters of the main diagonal of the matrix, by considering both the natural frequencies and inertia parameters estimation errors, are listed in the Table 3. *Estimated^{M1}* is the designation used to indicate that the first proposed stiffness estimation method which denotes the case that all the eigenfrequencies are accessed. *Estimated^{M2}* indicates that the second method of stiffness estimation is used, which is associated to the case that some of the frequency modes are not observable. In order to calculate *Estimated^{M2}*, two different cases in which only the first ten and eleven estimated modes are available are simulated, and the optimization problem defined by the LS estimator in Eq. (18) is solved. The results are listed in the Table 4. The estimation error of the associated underdetermined LS estimator reduces as more frequency modes are known. It is worth noting that our extensive simula-

Table 3 Estimation of the digital twin model parameters by using the proposed approach

Digital twin model parameters							
Inertia				Stiffness			
	Actual	Estimated	Error (%)		Actual	Estimated[M1]	Error (%)
j_1	57,231,658	57,233,035	0.0	$k_{R–R}$	933,804,433	930,722,878	0.3
j_2	11,378	11,406	0.2	$k_{c1–c1}$	68,281,102,095	68,267,445,875	0.0
j_3	218	218	0.2	$k_{p11–p11}$	4,410,016,425	4,402,960,399	0.2
j_4	218	218	0.2	$k_{p12–p12}$	4,410,016,425	4,402,960,399	0.2
j_5	218	218	0.2	$k_{p13–p13}$	4,410,016,425	4,402,960,399	0.2
j_6	351	352	0.2	k_{s1-s1}	8,654,576,740	8,643,325,790	0.1
j_7	2654	2658	0.1	k_{c2-c2}	93,292,722,671	93,115,466,498	0.2
j_8	214	215	0.1	$k_{p21-p21}$	12,852,864,792	12,820,732,630	0.3
j_9	214	215	0.1	$k_{p22-p22}$	12,852,864,792	12,820,732,630	0.3
j_{10}	214	215	0.1	$k_{p23-p23}$	12,852,864,792	12,820,732,630	0.3
j_{11}	15	15	0.2	k_{s2-s2}	1,975,494,815	1,971,938,924	0.2
j_{12}	1624	1625	0.0	k_{W-W}	6,005,991,606	5,992,778,424	0.2
j_{13}	8	8	0.0	k_{P-P}	469,799,716	468,390,317	0.3
j_{14}	537	537	0.0	k_{G-G}	95,646,856	95,321,657	0.3

Table 4 Estimated stiffness by Estimated[M2]

Error (%) by using Estimated[M2]													
Parameter	k_{R-R}	k_{c1-c1}	$k_{p11-p11}$	$k_{p12-p12}$	$k_{p12-p12}$	$k_{p13-p13}$	k_{s1-s1}	k_{c2-c2}	$k_{p21-p21}$	$k_{p23-p23}$	k_{s2-s2}	k_{W-W}	k_{P-P}	k_{G-G}
10 modes	0.9	27.4	26.8	26.8	26.8	27.1	2.1	1.7	1.7	1.7	1.8	1.9	1.9	0.9
11 modes	1.9	17.3	16.8	16.8	16.8	17.1	2.0	1.9	1.9	1.9	2.1	0.8	1.9	0.8

tions show that high values of error in the input data of stiffness estimation problem, namely the estimated eigenfrequencies and inertia parameters, may cause instability of the nonlinear numerical solver in Eq. (16). From the stability perspective, the LS estimator outperforms the first method which is based on solving the nonlinear Eq. (16), because the eigenfrequencies of Λ in the LS estimator are forced to be positive, which ensures the convergence of the solver. In other words, *Estimated*[M2] is more robust to the input data, and is recommended also when all the frequency modes are available.

3.3 Loads, contact stress and accumulated damage

Contact load and stress validation on the three different gear stages The input and output torques derived in Sect. 2.5 and shown in Eq. (25) can be seen in Fig. 6. There one can see that all torques present similar oscillating pattern around their nominal values, and that there is little difference between the output and input torques at adjacent stages, see Fig. 6d–b and e–c.

One can estimate the stresses at a sun-planet gear pair by inserting the input torques shown in Fig. 4 at the first order approximation using Taylor series, based on Eq. (12). The results for such operation can be seen in Fig. 7, together with results from high-fidelity Simpack multi-body simulation model of NREL 5 MW drivetrain reported in [12]. The

results show reasonable agreement between the results for the simplified torsional model and the high-fidelity multibody simulation platform. (Fig. 7).

Calculation of average accumulated damage The SN curve parameters for pitting fatigue damage calculations depend on material, load and gear geometry. These two parameters for the pitting fatigue damage estimated in this paper are obtained based on ISO 6336-2:2019 [14]. Only one region is assumed for the SN curve. For instance, k is 3.051×10^{44}, and m is 12 for the sun gear of the 2nd stage. The numbers of stress cycles of sun gear is multiplied by 3, because it meshes with 3 gears simultaneously at each revolution. The accumulated damage of 1st stage sun, 2nd stage sun and 3rd stage pinion gears due to contact fatigue for one hour of operation by using the proposed digital twin-based remaining useful lifetime monitoring approach is listed in the Table 5. The online accumulated damage D of the 2nd stage sun gear is shown in the Fig. 8.

Table 5 Accumulated damage for 3600 s of operation

Gear	Damage
1st stage sun	0.00000482
2nd stage sun	0.00000456
3rd stage pinion	0.00000759

Fig. 6 Comparison between nominal and estimated torques. **a**–**c** input torques and **d**–**f** output torques at the 1st, 2nd, and 3rd stages respectively
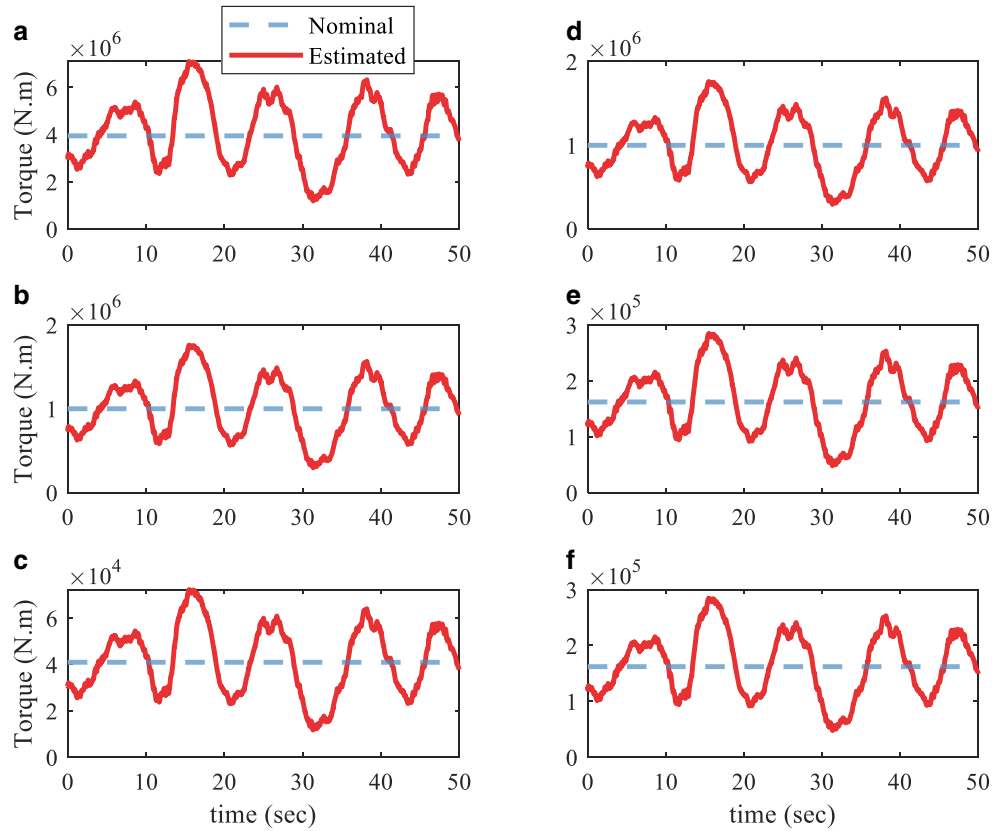
Fig. 7 Sun-planet contact stress comparison at second stage: results using estimated input torque and 1st order Taylor series (see Sect. 2.5) against Simpack simulation
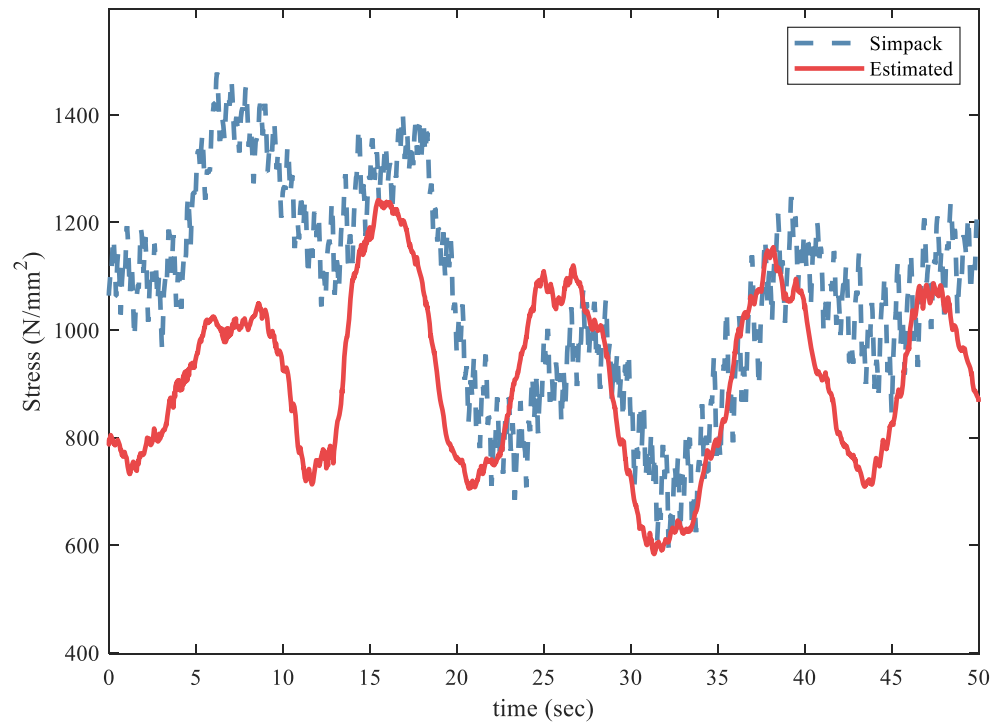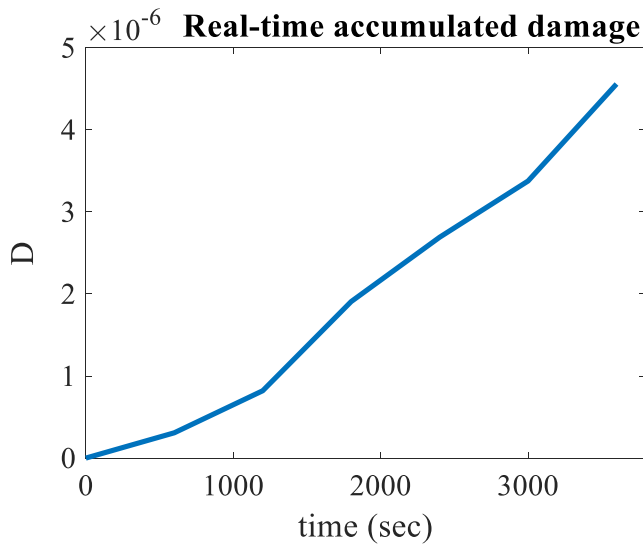
Fig. 8 Accumulated damage of the sun gear of the gearbox 2nd stage over one hour of operation

3.4 Practical implementation challenges and real-time performance

Challenges to overcome for practical implementation In practical turbines, uncertainties are an important issue. The different sources of uncertainties and their influences on the digital twin-based degradation estimation approach grounded on drivetrain torsional equivalent models for estimation of residual life of *the drivetrain shafts* are investigated by Moghadam and Nejad in [24]. The latter is performed by using stochastic models and statistical approaches, where uncertainties in both the model estimation and real-time measurements are addressed to mitigate their influence and improve the efficacy of the digital twin approach in presence of uncertainties. The statistical uncertainties in fatigue calculation due to material uncertainties can be accounted for by stochastic modelling of damage as explained by [24]. It is also worth noting that the employed least-square estimator is robust to the measurement noises, so that it can mitigate the influence of measurement noise in general, and cancel the influence of Gaussian noise [25]. The other issue in practical turbines is a call for computationally fast and inexpensive health monitoring techniques. The proposed algorithm is based on inexpensive linear equivalent models of the drivetrain, and the computational time depends on the sampling frequency and the number of input measurement samples. The algorithm can be managed to be executed only in order of a second. The method relies on a linear regression-based estimator and a few measurement samples, linear torque observers, and the real-time cycle counting of the estimated stress. The selected linear 14-DOF model of drivetrain is a compromise between simplicity and accuracy, which demonstrates similar results to

results of high fidelity models for estimation of load aimed at estimating the degradation of the gears. The practical implementation of the algorithm integrated with turbine fully automated control and monitoring systems means the automated pre-processing of the turbine measurements to continuously feed updated input data into the proposed system identification approach to estimate real-time equivalent model of drivetrain in the digital twin framework. This task aims at the optimization of data streaming and continuous processing architectures to deal with the experimental data aspect of the digital twin framework. Within the preprocessing stage, a significant step will be on time signal correlation methods to verify and ensure time synchronicity when the data comes from different sources by mapping them on common spaces and dealing with the effect of different samples rates. Another practical issue with the realization of this idea is the possible need for new sensor installations to access additional torsional measurements, communication links with the main operation control unit, and the processing power for the execution of algorithm in near real-time.

Real-time performance metrics As discussed in detail in Sect. 2, the digital twin algorithm consists of three main components, namely drivetrain equivalent model identification, load estimation and damage calculation. Therefore, for the simulation-based studies of this paper, the run-time of algorithm can be estimated by the summation of data recording time, time for estimation of equivalent model parameters, time for estimation of load and time for calculation of damage.

The model identification computation time, t^{model}, is mainly influenced by the linear regression estimator operation, which its computational time depends on the degree of model complexity and the number of input samples. For the case of 14-DOF model, the model identification algorithm can estimate the model parameters by less than 5% error only by using 14 data samples, where the estimation error can be reduced by increasing the number of input samples. By assuming 30 data samples, with the sampling frequency 300 Hz, the time length of recorded data block will be 0.1 sec. The processing time depends on the processing power, but can be managed to be only a fraction of a second. The estimation of load and stress computation time, t^{load}, is determined by the operation of designed torque observers which are simple arithmetic functions of estimated model parameters and measured real-time response, with the computational time very close to zero. The damage calculation computation time, t^{damage}, is reliant on real-time cycle counting operation on the data blocks of stress time series, which can be executed quite fast in fraction of a second.

["

12. Nejad AR, Guo Y, Gao Z, Moan T (2016) Development of a 5 MW reference gearbox for offshore wind turbines. Wind Energy 19(6):1089–1106

13. Kahraman A (1994) Natural modes of planetary gear trains. J Sound Vib 173(1):125–130

14. ISO6336-2 (2006) Calculation of load capacity of spur and helical gears, part 2: calculation of surface durability (pitting). International Organization for Standardization, Geneva

15. Pilanci M, Arikan O, Pinar MC (2010) Structured least squares problems and robust estimators. IEEE Trans Signal Process 58(5):2453–2465

16. Golub GH, Van Loan C (1979) Unsymmetric positive definite linear systems. Linear Algebra Appl 28:85–97

17. Nejad AR, Gao Z, Moan T (2014) On long-term fatigue damage and reliability analysis of gears under wind loads in offshore wind turbine drivetrains. Int J Fatigue 61:116–128

18. Downing SD, Socie DF (1982) Simple rainflow counting algorithms. Int J Fatigue 4(1):31–40

19. Manwell JF, McGowan JG, Rogers AL (2010) Wind energy explained: theory, design and application. John Wiley & Sons, United States

20. Liang X, Zhang H, Liu L, Zuo MJ (2016) The influence of tooth pitting on the mesh stiffness of a pair of external spur gears. Mech Mach Theory 106:1–15

21. Lei Y, Liu Z, Wang D, Yang X, Liu H, Lin J (2018) A probability distribution model of tooth pits for evaluating time-varying mesh stiffness of pitting gears. Mech Syst Signal Process 106:355–366

22. Liu J, Wang C, Wu W (2020) Research on meshing stiffness and vibration response of pitting fault gears with different degrees. Int J Rotating Mach. https://doi.org/10.1155/2020/4176430

23. Moghadam FK, Nejad AR (2020) Theoretical and experimental study of wind turbine drivetrain fault diagnosis by using torsional vibrations and modal estimation. J Sound Vib. (Under review)

24. Moghadam FK, Nejad AR (2020) Online condition monitoring of floating wind turbines drivetrain by using digital twin modeling approach. J Mech Syst Signal Process. (Under review)

25. Xu W, Chen W, Liang Y (2018) Feasibility study on the least square method for fitting non-Gaussian noise data. Phys A Stat Mech Appl 492:1917–1930

26. Duda KR, Prasov Z, York SP, West JJ, Robinson SK, Handley PM (2015) Development of an integrated simulation platform for real-time task performance assessment. In: 2015 IEEE Aerospace Conference. IEEE, , pp 1–9, New York City

27. Pratap V, Xu Q, Kahn J, Avidov G, Likhomanenko T, Hannun A, Collobert R et al (2020) Scaling up online speech recognition using convnets. arXiv:2001.09727