# Identification of wind turbine main shaft torsional loads from high-frequency SCADA measurements using an inverse problem approach

W. Dheelibun Remigius[1] and Anand Natarajan[1]

[1]Department of Wind Energy, Technical University of Denmark, Frederiksborgvej 399, 4000 Roskilde, Denmark.

**Correspondence:** W. Dheelibun Remigius (drwp@dtu.dk)

**Abstract.** To assess the structural health and remaining useful life of ~~a wind turbine~~ wind turbines within wind farms~~one would require~~ , the site-specific ~~dynamic quantities such as~~ structural response and modal parameters of the primary structures are required. In this regard, a novel inverse problem-based methodology is proposed here to identify the dynamic quantities of the ~~drive train~~ drivetrain main shaft, *i.e.,* torsional displacement and coupled stiffness. As a model-based approach, an inverse problem of a mathematical model concerning the coupled shaft torsional dynamics with high-frequency SCADA measurements as input is solved. It involves Tikhonov regularisation to ~~smoothen~~ minimize the measurement noise and irregularities on the shaft torsional displacement obtained from measured rotor and generator speed. Subsequently, the regularised torsional displacement along with necessary SCADA measurements is used as an input ~~for~~ to the mathematical model and a model-based system identification method called the ~~collage~~ Collage method is employed to estimate the coupled torsional stiffness. It is also demonstrated that the estimated shaft torsional displacement and coupled stiffness can be used to identify the site-specific main shaft torsional loads. It is shown that the torsional loads estimated by the proposed methodology is in good agreement with the aeroelastic simulations of the Vestas V52 wind turbine. Upon successful verification, the proposed methodology is applied to the V52 turbine ~~SCADA measurements~~ to identify the site-specific main shaft torsional loads and damage equivalent load. Since the proposed methodology does not require a design basis or additional measurement sensors, it can be directly applied to wind turbines within a wind farm ~~irrespective of their age~~that possess high-frequency SCADA measurements.

## 1 Introduction

Monitoring of wind turbines within wind farms is increasingly becoming very important due to the need to detect ~~specific turbines that show anomalous behavior~~anomalous behaviour, plan inspections or preventive maintenance, and ~~to~~ compute the remaining useful life of specific structures. ~~However,~~ The site-specific structural dynamic quantities such as structural response and modal parameters could assist in the condition monitoring of wind turbines.

The structural response of the wind turbines are measured using load instrumentation such as accelerometers (Koukoura et al., 2015; Pahr and an output-based operational modal analysis (OMA) (Wang et al., 2016) technique is employed for the identification of the modal parameters. OMA methods can be broadly classified into two categories: (i) Time domain-based methods, and (ii) Frequency domain-based methods (Zhang et al., 2010). A recent comprehensive review on various time domain and frequency domain-based OMA methods can be found in Zahid et al. (2020). Upon identification of the modal parameters together with the measured structural response, an inverse problem-based technique is employed for the estimation of the site-specific loads (Pahn et al., 2017). Alternatively, one can use strain gauge based load sensors to measure the site-specific loads directly (IEC 61400-13 .

The addition of new instrumentation to existing turbines, such as the installation of strain gauges ~~,~~ and accelerometers can be costly and also require repetitive calibration and synchronization of their measurement signals with the turbine computer. Most wind farm operators also do not possess the aeroelastic design parameters of their wind turbines and hence cannot simulate the mechanical loads acting on their wind ~~turbine. Monitoring of~~ turbines. Monitoring turbine primary structures through existing Supervisory Control and Data Acquisition (SCADA) ~~system based~~ system-based measurements can allow ~~cost effectiveness~~ cost-effectiveness and provide valuable information to the wind farm operator. Usually, such monitoring through SCADA only provides information on power performance and not regarding turbine structural integrity. ~~Here we develop mathematical models for~~ Since there are two SCADA signals (i.e., rotor speed and generator speed) related to torsional oscillations of the main shaft of ~~a wind turbine that can determine both the coupled torsional stiffness~~ wind turbine drivetrains, the same can be used to quantify the torsional dynamics of the main shaft.

For this purpose, an inverse problem-based approach is developed here to determine the torsional stiffness and response of the main shaft ~~and its torsional displacement in continuous time-domain~~of a wind turbine, using existing ~~time series measurements of~~ high-frequency SCADA measurements such as the rotor speed and generator speed. This is a ~~novel methodology that can potentially benefit wind farm owners, since both the property of the structure in terms of its stiffness, and the structural response can be determined without requiring additional sensors or information from~~ cost-effective alternative approach that is being proposed for the main shaft without using any additional measurement sensors or an aeroelastic design basis of the wind turbine~~manufacturer.~~

~~There are many studies available in the literature on system identification of wind turbines (Koukoura et al., 2015; Pahn et al., 2017; Noré . All of the studies employed output-based operational modal analysis (OMA) (Wang et al., 2016) technique on the measured structural response to estimate the modal parameters of the structure. OMA methods can be broadly classified into two categories: (i) Time domain-based methods, and (ii) Frequency domain-based methods (Zhang et al., 2010). Time-domain based OMA methods are based on the calculation of auto and cross-correlation functions of the response time histories. Since they possess similar properties as frequency response functions, modal parameters can be extracted from those correlation functions. Several algorithms are available to perform this task and few such methods are natural excitation technique (NExT) (James et al., , random decrement technique (RDT) (Ibrahim, 1977), auto-regression moving average vector model (ARMAV) (Andersen, 1997) and stochastic subspace identification (SSI) method (Van Overschee and De Moor, 1993). Frequency-domain based OMA methods are peak picking method, enhanced frequency domain decomposition method(FDD) (Brincker et al., 2001) and frequency and spatial domain decomposition method (FSDD) (Wang et al., 2005). A recent comprehensive review on various time domain and frequency domain-based OMA methods can be found in Zahid et al. (2020). However, to the best of the authors' knowledge, there is no such study available on estimating the structural response of a wind turbine component from SCADA measurements.~~

~~Thus, a novel~~ . The proposed inverse problem approach is a model-based approach ~~is proposed here for the wind turbine drive train main shaft to estimate both the structural response and the modal parameters simultaneously.~~

~~A~~ whereby a mathematical model concerning the shaft torsional dynamics will be utilized to obtain ~~the system response from SCADA measurements. In this context, the~~ both the torsional displacement and the coupled torsional stiffness in a continuous time domain. It involves Tikhonov regularisation (Tikhonov, 1963) for regularising the measurement data and the Collage method (Kunze and Vrscay, 1999) for estimating the torsional stiffness.

The concerned mathematical model comprised of differential equations will be solved for the shaft torsional velocity with high-frequency rotor and generator speed measurements as inputs. Subsequently, the main shaft torsional displacement is obtained by numerically integrating the shaft torsional velocity. However, ~~these time integration schemes are~~ numerical integration is based on time-marching algorithms, and the lack of initial conditions ~~make the~~ makes displacement reconstruction an ill-posed problem (Hong et al., 2008). ~~Further, these time marching algorithms are sensitive to measurement noise and they even~~ Since it is an ill-posed problem, influence of measurement noise will get amplified during the time marching procedure ~~which results in inadmissible errors in the reconstructed displacement. Also, this inaccurate displacement~~ and results in an erroneous displacement. This inaccurate displacement also leads to drastic errors in the system modal parameter estimation. Hence, one needs to go for regularisation techniques to smoothen the reconstructed displacement. Hansen (2005) discussed the nature of various ill-posed problems and presented ~~a number of~~ several solution methodologies. Though there are many regularisation techniques available, Tikhonov, ~~Truncated~~ truncated singular value decomposition (SVD) ~~,~~ and nuclear norm are a few of the popular techniques (Aarden, 2017). Among all these regularisation techniques, Tikhonov regularisation ~~(Tikhonov, 1963)~~ has been widely used in many engineering applications ~~(Ronasi et al., 2011; Hào and Quyen, 2012; Bangji et al., 2017; Nieminen et al., 2011) and also~~ (Ronasi et al., 2011; Hào and Quyen, 2012; Bangji et al., 2017; Nieminen et al., 2011) and it has been studied extensively in the field of inverse problems (Hansen, 2005) ~~.~~ as well. Further, digital filters and frequency domain integration approach (FDIA) are also widely used techniques in the literature to reconstruct displacements from measured accelerations (Hong et al., 2008; Brandt and Brincker, 2014; Qihe, 2019; Lee et al., 2010). However, digital filters such as impulse response filters (IIR) and finite response filters (FIR) have several drawbacks when reconstructing the low-frequency dominant displacements as is the case here (Lee et al., 2010). On the other hand, the FDIA methods are sensitive to the time interval of the measurements (Lee et al., 2010). It is shown by Lee et al. (2010) that the Tikhonov regularisation is better suited for low-frequency dominant structures. Hence, the same has been employed in the present work. Tikhonov regularization minimizes the error using the least-square criterion and by means of ~~a~~ numerical damping, it also minimizes the effect of ~~the~~ measurement noise.

~~Together with~~ Upon obtaining the regularised shaft torsional displacement, the same mathematical model is utilized to obtain the shaft stiffness. For this purpose, the Collage method - a model-based system identification technique ~~called the collage method is used in the present study~~ - is used (Kunze and Vrscay, 1999; Groetsch, 1993). ~~The model-based collage method have~~ This method has been successfully applied for the system identification in various differential ~~equations based~~ equations-based problems such as boundary value problems (Kunze et al., 2009), reaction-diffusion problems (Deng et al., 2008) and elliptic problems (Kunze and La Torre, 2016). The ~~collage method is used to convert the inverse problem of system identification~~ Collage method transforms the system identification problem into a minimization problem of a function of several variables (for example unknown system parameters) and then the ~~corresponding~~ minimization problem is solved using a ~~suitable algorithm . The minimization procedure is referred to as Collage coding and it~~ minimisation algorithm called the Collage coding (Kunze et al., 2009). Collage coding is a greedy algorithm that ~~seeks to construct an approximate solution to a target solution in one go. Hence, unlike other inverse problem methods , one need not solve the forward problem in an iterative manner (Deng and Liao, 2009).~~attempts to find the approximate solution in a single step without any need for iteration, as is the case for other inverse problem-based methods (Deng and Liao, 2009). Further, the model-based Collage method is simple, easy to implement and computationally inexpensive as compared to the output-based OMA methods.

~~One of the key benefits of estimating the~~ The estimated shaft torsional stiffness and displacement ~~is that it can be~~ are further used to identify the site-specific shaft torsional load~~and the remaining useful life (RUL) (Ziegler et al., 2018) of the main shaft. Also, the~~ . This novel methodology can potentially benefit wind farm owners since both the property of the structure in terms of its stiffness and the structural response and the site-specific load can be determined without requiring additional sensors or information from the wind turbine manufacturer. The main shaft torsional load ~~can significantly affect~~ affects the fatigue performance of other ~~drive train~~ drivetrain components such as gearbox and planetary bearings (Dong et al., 2012; Gallego-Calderon and Natarajan, 2015; Ding et al., 2018). Hence, the same site-specific torsional load ~~can also be used to quantify the RUL of gearbox and other drive train~~ may be used as an input for quantifying the remaining useful life (RUL) (Ziegler et al., 2018) of the main shaft, gearbox, and other drivetrain components as well. ~~Also, several older turbines possess only SCADA measurementsby default and they lack measurement sensors, the proposed methodology can~~ The estimation of RUL/ yearly damage does not require additional historical weather data and condition monitoring data as the wind speed and wind direction measurements are available in the SCADA measurements. However, this is beyond the scope of present work. Further, the proposed approach requires that the sampling frequency of the SCADA measurement be significantly higher than the dominant frequencies of the drivetrain torsional oscillations (i.e., 1p and 3p rotor excitation frequencies and torsional natural frequencies). As a result, the proposed method cannot be used for ~~estimation of the site-specific loads of them.~~ the turbines that have measurements in terms of 10-min SCADA statistics.

The rest of the paper is organised as follows: the problem formulation consisting of the Tikhonov regularisation and the ~~collage~~ Collage method is given in section 2; section 3 presents the verification of the proposed formulation; application of the proposed formulation on measurements are presented in section 4.

## 2 Problem formulation

As mentioned in the previous section, the main objective is to identify the shaft torsional displacement and coupled stiffness from SCADA measurements. This is achieved by solving the shaft torsional dynamical equations using a suitable inverse problem algorithm and the estimated shaft torsional displacement ~~$\theta$~~ $\boldsymbol{\theta}$ and torsional stiffness $K$ will be utilized for the shaft torsional load estimation. For this purpose, a two-mass model (refer Fig. 1) ~~(Boukhezzar et al., 2007; Berglind et al., 2015) which~~ (Boukhezzar et al., 2007; Girsang et al., 2013; Berglind et al., 2015) that governs the main shaft torsional dynamics subjected to the rotor and generator torques ~~$T_r$ and $T_g$~~ $\boldsymbol{T}_r$ and $\boldsymbol{T}_g$, respectively, is considered, and the mathematical model is given by Eqs.

(1-3). ~~It is assumed that the gearbox is perfectly stiff while transferring deformations on the main shaft and the gear ratio, $N$ is only considered as a parameter to proportionally adjust the force and torsional displacement between the high-speed and main shafts. The main shaft is modelled by an inertia free viscously damped torsional spring. Further, the edgewise flexibility of the blade and the torsional stiffness of~~ The governing equations are obtained in an inertial frame of reference and converted to the ~~main shaft is assumed to be combined in $K$~~ low-speed side of the drivetrain by means of the gear ratio (Boukhezzar et al., 2007). By assuming that the drivetrain components are in a series representation in terms of its modal quantities, effective values for the modal quantities are used in the two-mass model (Girsang et al., 2013).

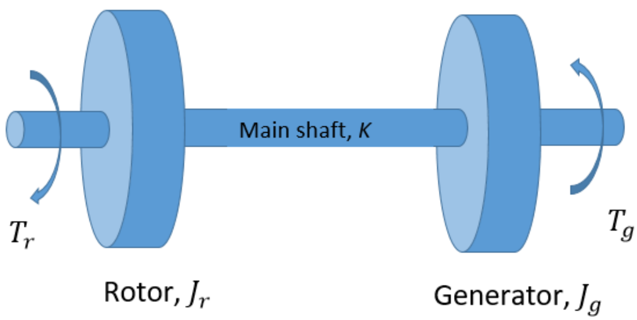



**Figure 1.** A two mass model of wind turbine ~~drive train~~drivetrain.

$$J_r\dot{\boldsymbol{\omega}}_r = T\boldsymbol{T}_r - K\theta\boldsymbol{\theta} - C\dot{\boldsymbol{\theta}}, \tag{1}$$

$$J_g\dot{\boldsymbol{\omega}}_g = -T - \boldsymbol{T}_g + K/N\theta\boldsymbol{\theta} + C/N\dot{\boldsymbol{\theta}}, \tag{2}$$

$$\dot{\boldsymbol{\theta}} = \omega\boldsymbol{\omega}_r - \omega\boldsymbol{\omega}_g/N. \tag{3}$$

Here, $J_r$ represents the inertia of the rotor, $J_g$ represents the collective inertias of the high-speed shaft (HSS), the gearbox, and the generator, ~~$\omega_r$ and $\omega_g$~~ $\boldsymbol{\omega}_r$ and $\boldsymbol{\omega}_g$ are the rotor and generator speeds, respectively, $K$ and $C$ ~~is the shaft damping coefficient and $\theta$ is the shaft torsional displacement.~~

~~For normal operation of the wind turbines, the shaft torsional dynamics is mainly influenced by the low frequency modes such as blade edgewise and shaft free-free modes and the high frequency gearboxdynamics do not play a significant role in it. Hence,~~ are the effective stiffness and damping of the drivetrain including the main shaft, HSS and gearbox, and $\boldsymbol{\theta}$ is the ~~two-mass model is sufficient enough to model the shaft torsional dynamics for the wind turbine normal operations as it includes both the low frequency modes. Also, given the system parameters and rotor and generator torques, the two-mass model is capable of predicting the shaft torsional displacement ($\theta$) as close as that of the full-fledged aeroelastic simulation as shown in Fig. ??. Here, the aeroelastic simulation is performed in the DTU in-house tool called HAWC2 (Larsen and Hansen, 2007) and the results are obtained for the Vestas V52 (Vestas) wind turbine at a mean wind speed of 8 $m/s$~~torsional displacement of the drivetrain. Throughout the article, all vector quantities have been marked with bold font.

~~Comparison of the estimated shaft torsional displacement ($\theta$) with the aeroelastic simulation result.~~

~~In forward problem approach, given~~

Given the modal parameters $(J_r, J_g, C, K)$ and external torques (~~$T_r$ and $T_g$~~$\boldsymbol{T}_r$ and $\boldsymbol{T}_g$), Eqs (1-3) are solved for ~~$\omega_r, \omega_g$ and $\theta$~~$\boldsymbol{\omega}_r, \boldsymbol{\omega}_g$ and $\boldsymbol{\theta}$, which is known as a forward problem approach (Pahn, 2013). But given only SCADA measurements, ~~one has to solve~~ Eqs. (1-3) ~~inversely for $\theta$~~ are solved inversely for $\boldsymbol{\theta}$ and modal parameters. ~~In general, the~~ The available SCADA measurements are ~~$\omega_r, \omega_g, P, \beta, U$~~$\boldsymbol{\omega}_r, \boldsymbol{\omega}_g, P, \beta, \boldsymbol{U}$. Here, ~~$P, \beta, U$~~ $P, \beta, \boldsymbol{U}$ are, respectively, the generator power, blade pitch angle and wind velocity. The proposed inverse problem approach consists of Tikhonov regularisation for regularising the measurement data and the Collage method for estimating the torsional stiffness and the entire methodology is shown as a flow chart in Fig. 2. In the following, implementation of the Tikhonov regularisation and Collage method on the drivetrain torsional

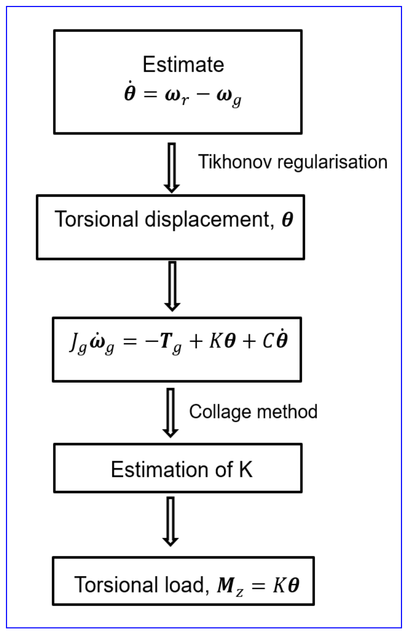



**Figure 2.** Flowchart depicting the inverse problem algorithm.

dynamics will be discussed in detail.

### 2.1 ~~Collage methodGiven $\omega_r$ and $\omega_g$, it is straightforward to use~~ Tikhonov Regularisation

Given $\boldsymbol{\omega}_r$ and $\boldsymbol{\omega}_g$, $\dot{\boldsymbol{\theta}}$ is obtained by using Eq. (3) ~~to obtain $\dot{\theta}$ and then $\theta$. The next step is to estimate the modal parameters that are required for the load calculation. The collage method (Kunze and Vrscay, 1999; Groetsch, 1993) is used for the same in the present study since it is a model-based approach. For a given initial value problem (IVP) ,~~

$$x(t) = f(x,t), x(0) = x_0,$$

~~that admits a target solution $x(t)$~~and then by numerically integrating $\dot{\boldsymbol{\theta}}$, the ~~associated Picard integral operator $T$ is given by,~~

$$(Tu)(t) = x_0 + \int\limits_0^t f(u(s),s)\,\mathrm{d}s.$$

shaft torsional displacement ($\boldsymbol{\theta}$) is obtained. The numerical integration schemes require initial conditions as they march on time. However, the initial conditions are unavailable or inaccurate in practice. The lack of initial conditions (assumed usually 0) on the displacement are inconsistent with the real values that result in the phenomenon of baseline shift or drift which causes the position error to grow with time during the integration (Pahn et al., 2017). The effect of the lack of initial conditions on the reconstructed displacement obtained using the time integration scheme is shown in Fig. 3. As seen in the figure, the numerical error is multiplicatively increased with time which results in a drift in the reconstructed displacement. As mentioned in the introduction, to minimize the numerical error due to the lack of initial conditions and to minimize the effect of measurement noise, a widely used regularization technique called Tikhonov regularization (Tikhonov, 1963) is employed here.

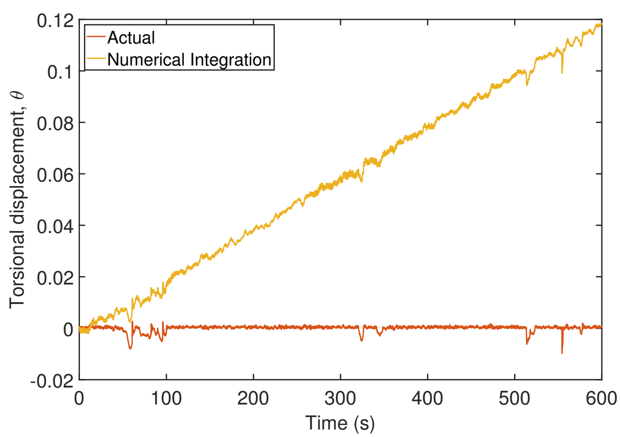



**Figure 3.** Comparison of time integration displacement with actual displacement.

Implementation of Tikhonov regularization on the velocity to obtain the displacement is not readily available in the literature and hence the same is presented in the Appendix A.1 for the sake of completeness. By following the procedure outlined in Appendix A.1, the regularised torsional displacement ($\boldsymbol{\theta}$) is obtained as,

$$\boldsymbol{\theta} = \left(\frac{\mathbf{L}^2}{4} + \lambda^2\mathbf{I}\right)^{-1}\frac{\mathbf{L}\mathbf{L}_a\dot{\boldsymbol{\theta}}\Delta t}{2}. \tag{4}$$

~~The assumptions regarding the parameter estimation problem using the collage method are listed as follow (Deng and Liao, 2009) : $x(t) \in [t_0, t_n]$ is a bounded solution; where, $t_0$~~ Here, the matrices $\mathbf{L}, \mathbf{L}_a$ and ~~$t_n$ are positive constants satisfying $t_0 < t_n$. $f_i(u, D, x, t, \lambda_1, \cdots, \lambda_m)(0 \leq i \leq n)$ are continuous, where, $\lambda_m$ is the unknown modal parameter. The exact solution $x(t)$ of the system (22) exists uniquely.~~ $\mathbf{I}$ are defined in the Appendix A.1, $\lambda$ is the regualrisation parameter, $\Delta t$ is the time interval and

$\dot{\boldsymbol{\theta}}$ is the torsional velocity obtained from Eq. (3). The obtained regularised torsional displacement is compared with the actual displacement in Fig. 4 along with the numerical integration result. As seen in Fig. 4, there is a close match between the result by Tikhonov regularisation and the actual displacement as compared to the numerical integration result.

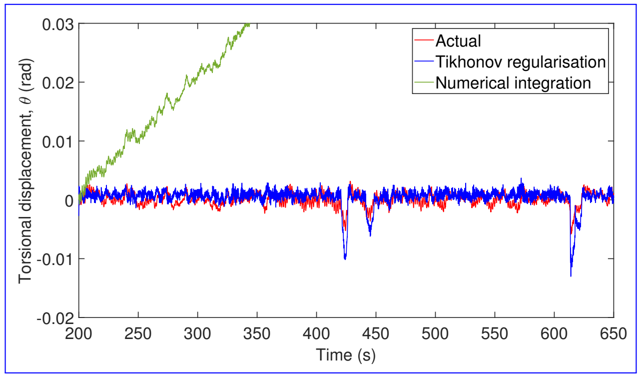



**Figure 4.** Comparison of the Tikhonov and time integration displacements with actual displacement.

~~It is important to note that the fixed point $\bar{u}(t)$ of this Picard operator is the unique solution of~~

## 2.2 Collage method

Upon estimating $\boldsymbol{\theta}$ using Tikhonov regularisation, next step is to estimate the ~~given IVP (Kunze et al., 2004). Accordingly, the collage distance becomes, $(x - Tx)$ and then the optimal solution is the one which minimizes the squared collage distance *i.e.,* $L^2$ collage distance. Also, unlike the conventional inverse problem which minimises the approximate error $d(x - \bar{x})$, the collage method minimizes the collage distance $d(x, Tx)$ which is an useful change as one cannot find $\bar{x}$ for a general $T$ (Kunze et al., 2004). Further, the optimality of the collage distance minimization is ensured as shown by Kunze et al. (2004). Minimising the $L^2$ collage distance using least square method yields a stationarity condition, $\frac{\mathrm{d}(x-Tx)}{\mathrm{d}\lambda_m} = 0$, which results in a set of linear equations in terms of the unknown modal parameters $(\lambda_m)$. By solving these equations, the~~ modal parameters ~~are estimated.~~

~~When concerning the shaft dynamics, the~~ using the Collage method (Kunze and Vrscay, 1999; Groetsch, 1993) - a model-based approach. The mathematical formulation of the Collage method and its implementation is given in Appendix A.2. The rotor equation (Eq. 1) cannot be used for the parameter estimation as there is no information about the rotor torque in the SCADA measurement. Instead, the ~~collage method is applied on~~ Collage method is employed on the generator equation (Eq. 2) ~~to estimate~~ for estimating the modal parameters as the generator torque (~~$T_g$~~$\boldsymbol{T}_g$) can be readily obtained from the SCADA data as, ~~$T_g = P/\omega_g$~~ $\boldsymbol{T}_g = P/\boldsymbol{\omega}_g$. Accordingly, ~~the Picard integral~~ for the target function $\boldsymbol{\omega}_g(t), t \in [t_0, t_n]$, the squared $\mathcal{L}^2$ collage distance (refer Appendix A.2) for Eq. (2) becomes,

$$\cancel{(Tu)(t)}\,\Delta^2 = \int_{t_0}^{t_n} \left[ J_g \cancel{[\omega_g - \omega_{g0}]_g} [\boldsymbol{\omega}_g - \boldsymbol{\omega}_{g0}] + \int_{\cancel{0}\, t_0}^{\cancel{t}\, t_n} \cancel{T}\boldsymbol{T}_g \,\mathrm{d}t - K \int_{\cancel{0}\, t_0}^{\cancel{t}\, t_n} \cancel{\theta \mathrm{d}}\boldsymbol{\theta} \mathrm{d}t - C \int_{\cancel{0}\, t_0}^{\cancel{t}\, t_n} \dot{\boldsymbol{\theta}} \mathrm{d}t \right]^2 \mathrm{d}t, \tag{5}$$

| Wind turbine | Design $K$ (~~$N.m$~~ $N.m/rad$) | Estimated $K$ (~~$N.m$~~ $N.m/rad$) | % error |
|---|---|---|---|
| DTU-10 MW | 2.317E09 | 2.1785E09 | 5.98 |
| Vestas V52 | - | - | 6.98 |

**Table 1.** Comparison of estimated $K$ with design $K$ for a turbine variant where only the main shaft is flexible.

where, ~~$t$ is the time variable and $\omega_{g0}$ is the~~ $\boldsymbol{\omega}_{g_0}$ is the generator speed at time, ~~$t=0$~~ $t=t_0$. Modal parameters will be obtained by ~~minimizing~~ minimising Eq. (5) ~~using the least square method.~~ with respect to $J_g, C$ and $K$. Upon obtaining $\boldsymbol{\theta}$ using the Tikhonov regularisation and $K$ using the Collage method, the torsional load is obtained as $\boldsymbol{M}_z = K\boldsymbol{\theta}$.

### 2.2.1 Application of the Collage method

To ~~the best of the authors' knowledge, this collage method has not been used in the context of wind turbine system identification and hence it is important to check the applicability of this method for the same~~test the applicability and efficiency of the Collage method for the wind turbine drivetrain system, a verification study is undertaken by comparing the main shaft torsional stiffness obtained using the Collage method with its design value for two different wind turbines. This is done ~~by estimating the modal parameters using the collage method~~ for the following two wind turbines~~and verifying with its design values~~, (i) DTU ~~10-MW~~10 MW (Bak et al., 2013) and (ii) Vestas-V52 (Vestas). For facilitating a comparison of the main shaft torsional response alone, the rigid variant of the turbine is chosen and this implies that the rotor and tower are rigid and the main shaft alone is considered to be flexible. By performing ~~HAWC2~~ aeroelastic simulations on these turbines, the shaft torsional displacement ~~$\theta$~~ $\boldsymbol{\theta}$ is obtained. ~~Then,~~ Throughout the article, the aeroelastic simulation is performed in the DTU in-house tool called HAWC2 (Larsen and Hansen, 2007). HAWC2 (Horizontal Axis Wind turbine simulation Code 2nd generation) is used for calculating the wind turbine aeroelastic loads and responses in time domain. It uses multibody formulation to model the structure, blade element momentum (BEM) theory-based models for modelling the wind effects and hydrodynamic models for modeling the hydro-effects (in the case of offshore turbines) on the structure. Control of the turbine is performed through the dynamic link libraries (DLLs). Using the main shaft torsional load time series obtained from HAWC2 and, by minimising Eq. (5) concerning the modal parameters, $K, C$, and $J_g$ are obtained. Since for the estimation of shaft torsional load, only $K$ is needed among all the modal parameters, hence the same is compared with the design values. The estimated shaft stiffness and the stiffness from the aeroelastic model of the DTU 10-MW turbine along with percentage error are tabulated in Table. ~~??. Due to confidentiality policy, only~~ 1. Only the percentage error is given for the Vestas V52 turbine ~~.~~ in Table. 1. As seen in the table, the estimated torsional stiffness values match well with the design values. If the torsional displacement ~~$(\theta)$~~ $(\boldsymbol{\theta})$ is known, then the determination of torsional stiffness ($K$) from Eq. (5) is readily feasible as explained. However in practice, the shaft torsional displacement is unknown, and therefore the ~~collage~~ Collage equations may not be directly used to determine the shaft stiffness. In the following, the entire proposed methodology will be verified with the aeroelastic simulation results of the Vestas V52 turbine.

## 2.3 ~~Regularisation~~

~~As explained earlier, with $\omega_r$ and $\omega_g$, $\dot{\theta}$ is obtained by using Eq. (3) and then by using time integration schemes on $\dot{\theta}$, the shaft torsional displacement $(\theta)$ is obtained. However, these time integration schemes are based on time-marching algorithms, they require initial condition on displacement which are usually unavailable or inaccurate in real situations. The effect of the lack of initial conditions on the reconstructed displacement obtained using the time integration scheme is shown in Fig. 3. As seen in the figure, the numerical error is multiplicatively increased with time which results in a drift in the reconstructed displacement. As mentioned in the introduction, to minimize the numerical error due to the lack of initial conditions and to subsidize the effect of measurement noise, a widely used regularization technique called Tikhonov regularization (Tikhonov, 1963) is employed here. Comparison of time integration displacement with actual displacement.~~

~~Implementation of Tikhonov regularization on the velocity to obtain the displacement is not readily available in the literature, the same is presented here for the sake of completeness. By definition, the velocity $\dot{\theta}$ is expressed as,~~

$$\underline{\dot{\theta}(t) = \frac{\mathrm{d}\theta}{\mathrm{d}t} \approx \dot{\tilde{\theta}}(t),}$$

~~where, $\dot{\tilde{\theta}}(t)$ is the velocity obtained from Eq. (3) which can be considered as a measured velocity. As explained earlier, the lack of initial conditions in addition to the measurement noise leads to erroneous displacement. In order to minimise the error, following minimisation problem has to be solved,~~

$$\underline{\mathrm{Min}} \; \underline{\Pi_E(\theta) = \frac{1}{2}\int_{t_1}^{t_2} (\dot{\theta}(t) - \dot{\tilde{\theta}}(t))^2 \, \mathrm{d}t.}$$

~~Here, $\dot{\theta}$ is the calculated velocity. By means of the trapezoidal rule, Eq. (12) is discretised as follows (Hong et al., 2008),~~

$$\underline{\Pi_E(\theta) \approx \|\mathbf{L}_a(\dot{\theta} - \dot{\tilde{\theta}})\|_2^2 \Delta t,}$$

~~where $\Delta t$ is the time interval of the discretization and $\mathbf{L}_a$ is the diagonal weighing matrix of order $(n+1)$ as,~~

$$\underline{\mathbf{L}_a = \begin{bmatrix} 1/\sqrt{2} & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1/\sqrt{2} \end{bmatrix}.}$$

~~Further, the calculated velocity $\dot{\theta}$ is discretised by the central difference rule and written in matrix form as,~~

$$\underline{\frac{1}{\Delta t}\mathbf{L}_c\theta = \dot{\theta},}$$

~~where, the central difference matrix $\mathbf{L}_c$ of size $(n+1)\times(n+3)$ and the displacement vector $\theta$ of size $(n+3)$ are given as,~~

$$\underline{\mathbf{L}}_c = \begin{bmatrix} 1 & 0 & 1 & \dots & & \\ & -1 & 0 & 1 & & \\ & & \ddots & & & \\ & -1 & 0 & 1 & & \\ & \dots & -1 & 0 & 1 \end{bmatrix}, \qquad \theta = \begin{pmatrix} \theta_{-1} \\ \theta_0 \\ \vdots \\ \theta_n \\ \theta_{n+1} \end{pmatrix}.$$

~~Here, the time steps denoted by $-1$ and $(n+1)$ are fictitious nodes. Substitution of Eq. (15) into Eq. (13) leads to,~~

$$\underline{\text{Min}}\ \Pi_E(\theta) \approx \frac{1}{2}\left\| \frac{1}{2\Delta t}\mathbf{L}_a\mathbf{L}_c\theta - \underline{\mathbf{L}}_a\ddot{\theta}\right\|_2^2 \Delta t = \frac{1}{2}\left\|\frac{1}{2}\mathbf{L}\theta - \underline{\mathbf{L}}_a\ddot{\theta}\Delta t\right\|_2^2 \frac{1}{\Delta t},$$

~~where, $\mathbf{L} = \mathbf{L}_a\mathbf{L}_c$. This minimisation problem is regularised for solution boundedness with a parameter $\lambda$, and given as,~~

$$\underline{\text{Min}}\ \Pi_E(\theta) \approx \frac{1}{2}\left\|\frac{1}{2}\mathbf{L}\theta - \underline{\mathbf{L}}_a\ddot{\theta}\Delta t\right\|_2^2 + \frac{\lambda^2}{2}\|\theta\|_2^2.$$

~~The above minimisation problem is known as the Tikhonov regularisation and $\lambda$ is referred to as the regularisation parameter. Minimising Eq. (18) as,~~

$$\frac{\mathrm{d}\Pi_E}{\mathrm{d}\theta} = \frac{1}{2}\left(\frac{\mathbf{L}^2\theta}{2} - \underline{\mathbf{L}\mathbf{L}}_a\ddot{\theta}\Delta t\right) + \lambda^2\theta = 0,$$

~~yields the following quadratic equation in $\theta$,~~

$$\theta = \left(\frac{\mathbf{L}^2}{4} + \lambda^2\mathbf{I}\right)^{-1}\frac{\mathbf{L}\mathbf{L}_a\ddot{\theta}\Delta t}{2},$$

~~where, $\mathbf{I}$ is the identity matrix of order $(n+3)$. Since ten-minute SCADA measurements with a sampling frequency of 50 $Hz$ are considered for $\theta(t)$ estimation, $n$ becomes 30000.~~

~~The choice of regularisation parameter $(\lambda)$ plays a crucial role in getting an optimal fit for the solution. Based on the knowledge about measurement errors, Hansen (2005) proposed two classes for the estimation of $\lambda$:~~

- ~~methods based on knowledge of measurement errors~~
- ~~methods that do not require details about measurement errors.~~

~~In the present scenario, the information regarding the measurement error is unknown, hence class two is used for the current study. In class two, there are three widely used methods (Nieminen et al., 2011): (i) quasi optimality criterion, (ii) Generalized gross validation (GCV), and (iii) L-curve method. Compared to the GCV method, the other two methods give a better estimate of $\lambda$ (Gao et al., 2016). Further, for larger problems, the quasi optimality method is computationally expensive than the L-curve~~

~~method. Owing to this fact, the L-curve method is used here for estimating $\lambda$. In L-curve method, the optimal $\lambda$ is the one which gives the maximum curvature in the L-curve between norm of the regularized solution $\alpha(\lambda) = \|\theta_{reg}\|_2$ and norm of the residual $\beta(\lambda) = \left\|\frac{1}{2}\mathbf{L}\theta - \mathbf{L}_a\dot{\tilde{\theta}}\Delta t\right\|_2$ and the curvature of the L-curve is given by (Nieminen et al., 2011),~~

$$\kappa(\lambda) = \frac{\ddot{\alpha}\dot{\beta} - \ddot{\beta}\dot{\alpha}}{[\dot{\alpha}^2 + \dot{\beta}^2]^{3/2}}.$$

~~Substituting the optimal $\lambda$ obtained by finding the maximum curvature of Eq. (21) and $\dot{\tilde{\theta}}(t)$ obtained from Eq. (3) in Eq. (4), the regularised displacement ($\theta_{reg}$) is obtained. The obtained regularised torsional displacement is compared with the actual displacement in Fig. 4.Also, the displacement obtained from the numerical integration technique also presented in the same figure. As seen in the figure, result using Tikhonov regularisation matches closely with the actual displacement as compared to the displacement obtained by numerical integration. Comparison of the Tikhonov and time integration displacements with actual displacement.~~

~~Upon obtaining $\theta$ using the Tikhonov regularisation, Eq. (2) is then solved inversely with regularized $\theta$ using the collage method for the estimation of $K$ and then the torsional load is obtained as $M_z = K\theta$. This entire methodology is shown as a flow chart in Fig. 2. Flowchart depicting the inverse problem algorithm.~~

## 3 Verification of the method on V52 turbine simulations

~~To verify the proposed methodology for the torsional load estimation described in section 2, aeroelastic~~ Aeroelastic simulations are performed for the Vestas V52 turbine corresponding to the design load case (DLC 1.2) ~~(Hansen et al., 2015)~~ (IEC 61400-1) in HAWC2. The DLC 1.2 ~~consists of 18~~ was run with eighteen 10-minute load simulations (three yaw directions: 0 deg. $\pm$ 10 deg and six turbulent wind seeds ) for each mean wind speed ~~and the mean wind speed is~~ ranging from 4 $m/s$ to 26 $m/s$ in the interval of 2 $m/s$, which results in total 216 simulations. Each simulation uses 10-minute normal wind turbulence inflow with a sampling frequency of 50 Hz. From these 216 simulations, the inputs (~~$\omega_r, \omega_g$, and $T_g$~~ $\boldsymbol{\omega}_r$, $\boldsymbol{\omega}_g$, and $\boldsymbol{T}_g$) for the proposed methods are obtained and by following the procedure depicted in Fig. 2, the torsional loads are obtained and the same is compared with the simulation results. From the simulated ~~$\omega_r$ and $\omega_g$~~ $\boldsymbol{\omega}_r$ and $\boldsymbol{\omega}_g$, the torsional velocity ~~$\dot{\theta}$~~ $\dot{\boldsymbol{\theta}}$ is obtained using Eq. (3) and then the corresponding ~~$\theta$~~ $\boldsymbol{\theta}$ for each simulation is obtained by using the Tikhonov regularisation ~~.~~ (Eq. 4). By assuming that the same level of ~~noise~~ numerical noise (noise due to lack of initial conditions) presents in all the simulations of a particular mean wind speed, the optimal $\lambda$ parameter for regularization (refer Appendix A. 1 ) is estimated only once per mean wind speed. The estimated regularisation parameter ($\lambda$) for each mean wind speed is presented in Fig. 5. As seen in the figure, a higher value of $\lambda$ at higher wind speeds indicates that more numerical damping is needed to suppress the numerical noise.

~~At this point, it is important to realize that the static~~ Since the displacement is reconstructed from the velocity- a dynamic quantity, the reconstructed displacement will have zero mean and this displacement component is referred to as a dynamic component of the displacement ~~does not have any effect on the torsional velocity and hence only the~~ ($\boldsymbol{\theta}_{dyn}$). However, the torsional displacement will have a contribution from the static load and the displacement due to the static load is referred to

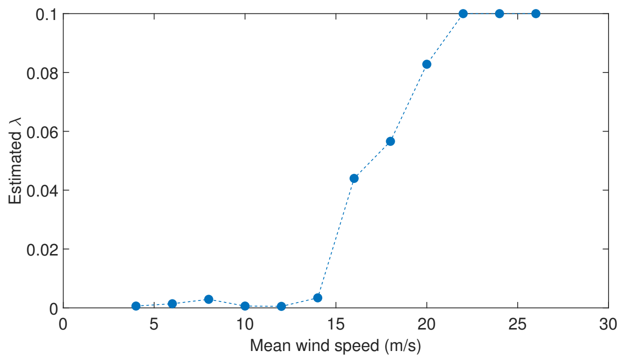
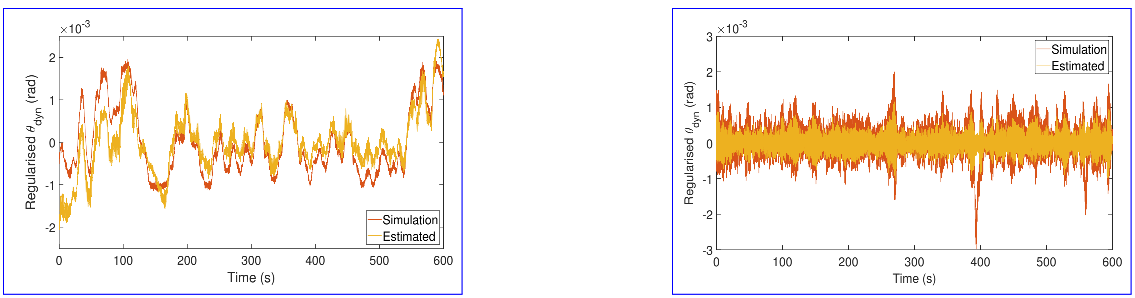



**Figure 5.** Estimated $\lambda$ for each mean wind ~~speeds~~speed.




(a) WSP $= 4\ m/s$

(b) WSP $= 16\ m/s$

**Figure 6.** Comparison regularized $\cancel{\theta_{dyn}}$ $\boldsymbol{\theta}_{dyn}$ with the results from aeroelastic simulation.

as a static component of the displacement ($\boldsymbol{\theta}_{stat}$) (*i.e.*, the mean value of the displacement). The dynamic component of the ~~torsional displacement $\theta_{dyn}$ can be reconstructed from $\dot{\theta}$ using the Tikhonov regularisation~~displacement oscillates about the static component of the displacement. The regularized $\cancel{\theta_{dyn}}$ $\boldsymbol{\theta}_{dyn}$ for two representative mean wind speeds are compared with the dynamic component of simulated torsional displacement in Fig. 6. By removing the static displacement from the simulated displacement, the resulting dynamic component can be compared with the regularised $\cancel{\theta_{dyn}}$ $\boldsymbol{\theta}_{dyn}$ as shown in Fig. 6. As seen in the figure, the regularised $\cancel{\theta_{dyn}}$ $\boldsymbol{\theta}_{dyn}$ matches well with the simulated dynamic displacement for most of the time except for the peak amplitudes.

~~After obtaining $\theta_{dyn}$,~~ Upon ensuring the correctness of $\boldsymbol{\theta}_{dyn}$, Collage method can be employed for $K$ estimation using the results of all the 216 simulations. Since $\boldsymbol{\theta}_{dyn}$ only available, Eq. (~~2) is solved inversely using the collage method for~~ 5) is modified as to have the dynamic components alone for the estimation of $K$ as,

$$\Delta^2 = \int_{t_0}^{t_n} \left[ J_g[\boldsymbol{\omega}_g - \boldsymbol{\omega}_{g0}] + \int_{t_0}^{t_n} (\boldsymbol{T}_g - \boldsymbol{T}_{g_{mean}})\,\mathrm{d}t - K \int_{t_0}^{t_n} \boldsymbol{\theta}_{dyn}\mathrm{d}t - C \int_{t_0}^{t_n} \dot{\boldsymbol{\theta}}\mathrm{d}t \right]^2 \mathrm{d}t. \quad (6)$$

By minimising Eq. (6), $K$ will be obtained for all the 216 simulations. Owing to uncertainty associated with ~~$\theta_{dyn}$~~ $\boldsymbol{\theta}_{dyn}$ estimation along with the ~~collage~~ Collage method, three skewed ~~(outliers)~~ distributions for $K$ can be determined at different mean wind speeds. For ~~the distributionswith outliers, mode is the better~~ these distributions, the mode is a more stable estimate of the central tendency as it is least biased by ~~the~~ outliers (Hedges and Shah, 2003). ~~Ideally, there exists only one torsional $K$ value for the wind turbine main shaft and hence by taking the~~ The mean of modes of the resulting ~~pdfs some stability can be gained for~~ probability density functions (pdfs) provides the resultant estimate of $K$. The relative error between the estimated $K$ value obtained by taking the mean of modes and the design value of the turbine is 12.06 %~~and the values are not presented here due to confidentiality policy~~. At this point, it is important to remember that ~~the estimated $K$ value is a combination~~ since the inputs are from the HAWC2 aeroelastic simulation which does not account for gearbox dynamics, the estimated stiffness value has the contributions from the rotor and the main shaft only. As a result, the estimated $K$ is the resultant torsional stiffness of the main shaft ~~and~~ including the blade edgewise stiffness contribution.

~~In order to obtain the site-specific torsional load,~~ Even though the dynamic component of the displacement is sufficient enough for $K$ estimation as explained above, the mean value (i.e., the static component) of the displacement has a significant effect on the fatigue damage (Veldkamp, 2006). Hence, it is important to estimate the ~~static displacement ($\theta_{stat}$), since the torsional load is given by $M_z = K\theta = K(\theta_{stat} + \theta_{dyn})$. This is achieved by~~ same and this is obtained by solving the static problem of the drivetrain. By considering the static equilibrium of Eq. (1) with all parameters expressed on the low-speed side as follows,

$$K\theta\boldsymbol{\theta}_{stat} = T\boldsymbol{T}_{r_{mean}} = \frac{T_{g_{mean}}}{\eta_{gen}},\frac{\boldsymbol{T}_{g_{mean}}}{\eta_{gen}}. \tag{7}$$

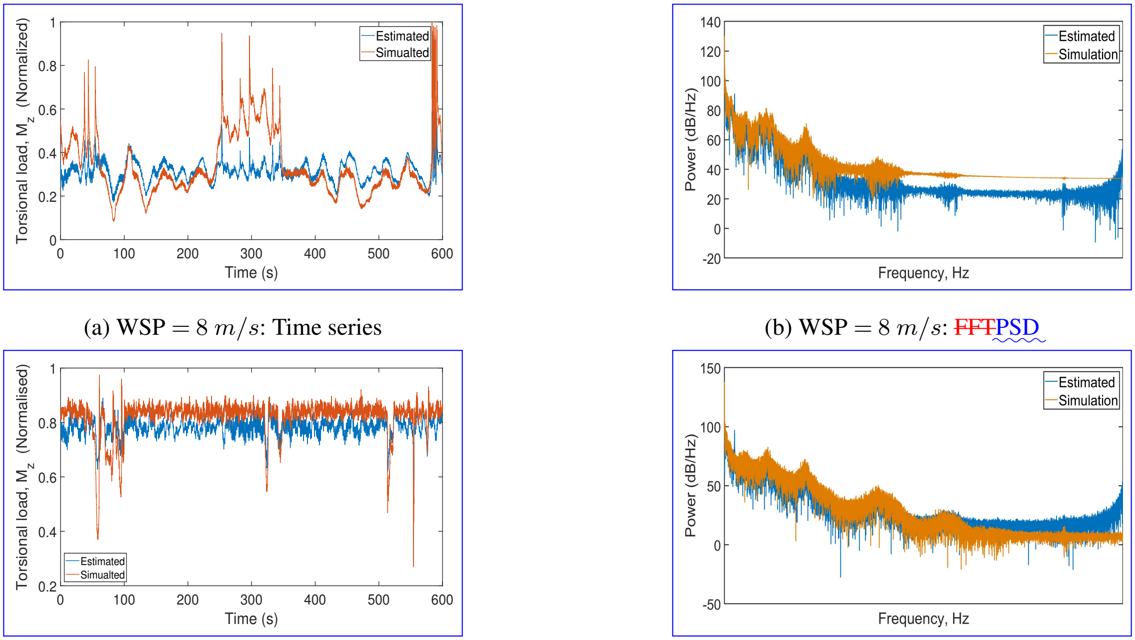
~~where~~Here, $\eta_{gen}$ is the generator efficiency, which is 94.4 % for the V52 turbine~~and from~~. As an alternative way, one can use the overall efficiency of the wind turbine as well for the sake of completeness. However, the use of different efficiencies will not significantly affect the outcome. From Eq. (7), ~~$\theta_{stat}$~~ $\boldsymbol{\theta}_{stat}$ is obtained as,

$$\theta\boldsymbol{\theta}_{stat} = \frac{T_{g_{mean}}}{K\eta_{gen}}\frac{\boldsymbol{T}_{g_{mean}}}{K\eta_{gen}}. \tag{8}$$

~~Hence, by using $\theta_{stat}$ and $\theta_{dyn}$, the total torsional load is obtained as ,~~ Finally, the static and dynamic components are superimposed to get the actual displacement ($\boldsymbol{\theta}$). This actual displacement will then be used to estimate the site-specific main shaft torsional load as shown by Eq. (9).

$$M\boldsymbol{M}_z = K\boldsymbol{\theta} = K(\theta\boldsymbol{\theta}_{stat} + \theta\boldsymbol{\theta}_{dyn}), \tag{9}$$

At this stage, the estimated torsional loads can be compared with the simulated torsional loads. The comparison of the torsional loads for two representative mean wind speeds are shown in Fig. (7). ~~In the figure, confidential values such as load amplitude and frequency are not presented.~~ As seen in ~~the figure~~Fig. (7), all the important aspects of the time-series variation and the dominant frequency dynamics (low-frequency components - ~~upto~~ up to first three peaks ) are captured quite well in the estimated torsional load. The computational time for identifying the regularisation parameter for each mean wind speed




(a) WSP $= 8\ m/s$: Time series

(b) WSP $= 8\ m/s$: ~~FFT~~PSD

(c) WSP $= 14\ m/s$: Time series

(d) WSP $= 14\ m/s$: ~~FFT~~PSD

**Figure 7.** Comparison reconstructed time series and ~~frequency spectra~~ power spectral density (PSD) of the torsional load for two different mean wind speeds.

is about 40 minutes in real-time and the stiffness estimation for the 10-minute simulation takes 14 s in real-time. The above computations are performed on a single node of the high-performance computing cluster of DTU. The cluster has 320 nodes in total, each node consists of two Intel Xeon E5-2680v2 processors, and each processor consists of 10 cores running at 2.8 GHz. If the wind farm is of same type of turbine, then it is sufficient to estimate the stiffness for one of the turbines. However, the shaft torsional displacement needs to be estimated for all the turbines.

Upon estimating the torsional load, the torsional damage equivalent load (DEL) at each mean wind speed is calculated using the following equation:

$$\mathrm{DEL} = \left( \frac{1}{N_{ref}} \sum_{i=1}^{N_{sim}} \left( \frac{1}{N_{sim}} \right) \sum_{k=1}^{k_n} \frac{N_{i,k} S_{i,k}^m(0)}{T_{sim}} \right)^{\frac{1}{m}}, \tag{10}$$

where, $T_{sim}$ is the duration of the load case, $N_{sim}$ is the number of simulations at each mean wind speed, $k_n$ is the total number of load cycles in a given time series, $N_{i,k}$ are the number of cycles at load amplitude $S_{i,k}(\sigma)$ determined with rain flow counting and $m$ is the Wöhler exponent. The zero mean load amplitude is obtained as (Veldkamp, 2006), $S_{i,k}(0) = S_{i,k}(\sigma) + MS_m$, where, $S_m$ is the mean load and $M$ is the mean stress sensitivity. The turbine shaft is made up of cast iron and Veldkamp (2006) has reported that for such material, the mean stress correction factor is $M = 0.19$. ~~For the 1 $Hz$ torsional DEL of a ten-minute time series, $N_{ref}$ becomes 600 cycles and~~ The Wöhler exponent for the cast iron of $m = 6$ ~~is used for the present computation. Using~~ (Veldkamp, 2006) is used in the present study. When $N_{ref}$ in Eq. (10) ~~, the~~ equals the component's lifetime

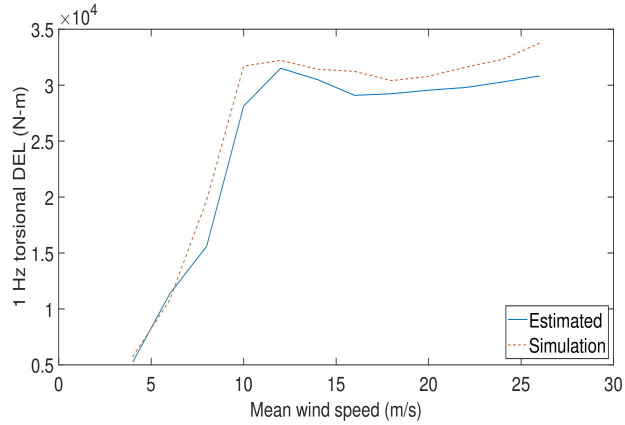



**Figure 8.** Comparison of the predicted DEL with the DEL computed from aeroelastic simulations over all mean wind speeds.

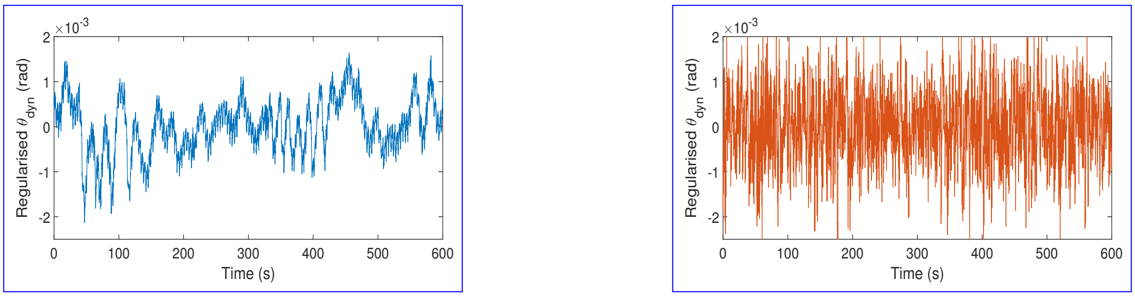
in seconds, then the DEL has the frequency of 1 Hz. The computed torsional DEL using Eq. (10) for DLC 1.2 is compared with the simulated DEL and is shown in Fig. (8). For most of the ~~means~~ mean wind speeds, the estimated DEL is in good agreement with the simulated DEL and the absolute error between these two at each mean wind speed ranges from 4% to 12 %. Higher error at higher mean wind speeds is due to the fact that the peak amplitudes in $\cancel{\theta_{dyn}}$ $\boldsymbol{\theta}_{dyn}$ are not captured well using the Tikhonov regularisation. ~~This~~ The fact that the peak amplitudes are not captured in the reconstructed torsional displacements is typical for Tikhonov regularisation as there is a slight mismatch in the frequency spectra between the actual and reconstructed torsional oscillations (Lee et al., 2010). Another reason could be due to the fact that the pitch ~~control will be active~~ angle influences the main shaft torsional oscillation through the rotor torque and the rotor speed. However, the controller is maintaining a constant rotational speed beyond the rated wind speed~~and~~ , hence the instantaneous changes in the pitch angle ~~affects~~ are not propagated through the rotor speed ~~.~~ to the main shaft oscillations. All these effects, in addition to the uncertainty in $K$ estimation, may resulted in a maximum error of 12 % in the estimated DEL.

## 4 Application on V52 turbine measurements

~~Now, using the verified methodology, the torsional loads for the drive train main shaft~~ Upon verifying the proposed method, the drivetrain main shaft torsional loads are estimated from the SCADA measurements ~~. For this purpose, SCADA measurements of~~ without any need for the aeroelastic model. SCADA measurements taken during January 2019 for the Vestas V52-850 kW research turbine installed at the DTU Risø site ~~will be utilized . In particular, measurements taken for the period of January 2019 consisting of~~ is utilized for this purpose. The measurement data consists of 4459 ten minute recorded cases ~~are used for this study. Since the interest is on normal operations of the turbine~~with 50 Hz sampling frequency. Generator torque is used as one of the SCADA signal instead of the generator speed for this part of the study and the generator speed is obtained from the generator power and generator torque (on the low-speed side) SCADA signals as $\boldsymbol{\omega}_g = P/\boldsymbol{T}_g$ . Further, the measurement data corresponds to normal operation is filtered based on the conditions given in Table 2 that results in 627 cases. It is important to

| | |
|---|---|
| Minimum rotor speed | 1.7 ~~$(rad/s)$~~ $rad/s$ |
| Minimum power | 100 ~~W~~ W |
| ~~Mean~~ Minimum wind speed | 4 ~~$(m/s)$~~ $m/s$ |
| Wind direction | [280°-320°] |

**Table 2.** Normal operation filter conditions.




(a) WSP $= 6\ m/s$

(b) WSP $= 18 m/s$

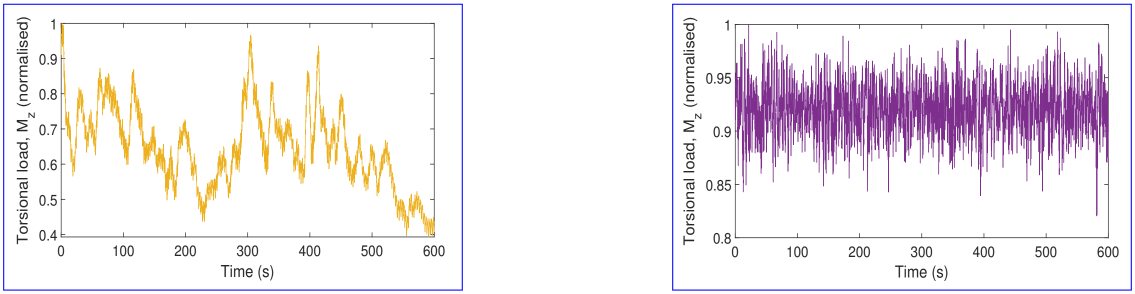
**Figure 9.** ~~Regualrized $\theta_{dyn}$~~ Regularised $\boldsymbol{\theta}_{dyn}$ for two mean wind speeds.

note that based on the site sector assessment, the V52 turbine is in wake free condition between 280° and 320° wind directions.

~~It is important to note that based on the site sector assessment, the V52 turbine is in wake free condition between 280° and 320° wind directions. Further, speed sensors of the rotor measures the rotational speed with a low sample rate which results in a piecewise constant signal. This type of signal has to be converted as a differentiable signal which is achieved by increasing the sample rate followed by a spline interpolation.~~

~~Now, the filtered SCADA measurements are utilized to obtain the torsional loads using the proposed methodology. Using the measured~~ Using the rotor speed and generator speed, ~~$\theta_{dyn}$~~ $\boldsymbol{\theta}_{dyn}$ is calculated using the Tikhonov regularisation method ~~. The obtained $\theta_{dyn}$~~ and the obtained $\boldsymbol{\theta}_{dyn}$ for two representative mean wind speeds are shown in Fig. 9.

~~Similar to the V52 simulations, the~~ The regularisation parameter is obtained for each ~~wind speed measurements~~ mean wind speed measurement using the L-curve criterion and the obtained values not presented here for the sake of brevity. Subsequently, by applying the ~~collage~~ Collage method on Eq. (~~2~~6) for all the 627 cases, the $K$ values are estimated for each load case and then the resultant $K$ value is obtained by taking the mean of modes of the resulting pdf as described in the previous section. ~~The relative error between the estimated and design~~ At this point, it is important to remember that the inputs are from measurement, hence the estimated $K$ ~~values is 3.6 %.~~ is a collective stiffness that has a contribution from all the drivetrain components including the gearbox. With the estimated $K$ value, ~~$\theta_{stat}$~~ $\boldsymbol{\theta}_{stat}$ is calculated for each load case using Eq. (8) and then the torsional load is obtained from Eq. (9). The calculated torsional loads for two representative wind speeds are shown in Figs. (10a, 10b). After estimating the torsional loads for all the 627 cases, the identified loads are grouped according to the mean




(a) WSP $= 8\ m/s$: Time series

(b) WSP $= 14\ m/s$: Time series

**Figure 10.** Identified torsional loads ($M_z$) for two mean wind speeds.

wind speeds range from 6 $m/s$ to 22 $m/s$ which are subdivided into 9 wind speed bins of 2 $m/s$ width each. Subsequently, the ~~1 $Hz$~~ torsional DEL is calculated for each mean wind speed using Eq. (10) and the same is shown in Fig. (11). It is important to note that the DEL given in Fig. 8 (~~Simulation~~simulation) being the design load and the DEL given in Fig. 11 being the site-specific load, the difference between these two ~~will~~ can give an estimate about the remaining torsional fatigue life of the main shaft under normal operating conditions (Ziegler et al., 2018). ~~Decisions regarding RUL of the drivetrain may be taken considering the~~ Further, the estimated main shaft torsional DEL ~~as an input.~~ may be used to quantify the RUL of the drivetrain (Pagitsch et al., 2020). However, this is beyond the scope of the current work. It is also feasible to reconstruct the torsional ~~moment~~ load time series, which may be used an input to gearbox design tools to predict the loading within the gearbox.

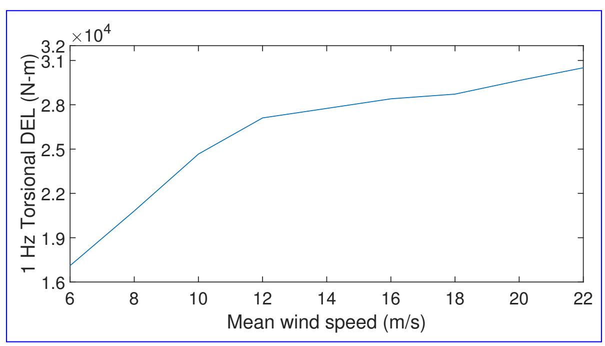



**Figure 11.** Estimated torsional DEL from SCADA measurements of Vestas V52 turbine.

## 5 Summary

A novel inverse problem-based approach is developed for estimating the main shaft torsional displacement and stiffness by using the high-frequency SCADA measurements. A mathematical model describing the coupled shaft torsional dynamics is used for this purpose. Numerical errors and the effect of measurement noise on the torsional displacement reconstruction are minimized through the Tikhonov regularization technique. Subsequently, the ~~collage~~ Collage method is used to estimate the main shaft coupled torsional stiffness. The estimated main shaft quantities are then used to identify the main shaft site-specific torsional load. The proposed formulation is successfully verified for the main shaft torsional loads with the aeroelastic simulation of the Vestas V52 turbine. Upon verification, the methodology is extended to identify the site-specific main shaft torsional loads of the same turbine by using SCADA measurements. For this purpose, the measurement data from the DTU Risø site is utilized and the measurement data is filtered and calibrated for the turbine normal operation. Using the identified torsional loads, the torsional DEL is obtained. Depending on the prior information about the stiffness value, one can either use the entire proposed methodology or follow the torsional displacement estimation part of the proposed methodology for the torsional load identification. Since the site-specific SCADA measurements are used in the analysis, the obtained loads can give a better estimate of the remaining fatigue life ~~which can help in the life extension decision-making process. Besides monitoring~~ of drivetrain components. Monitoring the estimated loads can help in inspection planning and scheduling maintenance activities. As the proposed methodology does not require any design basis or an aeroelastic design basis, it can be used for ~~any older turbines~~ wind turbines that possess high-frequency SCADA measurements for the estimation of the main shaft torsional ~~RUL~~load and DEL.

## Appendix

### A.1 Displacement Reconstruction using the Tikhonov Regularisation

By definition, the velocity $\dot{\boldsymbol{\theta}}$ is expressed as,

$$\dot{\boldsymbol{\theta}}(t) = \frac{\mathrm{d}\boldsymbol{\theta}}{\mathrm{d}t} \approx \dot{\hat{\boldsymbol{\theta}}}(t), \tag{11}$$

where, $\dot{\hat{\boldsymbol{\theta}}}(t)$ is the velocity obtained from Eq. (3) which can be considered as a measured velocity. As explained in Section 2.1, the lack of initial conditions in addition to the measurement noise leads to erroneous displacement. In order to minimise the error between the actual and measured velocities in the least square sense, following minimisation problem has to be solved,

$$\text{Min} \;\; \Pi_E(\boldsymbol{\theta}) = \frac{1}{2}\int_{t_0}^{t_n} (\dot{\boldsymbol{\theta}}(t) - \dot{\hat{\boldsymbol{\theta}}}(t))^2 \, \mathrm{d}t. \tag{12}$$

Here, $\dot{\boldsymbol{\theta}}$ is the calculated velocity. In other words, Eq. (12) gives a measure of how well the actual velocity approximates the measured velocity. By means of the trapezoidal rule, Eq. (12) is discretised as follows (Hong et al., 2008),

$$\Pi_E(\boldsymbol{\theta}) \approx \|\mathbf{L}_a(\dot{\boldsymbol{\theta}} - \dot{\hat{\boldsymbol{\theta}}})\|_2^2 \Delta t, \tag{13}$$

where $\Delta t$ is the time interval of the discretization and $\mathbf{L}_a$ is the diagonal weighing matrix of order $(N+1)$ as,

$$\mathbf{L}_a = \begin{bmatrix} 1/\sqrt{2} & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1/\sqrt{2} \end{bmatrix}. \tag{14}$$

Here, $N$ is the number of data points in $\boldsymbol{\theta}$ and for a 10-minute simulation with 50 Hz sampling frequency $N$ becomes 30000. Further, the calculated velocity $\dot{\boldsymbol{\theta}}$ is discretised by the central difference rule and written in matrix form as,

$$\frac{1}{\Delta t}\mathbf{L}_c\boldsymbol{\theta} = \dot{\boldsymbol{\theta}}, \tag{15}$$

where, the central difference matrix $\mathbf{L}_c$ of size $(N+1)\times(N+3)$ and the displacement vector $\boldsymbol{\theta}$ of size $(N+3)$ are given as,

$$\mathbf{L}_c = \begin{bmatrix} 1 & 0 & 1 & \dots & \\ & -1 & 0 & 1 & \\ & & \ddots & & \\ & -1 & 0 & 1 & \\ & \dots & -1 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_{-1} \\ \boldsymbol{\theta}_0 \\ \vdots \\ \boldsymbol{\theta}_n \\ \boldsymbol{\theta}_{n+1} \end{pmatrix}. \tag{16}$$

In the finite difference discretization of Eq. (16), some defined nodes are located outside of the domain considered (*i.e.*, domain here is time, $t \in [t_0, t_n]$ satisfying $t_0 \leq t \leq t_n$). These nodes defined by time steps $i = -1$ and $i = n+1$ are fictitious. These nodes are the dummy nodes that are used in solving the differential equations by the finite difference method (Lapidus and Pinder, 2011) . Substitution of Eq. (15) into Eq. (13) leads to,

$$\text{Min} \;\; \Pi_E(\boldsymbol{\theta}) \approx \frac{1}{2}\left\| \frac{1}{2\Delta t}\mathbf{L}_a\mathbf{L}_c\boldsymbol{\theta} - \mathbf{L}_a\dot{\tilde{\boldsymbol{\theta}}} \right\|_2^2 \Delta t = \frac{1}{2}\left\| \frac{1}{2}\mathbf{L}\boldsymbol{\theta} - \mathbf{L}_a\dot{\tilde{\boldsymbol{\theta}}}\Delta t \right\|_2^2 \frac{1}{\Delta t}, \tag{17}$$

where, $\mathbf{L} = \mathbf{L}_a\mathbf{L}_c$. This minimisation problem is regularised for solution boundedness with a parameter $\lambda$, and given as,

$$\text{Min} \;\; \Pi_E(\boldsymbol{\theta}) \approx \frac{1}{2}\left\| \frac{1}{2}\mathbf{L}\boldsymbol{\theta} - \mathbf{L}_a\dot{\tilde{\boldsymbol{\theta}}}\Delta t \right\|_2^2 + \frac{\lambda^2}{2}\|\boldsymbol{\theta}\|_2^2. \tag{18}$$

The above minimisation problem is known as the Tikhonov regularisation and $\lambda$ is referred to as the regularisation parameter. Minimising Eq. (18) as,

$$\frac{\mathrm{d}\Pi_E}{\mathrm{d}\boldsymbol{\theta}} = \frac{1}{2}\left( \frac{\mathbf{L}^2\boldsymbol{\theta}}{2} - \mathbf{L}\mathbf{L}_a\dot{\tilde{\boldsymbol{\theta}}}\Delta t \right) + \lambda^2\boldsymbol{\theta} = 0, \tag{19}$$

yields the following quadratic equation in $\boldsymbol{\theta}$,

$$\boldsymbol{\theta} = \left( \frac{\mathbf{L}^2}{4} + \lambda^2\mathbf{I} \right)^{-1} \frac{\mathbf{L}\mathbf{L}_a\dot{\tilde{\boldsymbol{\theta}}}\Delta t}{2}, \tag{20}$$

where, $\mathbf{I}$ is the identity matrix of order $(N+3)$.

The choice of regularisation parameter $(\lambda)$ plays a crucial role in getting an optimal fit for the solution. Based on the knowledge about measurement errors, Hansen (2005) proposed two classes for the estimation of $\lambda$:

- methods based on knowledge of measurement errors
- methods that do not require details about measurement errors.

In the present scenario, the information regarding the measurement error is unknown, hence class two is used for the current study. In class two, there are three widely used methods (Nieminen et al., 2011): (i) quasi optimality criterion, (ii) Generalized gross validation (GCV), and (iii) L-curve method. Compared to the GCV method, the other two methods give a better estimate of $\lambda$ (Gao et al., 2016). Further, for larger problems, the quasi optimality method is computationally more expensive than the L-curve method. Owing to this fact, the L-curve method is used here for estimating $\lambda$. In L-curve method, the optimal $\lambda$ is the one which gives the maximum curvature in the L-curve between norm of the regularized solution $\alpha(\lambda) = \|\boldsymbol{\theta}_{reg}\|_2$ and norm of the residual $\beta(\lambda) = \left\| \frac{1}{2}\mathbf{L}\boldsymbol{\theta} - \mathbf{L}_a\dot{\tilde{\boldsymbol{\theta}}}\Delta t \right\|_2$ and the curvature of the L-curve is given by (Nieminen et al., 2011),

$$\kappa(\lambda) = \frac{\ddot{\alpha}\dot{\beta} - \ddot{\beta}\dot{\alpha}}{[\dot{\alpha}^2 + \dot{\beta}^2]^{3/2}}. \tag{21}$$

Substituting the optimal $\lambda$ obtained by finding the maximum curvature of Eq. (21) and $\dot{\tilde{\boldsymbol{\theta}}}(t)$ obtained from Eq. (3) in Eq. (20), the regularised torsional displacement $(\boldsymbol{\theta})$ is obtained.

## A.2 Modal Parameters Estimation using the Collage Method

For a given initial value problem (IVP),

$$x(t) = f(x,t), x(0) = x_0, \tag{22}$$

that admits a target solution $x(t)$, the associated Picard integral operator T is given by,

$$(\mathrm{T}u)(t) = x_0 + \int_{t_0}^{t_n} f(u(s),s)\mathrm{d}s. \tag{23}$$

The assumptions regarding the parameter estimation problem using the Collage method are listed as follow (Deng and Liao, 2009) :

1. $x(t) \in [t_0, t_n]$ is a bounded solution; where, $t_0$ and $t_n$ are positive constants satisfying $t_0 < t_n$.
2. $f(u,x,t,\gamma_1,\cdots,\gamma_m)$ is continuous, where, $\gamma_i, i = 1,\cdots,m$ are the unknown modal parameters.
3. The exact solution $x(t)$ of the system (22) exists uniquely.

Here, the unique solution of the considered IVP is given by the fixed point $\bar{u}(t)$ of this Picard operator (Kunze et al., 2004) . Accordingly, the Collage distance becomes, $(x - \mathrm{T}x)$ and then the optimal solution is the one which minimizes the squared $\mathcal{L}^2$ Collage distance (Here, $\mathcal{L}^2$ is 2-norm or square norm of a function). Also, unlike the conventional inverse problem which minimises the approximate error $\mathrm{d}(x-\bar{x})$, the Collage method minimizes the Collage distance $\mathrm{d}(x,\mathrm{T}x)$ which is an useful change as one cannot find $\bar{x}$ for a general T (Kunze et al., 2004). Further, the optimality of the Collage distance minimization is ensured as shown by Kunze et al. (2004). Accordingly, the $\mathcal{L}^2$ Collage distance has the form,

$$\Delta = \left(\int_{t_0}^{t_n} (x(t) - (\mathrm{T}x)(t))^2 \mathrm{d}t\right)^{\frac{1}{2}}. \tag{24}$$

Minimising the squared $\mathcal{L}^2$ Collage distance yields a stationarity condition, $\frac{\mathrm{d}\Delta^2}{\mathrm{d}\gamma_m} = 0$, that results in a set of simultaneous linear equations as a function of unknown modal parameters $(\gamma_m)$. The modal parameters are then obtained by solving those set of linear equations.

*Author contributions.* **W. Dheelibun Remigius**: Methodology, Formal analysis, Investigation, Validation, Writing - original draft. **Anand Natarajan**: Original idea, developing the scientific methods, Writing - review and editing, Supervision.

*Competing interests.* 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.

*Acknowledgements.* This work has been ~~fully~~ funded by the Energy Technology Development and Demonstration Programme, Denmark (EUDP) [Grant No. 64017 -05114]. Investigations have been carried out at Department of Wind Energy, Technical University of Denmark.

## References


Aarden, P.: System Identification of the 2B6 Wind Turbine: A regularised PBSIDopt approach, Master's thesis, Delft University of Technology, 2017.

Andersen, P.: Identification of civil engineering structures using vector ARMA models, Ph.D. thesis, Department of Building Technology and Structural Engineering, Aalborg University, 1997.

Bak, C., Zahle, F., Bitsche, R., Kim, T., Yde, A., Henriksen, L. C., Natarajan, A., and Hansen, M.: Description of the DTU 10 MW reference wind turbine, Tech. Rep. I-0092, DTU Wind Energy, 2013.

Bangji, Z., Shouyu, Z., Qingxi, X., and Nong, Z.: Load identification of virtual iteration based on Tikhonov regularization and model reduction, Hong Kong Journal of Social Sciences, 44, 53 – 59, 2017.

Berglind, J. J. B., Wisniewski, R., and Soltani, M.: Fatigue load modeling and control for wind turbines based on hysteresis operators, in: 2015 American Control Conference (ACC), pp. 3721–3727, IEEE, 2015.

Boukhezzar, B., Lupu, L., Siguerdidjane, H., and Hand, M.: Multivariable control strategy for variable speed, variable pitch wind turbines, Renewable Energy, 32, 1273–1287, 2007.

Brandt, A. and Brincker, R.: Integrating time signals in frequency domain–Comparison with time domain integration, Measurement, 58, 511–519, 2014.

Brincker, R., Ventura, C., and Andersen, P.: Damping estimation by frequency domain decomposition, in: Proceedings of the 19th International Modal Analysis Conference (IMAC), pp. 5–8, Orlando, FL, USA, 2001.

Deng, X. and Liao, Q.: Parameter estimation for partial differential equations by collage-based numerical approximation, Mathematical Problems in Engineering, 2009.

Deng, X., Wang, B., and Long, G.: The Picard contraction mapping method for the parameter inversion of reaction-diffusion systems, Computers & Mathematics with Applications, 56, 2347–2355, 2008.

Ding, F., Tian, Z., Zhao, F., and Xu, H.: An integrated approach for wind turbine gearbox fatigue life prediction considering instantaneously varying load conditions, Renewable energy, 129, 260–270, 2018.

Dong, W., Xing, Y., and Moan, T.: Time domain modeling and analysis of dynamic gear contact force in a wind turbine gearbox with respect to fatigue assessment, Energies, 5, 4350–4371, 2012.

Gallego-Calderon, J. and Natarajan, A.: Assessment of wind turbine drive-train fatigue loads under torsional excitation, Engineering Structures, 103, 189–202, 2015.

Gao, W., Yu, K., and Wu, Y.: A new method for optimal regularization parameter determination in the inverse problem of load identification, Shock and Vibration, pp. 7328 969–1–16, 2016.

Girsang, I. P., Dhupia, J. S., Muljadi, E., Singh, M., and Pao, L. Y.: Gearbox and drivetrain models to study dynamic effects of modern wind turbines, in: 2013 IEEE Energy Conversion Congress and Exposition, pp. 874–881, https://doi.org/10.1109/ECCE.2013.6646795, 2013.

Groetsch, C. W.: Inverse problems in the mathematical sciences, vol. 52, Springer, 1993.

Hansen, M. H., Thomsen, K., Natarajan, A., and Barlas, A.: Design load basis for onshore turbines-revision 00, Tech. rep., DTU Wind Energy, 2015.

Hansen, P. C.: Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion, vol. 4, SIAM, 2005.

Hào, D. N. and Quyen, T. N. T.: Convergence rates for Tikhonov regularization of a two-coefficient identification problem in an elliptic boundary value problem, Numerische Mathematik, 120, 45–77, 2012.

Hedges, S. B. and Shah, P.: Comparison of mode estimation methods and application in molecular clock analysis, BMC Bioinformatics, 4, 31–1–11, 2003.

Hong, Y. H., Park, H. W., and Lee, H. S.: A regularization scheme for displacement reconstruction using acceleration data measured from structures, in: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2008, vol. 6932, p. 693228, International Society for Optics and Photonics, 2008.

Ibrahim, S. R.: Random decrement technique for modal identification of structures, Journal of Spacecraft and Rockets, 14, 696–700, 1977.

IEC 61400-1: Wind energy generation systems –Part 1: Design requirements, Standard, International Electrotechnical Commission, Geneva, CH, 2019.

IEC 61400-13: Wind turbines – Part 13: Measurement of mechanical loads, Standard, International Electrotechnical Commission, Geneva, CH, 2016.

James, G. H., Carne, T. G., and Lauffer, J. P.: The natural excitation technique (NExT) for modal parameter extraction from operating structures, Modal Analysis- The International Journal of Analytical and Experimental Modal Analysis, 10, 260–277, 1995.

Koukoura, C., Natarajan, A., and Vesth, A.: Identification of support structure damping of a full scale offshore wind turbine in normal operation, Renewable Energy, 81, 882–895, 2015.

Kunze, H. and La Torre, D.: Computational Aspects of Solving Inverse Problems for Elliptic PDEs on Perforated Domains Using the Collage Method, in: Mathematical and Computational Approaches in Advancing Modern Science and Engineering, pp. 113–120, Springer, 2016.

Kunze, H., La Torre, D., and Vrscay, E.: A generalized collage method based upon the Lax–Milgram functional for solving boundary value inverse problems, Nonlinear Analysis: Theory, Methods & Applications, 71, e1337–e1343, 2009.

Kunze, H. E. and Vrscay, E. R.: Solving inverse problems for ordinary differential equations using the Picard contraction mapping, Inverse Problems, 15, 745–770, 1999.

Kunze, H. E., Hicken, J. E., and Vrscay, E. R.: Inverse problems for ODEs using contraction maps and suboptimality of the 'collage method', Inverse Problems, 20, 977–991, 2004.

Lapidus, L. and Pinder, G. F.: Numerical solution of partial differential equations in science and engineering, John Wiley & Sons, 2011.

Larsen, T. J. and Hansen, A. M.: How 2 HAWC2, the user's manual, Tech. rep., Risø National Laboratory, Technical University of Denmark, 2007.

Lee, H. S., Hong, Y. H., and Park, H. W.: Design of an FIR filter for the displacement reconstruction using measured acceleration in low-frequency dominant structures, International Journal for Numerical Methods in Engineering, 82, 403–434, 2010.

Nieminen, T., Kangas, J., and Kettunen, L.: Use of Tikhonov regularization to improve the accuracy of position estimates in inertial navigation, International Journal of Navigation and Observation, pp. 450 269–1–10, 2011.

Norén-Cosgriff, K. and Kaynia, A. M.: Estimation of natural frequencies and damping using dynamic field data from an offshore wind turbine, Marine Structures, 76, 102 915, 2021.

Pagitsch, M., Jacobs, G., and Bosse, D.: Remaining Useful Life Determination for Wind Turbines, Journal of Physics: Conference Series, 1452, 012 052, 2020.

Pahn, T.: Inverse load calculation for offshore wind turbines, Ph.D. thesis, Hannover: Gottfried Wilhelm Leibniz Universität Hannover, 2013.

Pahn, T., Rolfes, R., and Jonkman, J.: Inverse load calculation procedure for offshore wind turbines and application to a 5-MW wind turbine support structure, Wind Energy, 20, 1171–1186, 2017.

Qihe, L.: Integration of vibration acceleration signal based on labview, Journal of physics: conference series, 1345, 042 067, 2019.

Ronasi, H., Johansson, H., and Larsson, F.: A numerical framework for load identification and regularization with application to rolling disc problem, Computers & Structures, 89, 38–47, 2011.

Tikhonov, A. N.: Solution of incorrectly formulated problems and the regularization method, Soviet Mathematics - Doklady, 4, 1035–1038, 1963.

Van Overschee, P. and De Moor, B.: Subspace algorithms for the stochastic identification problem, Automatica, 29, 649–660, 1993.

Veldkamp, H. F.: Chances in wind energy: A probabilistic approach to wind turbine fatigue design, Ph.D. thesis, Faculty of Mechanical Maritime and Materials Engineering, TU Delft, 2006.

Vestas: Vestas V52, https://en.wind-turbine-models.com/turbines/71-vestas-v52 (08-04-2020).

Wang, T., Zhang, L., and Tamura, Y.: An operational modal analysis method in frequency and spatial domain, Earthquake Engineering and Engineering Vibration, 4, 295–300, 2005.

Wang, T., Celik, O., Catbas, F. N., and Zhang, L. M.: A frequency and spatial domain decomposition method for operational strain modal analysis and its application, Engineering Structures, 114, 104–112, 2016.

Zahid, F. B., Ong, Z. C., and Khoo, S. Y.: A review of operational modal analysis techniques for in-service modal identification, Journal of the Brazilian Society of Mechanical Sciences and Engineering, 42, 1–18, 2020.

Zhang, L., Wang, T., and Tamura, Y.: A frequency–spatial domain decomposition (FSDD) method for operational modal analysis, Mechanical systems and signal processing, 24, 1227–1239, 2010.

Ziegler, L., Gonzalez, E., Rubert, T., Smolka, U., and Melero, J. J.: Lifetime extension of onshore wind turbines: A review covering Germany, Spain, Denmark, and the UK, Renewable and Sustainable Energy Reviews, 82, 1261–1271, 2018.