Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

Experimental validation of virtual torque sensing for wind turbine gearboxes based on strain measurements

Jelle Bosmans*[1,2] | Simone Gallas[1,2] | Victor Smeets[1,2] | Matteo Kirchner[1,2] | Luk Geens[3] | Jan Croes[4] | Wim Desmet[1,2]

[1]Department of Mechanical Engineering, KU Leuven, Belgium
[2]Flanders Make@KU Leuven
[3]ZF Wind Power Antwerpen NV, Lommel, Belgium
[4]Forcebit BV, Kessel-Lo, Belgium

Correspondence
*Jelle Bosmans, Department of Mechanical Engineering, KU Leuven
Email: jelle.bosmans@kuleuven.be

Abstract

In efforts to reduce the operation and maintenance cost of wind turbines, there is an increasing interest to monitor key turbine quantities such as the torque load on the gearbox. Monitoring the torque paves the way for the calculation of remaining useful lifetime, leading to cost reductions through improved reliability and maintenance planning.

In order to avoid expensive direct torque sensors, this paper investigates the potential of virtual torque sensing, a technique based on 3 basic components: firstly, a set of non-intrusive sensors installed on the gearbox. Three groups of strain gauges on the gearbox as well an angular encoder are considered in this paper. Secondly, a physics-based model, capable of predicting the response of aforementioned sensors. These models are constructed with a purposeful balance between accuracy and computational cost. Model validation and updating are performed to ensure efficient and accurate prediction of the sensor output. Finally, an Augmented Extended Kalman Filter (AEKF) is used to combine the measured response with predictions from the model to infer the gearbox input torque. Since a key factor determining the performance of the AEKF is the tuning of the AEKF covariance matrices, multiple methods are introduced to systematically tune the covariance matrices. Experimental validation results show that the virtual torque sensor can detect the load torque with a Normalized Mean Absolute Error (NMAE) between 3.41 % to 7.47 %, depending on the sensor set. The influence of the amount of sensors used and the tuning method are also investigated.

KEYWORDS:
Virtual torque sensing, Gearbox, Extended Kalman Filter, Strain gauges

1 | INTRODUCTION

Over the last decade, the total installed capacity of wind power generation has increased significantly, reaching 769GW and 55.7GW for onshore and offshore wind, respectively[1]. Driven by ambitious policy targets for renewable electricity generation, this share is expected to continue increasing according to recent forecasts[2]. Meanwhile, the Levelized Cost of Electricity (LCoE) of wind power generation has dropped by 68 % for onshore wind and 60 % for offshore wind[1]. This cost reduction has been driven by three main factors. Firstly, the cost of new wind turbine installations has decreased[1]. Secondly, the average size of wind turbines has increased, resulting in greater efficiency, cost savings due to economies of scale,

[0]**Abbreviations:** LCoE, Levelized Cost of Electricity; RUL, Remaining Useful Life; TA, Torque Arm; RG, Ring Gear; HSH, High-Speed Housing; SNR, Signal-to-Noise Ratio; FE, Finite Element; RMSE, Root Mean Square Error; MAE, Mean Absolute Error; NMAE, Normalized Mean Absolute Error; AEKF, Augmented Extended Kalman Filter; FRF, Frequency Response Function; TF, Transfer Function

and higher capacity factors[3]. Finally, improvements in reliability and maintenance have reduced unplanned downtime and associated production losses[4]. The latter trend is the focus of this paper, as further development in these areas is necessary for the continued success of wind power.

A large portion of the installed wind turbines employ a gearbox in the drivetrain to convert the slow rotation speed of the rotor to a higher velocity appropriate for the generator[5,6]. In these turbines, gearbox repairs can have a significant contribution to maintenance costs and downtime[7,6]. Hence, improving gearbox reliability can effectively reduce the LCoE. In this context - and in the broader context of digitalization of industrial products, there is an increasing interest to monitor key gearbox quantities during operation. Through monitoring, condition-based predictive gearbox maintenance is enabled, with the goal of bringing down the amount of maintenance interventions and unplanned downtime[3]. For this purpose, the operational loads applied to the gearbox are among the most valuable quantities to be monitored, since the loads directly influence the consumption of fatigue life of gearbox components. Knowledge of the operational loads enables the calculation of the Remaining Useful Life (RUL) of the gearbox components.

Operational loads can be measured directly with heavy duty sensor solutions[8,9]. However, these sensors are typically reserved for lab settings and prototypes, since they are intrusive and expensive. Virtual load sensing aims to provide a more cost-effective alternative by using indirect sensors which are economically more viable and can be placed in more favourable positions[10]. The indirect measurements are combined with a model in order to infer operational loads. This model-based virtual load sensing approach has been successfully demonstrated in many applications, including civil structures[11,12], for a $5.5\,\mathrm{kW}$ mechatronic drivetrain[13], and for an automotive suspension system[14]. Applications of virtual load sensing in wind turbine drivetrains have been conceptualized in multiple works[15,16,17,18]. However, these studies are purely numerical and do not include experimental validation.

This paper focuses on the experimental validation of a model-based virtual load sensor to detect the loads applied to a wind turbine gearbox, focusing on the operational torque load. A target frequency range for the virtual torque sensor from $0\,\mathrm{Hz}$ to $10\,\mathrm{Hz}$ is considered, which was defined by the gearbox manufacturer. In an earlier contribution by the authors, the proposed methodology has been presented and validated using numerical data[17]. Furthermore, a sensor response sensitivity study, and model validation and updating has been performed based on the experimental data[19]. Since the experimental virtual torque sensing results presented in this paper depend heavily on these preceding research results, the main conclusions of these contributions will be repeated.

This paper is organized as follows: section 2 introduces the drivetrain test setup (including the gearbox) that was used in this research. Measurements are carried out on a lab test setup in order to allow greater control over the applied operating conditions and loads. The following sections introduce the three components that make up the proposed virtual torque sensor:

- The sensors installed on the gearbox are discussed in section 3, distinguishing between indirect sensors which will be used in the virtual torque sensor, and direct torque sensors that are used only for validation and model calibration purposes. This section also contains an analysis of the measurements which validates that the sensor response meets the expectations required for virtual torque sensing[19].
- A physics-based model of the gearbox tailored to virtual sensing is introduced in section 4. Model validation and updating are also included in this section[19].
- The estimation algorithm which infers the unknown torque from the measured response and the model predictions is discussed in section 5. Specific attention is paid to systematically tune the settings of the estimation algorithm in order to achieve comparable results among different measurement sets, allowing for a performance comparison.

Section 6 presents the experimental virtual torque sensing results, which is to the authors knowledge the first demonstration of model-based virtual torque sensing for wind turbine gearboxes. This section brings together all elements discussed in earlier sections, investigating the virtual torque sensing performance using different sets of measurements and with different tuning methods for the estimation algorithm. Finally, section 7 wraps up the paper with conclusions and suggestions for further research.

2 | TEST SETUP, DEVICE UNDER TEST AND TEST CAMPAIGN

For the research presented in this paper, the $4\,\mathrm{MW}$ drivetrain testing infrastructure at the Center for Wind Power Drives (CWD) was used[20]. A picture of the testing infrastructure with the major subsystems is shown in fig. 1. This paper will only briefly touch upon the function and capabilities of each subsystem, for further details the reader is referred to Averous et al[21]. Moving from left to right in fig. 1, we first encounter the drive motor, a $4\,\mathrm{MW}$ direct-drive permanent-magnet machine. This drive motor is able to emulate the high torque at low rotation speed that would be generated by the rotor in a wind turbine. The next subsystem is the non-torque load application unit, in which 6 actuators generate both axial and radial forces, and bending moments. Together with the drive motor, this unit allows the application of realistic operating loads to the gearbox.

The gearbox considered in this research is originally from an $850\,\mathrm{kW}$ Vestas V52 turbine. Next to the original bedplate, main shaft, and main bearings are also used. The generator is replaced by a higher powered model, which combined with the $4\,\mathrm{MW}$ motor of the test infrastructure

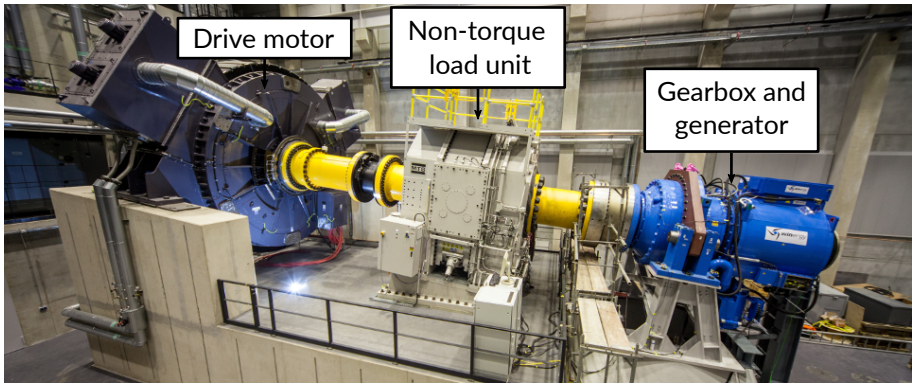


Figure 1 Picture of the 4 MW drivetrain testing infrastructure at CWD indicating the major subsystems. Image source: CWD [20].

allows the gearbox to be run in overloaded conditions. This is not a relevant operating condition for this research, but overload conditions were considered in the context of metrological gearbox robustness evaluation in other research using the same test setup [22]. The gearbox itself is a three-stage gearbox consisting of one planetary stage and two helical stages. The gearbox housing consists of three main components: the Torque Arm (TA) which forms the connection point of gearbox to the nacelle mainframe; the static Ring Gear (RG) of the planetary stage; and the High-Speed Housing (HSH) which supports the bearings for the helical stages.

Several load scenarios were considered during the test campaign, with variations of the operating speed, torque, and non-torque loads. A full overview of the design of experiments and the load scenarios investigated falls outside the scope of this paper, but can be found by Azzam et al [23]. For this paper, three different segments of measured data are considered. Since the focus of this paper is on torque estimation, the non-torque loads are kept at zero for all the selected segments. The segments have the following properties:

1. A segment with slowly increasing and decreasing torque, and constant speed, which is used to investigate the response of the sensors to torque (section 3.4).
2. A segment with constant torque and speed, which is used to validate and update the models (section 4.2).
3. A segment containing a start-stop cycle where the speed and torque are increased rapidly, then held for a period before being decreased again. This segment is the validation case on which the performance of the virtual torque sensor is assessed.

3 | SENSORS

This section discusses the sensors that were installed on the gearbox for this research. A distinction is made between indirect sensors which will be used in the virtual torque sensor, and direct torque sensors that are used only for validation and model calibration purposes. The direct sensors are discussed first, in section 3.1. Before discussing the indirect sensors, it is necessary to define which characteristics of these sensor and their response are desirable in the context of virtual torque sensing [19,10]. This is covered in section 3.2. The indirect sensors themselves are discussed in section 3.3. Finally, section 3.4 will validate some of the indirect sensor requirements based on a sensor response sensitivity study.

3.1 | Direct torque sensors

Direct torque transducers are mounted on the gearbox input and output shafts. These sensors are exactly the intrusive sensors that should be avoided, and as such they are only included for validation purposes (i.e. they are not directly used in the virtual torque sensing algorithm). However, they are used to analyze the torque-response relationship for the strain gauges (section 3.4), for model validation and updating (section 4.2), and as a reference signal to quantify the accuracy of the virtual torque sensing results (section 6).

3.2 | Requirements for indirect sensors

Firstly, in order achieve the envisioned advantages vis-a-vis direct torque measurements, the sensors needs to be economic, reliable, and non-intrusive. The latter is defined here as neither interfering with the function of critical components in the gearbox, nor requiring significant design

Table 1 Overview of all sensors used in the test campaign.

Sensor purpose	Sensor type	Measured quantity	Location	No. of channels
Validation	Torque flange	Torque	Main Shaft	1
			High-Speed Shaft	1
Virtual sensing	Delta rosette strain gauges	Strain	Torque Arm	17
			Ring Gear	55
			High-speed housing	29
	Incremental encoder	Angular velocity	Main Shaft	1
			High-Speed Shaft	1

changes or additional equipment in order to accommodate for the sensor. This implies that non-rotating sensors which are accessible (and thus serviceable) on the exterior surface of the gearbox housing are preferred. Sensors placed inside the gearbox will have to adhere to much stricter demands on sealing and mechanical integrity, making them less suitable. The aspects of economic feasibility and reliability for industrial deployment are not discussed in detail in this paper, since they strongly depend on aspects that are outside of the research scope (e.g. economies of scale, logistics, installation procedures, and quality control).

Secondly, the sensor response should be torque-driven with a sufficiently high Signal-to-Noise Ratio (SNR). If not, the response will not contain sufficient information to infer the applied torque. Thirdly, this torque-driven response should be repeatable, with a certain torque level and gearbox state ideally leading to the same sensor output. If this is not the case, for example due to the presence of some additional loads not taken into account, the accuracy of the virtual torque sensor will suffer.

Finally, since a model-based approach is used, it is also crucial that the chosen model can reliably predict the sensor output. This implies that the choice of modelling formalism and the selection of indirect sensor types are coupled. For example, the use of strain gauges on the gearbox housing necessitates the use of a Finite Element (FE) model for the gearbox housing. This criterion is also linked to the previously mentioned repeatability: in case the response of a certain sensor is not repeatable due to the presence of an additional load, it can be worthwhile to include said load in the model.

3.3 | Indirect sensors

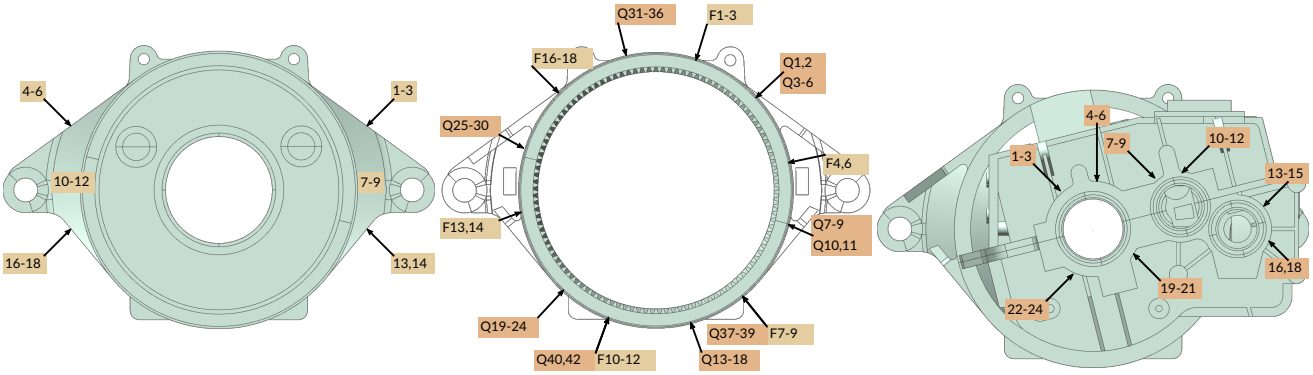
With the previously described sensor requirements in mind, a sensor installation has been realized on the gearbox. An overview of all available sensors is given in table 1. The following sections will discuss the different sensor types used in this research and discuss how their characteristics meet the criteria defined above. Section 3.4 will further validate the conditions of repeatable and torque-driven response.

3.3.1 | Strain gauges

Strain gauges are installed at several locations on the gearbox housing, divided into three main groups based on their location. Delta rosettes are used for all strain measurements, meaning that the strain is measured in three directions ($-45°$, $0°$ and $45°$) at each location. Each direction is considered as a separate measurement channel. Since strain gauges are simply glued to the existing housing, they are clearly non-intrusive sensors. The main loading mechanism linking torque to strain will be described for each group below, indicating expectations with respect to the criteria on torque-driven response and repeatability.

The first group of strain gauges installed on the TA and is visualized in fig. 2a. These sensor locations were derived from a sensor selection study, which was presented in an earlier publication by the authors [17]. The locations where adapted from the results of that study to realize a symmetrical layout in order to allow for easier interpretation of the measurements. The chief loading mechanism for the TA is the reaction torque produced by the gearbox due to the torque reduction. To counteract this torque, forces are generated at the connection points where the ends of the TA are connected to the bedplate. Due to these forces, one side of the TA undergoes compression while the other side is in tension, leading to a constant strain that scales with torque. One of the channels on the left side of the TA (15) malfunctioned during testing and was not further considered. This is indicated in fig. 2a by omitting the relevant label, and the same way of reporting will be used for the other strain gauge groups. The amount of channels reported in table 1 includes only the healthy channels.

A second group of strain gauges is installed on the external surface of the RG of the planetary stage. These strain gauges are installed on the middle line of the RG (in the axial direction) and at 12 approximately evenly spaced locations in the circumferential direction (see fig. 2b). The ring gear is loaded as its internal gear geometry meshes with the planet gears mesh, causing axial, radial and tangential force as well as bending moment



(a) Torque Arm **(b)** Ring Gear **(c)** High-Speed Housing

Figure 2 Strain gauge positions on different parts of the gearbox housing.

on the tooth root, all of which scale with the applied torque. This load causes a high but intermittent response, with a peak in the strain that occurs when one of the planet gears passes the sensor position [24].

The third group of strain gauges is installed on the HSH, on the structural parts supporting the bearings for the helical stages of the gearbox. These sensors are visualized in fig. 2c. Due to the radial and axial forces generated by the helical stages the strain on these parts will be driven by the load torque. However, the response is lower than on the TA because the force levels are considerably lower than in the planetary stage.

3.3.2 | Encoders

The sensor selection study also indicated the importance of encoder measurements measuring the rotation speed and angle of the gearbox [17]. While the response of these sensors is not directly torque-driven, these measurements are important to support the interpretation of the strain measurements, since the angular position of the planetary stage determines the force path from the planet gears to the RG and the rest of the gearbox housing. Hence, incremental encoders are included on the gearbox input and output shafts, measuring angular velocity. This velocity can be integrated in order to obtain an angle measurement, however this approach only yields a relative angle which can also be susceptible to drift. In order to provide an absolute angle that can be linked to the position of the planetary stage, a single pulse per rotation sensor was also included. Because the pulse occurs at a known angular position of the planetary stage, the pulses can be used to home the angle signal in post-processing.

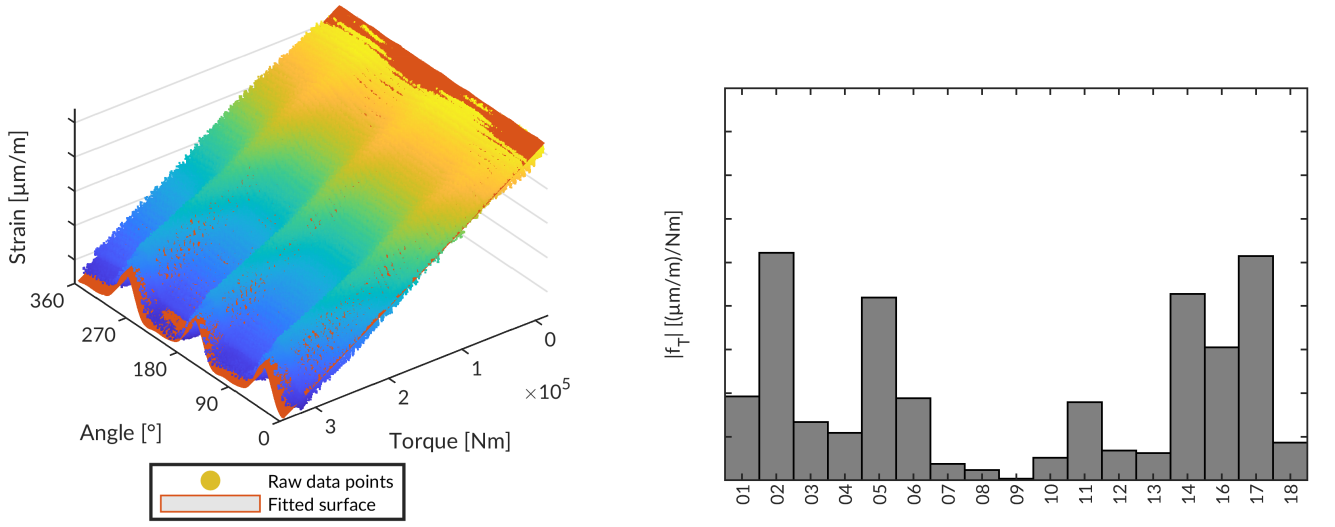
Encoder measurements are somewhat intrusive, since they must be installed on the rotating shafts. However in most encoders only passive elements (e.g. permanent magnets or an optical code disc) are added to the shaft, so the required modification is limited. Furthermore, in modern variable speed turbines at least one encoder is already present for control purposes, so conceptually this measurement could be re-used for virtual torque sensing as well.

3.4 | Strain gauge response study

In an earlier contribution, a methodology was proposed to validate the conditions of repeatable and torque-driven response for the strain gauges [19]. The approach is based on fitting a surface parametrized by the torque level T and the planet carrier angle θ_{PC} through the strain response measured at varying torque level. T and θ_{PC} represent a minimal selection of variables which can reasonably be expected to explain the variations in the measured response. On the one hand the torque is included, as main load component driving the response. On the other hand, the planet carrier angle determines where the contact force applied by the planet gears is exerted on the ring gear. This in turn drives global deformations and strain in the housing. In general, the fitted equation has the following form:

$$\epsilon\left(T, \theta_{\mathrm{PC}}\right) = f_T T f_\theta\left(\theta_{\mathrm{PC}}\right) + d, \tag{1}$$

where f_T is a constant factor describing torque sensitivity, $f_\theta\left(\theta_{\mathrm{PC}}\right)$ is a periodic function with period $360°$, and d is a constant that models an offset in the response that is not torque-driven. Only if the surface fit is of high quality, the response can be considered repeatable and torque-driven, since high fit quality implies that the chosen set of independent variables adequately explains the variance in the measured response. Surface fit quality is measured by the Root Mean Square Error (RMSE) and adjusted coefficient of determination ($\bar{R}^2$). Using this approach, it was concluded that for all strain gauges listed in table 1 the response is torque-driven and repeatable. The results of the approach are visualized in fig. 3a for a strain gauge on the TA.



(a) Torque arm strain gauge response sensitivity study.

(b) Overview of response sensitivity values for all torque arm strain gauges.

Figure 3 Results of the strain gauges response sensitivity study [19]. Numerical strain and strain sensitivity values have been redacted for reasons of confidentiality.

Using the same methodology, sensors can also be ranked according to the response magnitude, based on the linear sensitivity parameter f_T of the fitted surface (see eq. (1)). An overview of the sensitivity values of the strain gauges on the TA is shown in fig. 3b. According to the full ranking of all strain gauge channels, the strain gauges on the RG show the strongest response to torque. However, the response is intermittent with a large tensile strain peak when a planet gear is in contact close to the sensor location. Between the peaks the strain level is low. The strain gauges on the TA take the second place in the ranking, but these strain gauges do have the advantage of a more constant response. Finally, the strain gauges on the HSH show the lowest response sensitivity, due to the lower force levels in this part of the gearbox. Despite the differences in sensitivity, the sensitivity for all sensors is high enough to achieve an acceptable SNR at typical torque levels.

4 | GEARBOX SIMULATION MODEL IN SERVICE TO VIRTUAL SENSING

A physics-based model of the gearbox suitable needs to meet a few specific criteria in order to be suitable for virtual torque sensing. Firstly, the model should be able to predict the output from all sensors used in the measurement campaign within the frequency range of interest. Secondly, model fidelity (and expected accuracy) should be balanced against the implementation complexity and computational burden (since real-time execution in the context of a monitoring solution is envisioned). System behaviour that occurs outside the frequency range of interest or that has no detectable influence on the sensor output should be omitted from the model since it would needlessly increase computational burden and may render the system unobservable, leading to poor estimation results [10]. A model of the drivetrain and gearbox that fulfils these requirements is introduced in section 4.1. The correlation of strain measurements to the model predictions will be discussed in section 4.2, along with a model updating procedure [17,19].

4.1 | Model description

The model consists of two main components, which reflects two requirements that stem from the application and the chosen sensor set: The first model component is a lumped-parameter torsional model of the drivetrain. This component is necessary since the encoders in the test setup directly measure torsional states (i.e. angular velocity and position). Including torsional dynamics is also critical to achieve accurate torque estimation, since torsional stiffness and inertia influence the transmission of gearbox input torque to internal gearbox loads and sensor responses [25,13]. The second model component is an FE model of the gearbox housing which is necessary to predict the deformation of the housing and the resulting strain. Representative forces are applied to the FE model in the gear contact and bearing locations.

Three main model assumptions are used to link the model components and limit the model complexity: firstly, the torsional dynamics are partially decoupled from the housing deformation. Only one-way coupling from the torsional dynamics and resulting gear and bearing forces to the housing

deformation is included, meaning that the influence of the housing deformation on the torsional dynamics is omitted. This assumption is motivated by the fact that the interaction of housing deformation with torsional dynamics only has a limited, localized effect on contact force patterns in gears and bearings and the resulting local strain field. Due to the principle of St-Venant, a model with this assumption can still be expected to predict the strain with reasonable accuracy, as long as the measurement location is sufficiently far from the contact zones.

As a second assumption, linear, quasi-static behaviour is assumed for all non-torsional effects in the gearbox. With this assumption, the one-way coupling from torsional dynamics to housing deformation can be achieved by considering the static force and moment equilibrium for the shaft-gear-bearing system. Thirdly, linear, quasi-static behaviour is also assumed for the gearbox housing deformation. This means the gearbox housing deformation and strain model can be reduced to a generalized spring model without internal states. These assumptions of quasi-static behaviour were validated in a previous numerical and experimental analysis, which showed that the gearbox has no flexible eigenmodes in the frequency range of interest [26].

The final combined model thus uses the torsional model predict the encoder measurements, and, through combination with quasi-static equilibrium equations of the shaft-gear-bearing system, the forces applied to the gearbox housing. These housing forces are then applied to the statically reduced FE model of the housing to efficiently predict the strain output.

4.2 | Correlation of strain measurements to model and model updating

In an earlier publication, the correlation of the strain measurements to the strain predicted by the model introduced in the previous section was examined [19]. The correlation was studied for measurements taken at constant load and speed. The same conditions are reproduced in the model by applying the torque measured by the reference sensors on the main shaft. On the high-speed side of the gearbox, a viscous damping element models the generator counter-torque. The damper rating is tuned to reach the measured generator speed.

The measured strain is then compared to the simulated strain, this comparison is made in the angular domain of the planet carrier (θ_{PC}). Using the angular domain is possible since the measured strain is periodic and shows good repeatability, as demonstrated in section 3.4. The model results are completely periodic by definition through the model assumptions.

After the baseline correlation between the model and the measurements is established, a model updating step takes place in order to improve the correlation. The exact location of the strain gauges is subject to some uncertainty because of geometrical and measurement errors, hence the strain gauge positions are a logical target for the model update step [19,14]. An optimisation approach is used to find the node with the lowest Mean Absolute Error (MAE) in a patch of candidate nodes. This patch is defined as a collection of nodes around the nominal position of the strain gauge. Several patches are indicated on the FE model in fig. 4a.

From this study, it is concluded that the TA strain gauges already show a good correlation before the model update. After the model update, the results can still be improved by optimizing the strain gauge position. Figure 4b shows a comparison between the measured strain on the TA, and the simulated strain at the nominal strain gauge position, at the patch corner points, and at the optimized position. The results at the patch corner points are provided in order to illustrate the variability of the simulated strain over the patch.

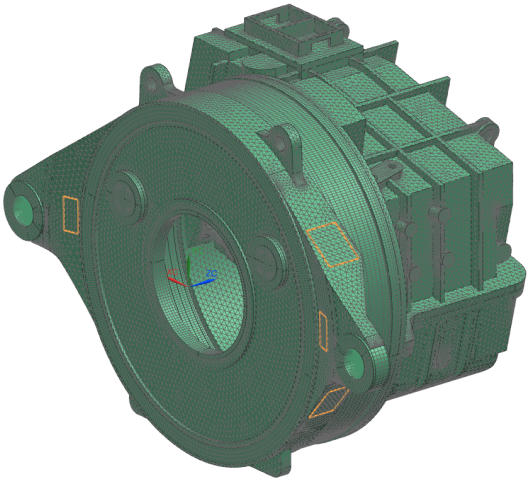
For the strain gauges on the RG, the model overestimates the strain peak due to the assumptions of the model, which represent the force applied by the planets on the RG as a single point load while in fact there exists a distributed contact load over the flank width of the teeth in contact. The optimized strain gauge position is found by moving in the axial direction, away from the middle of the RG. This lowers the response since the measurement node is now further away from the load concentration. This update does not necessarily represent the physical location of the strain gauge any more, but rather should be seen as an empirical correction on the model. However, due to the linearity in the strain response, this correction can be expected to hold at varying torque levels.

For the strain gauges on the HSH, the simulated strain at the nominal position shows a large offset in the mean strain. This offset is removed to a large extent when the position is optimized. A large spread on the simulated strain values exists across the patch, which implies that a large strain gradient exists in this location, due to complex geometry of the HSH and the local load of the bearing force.

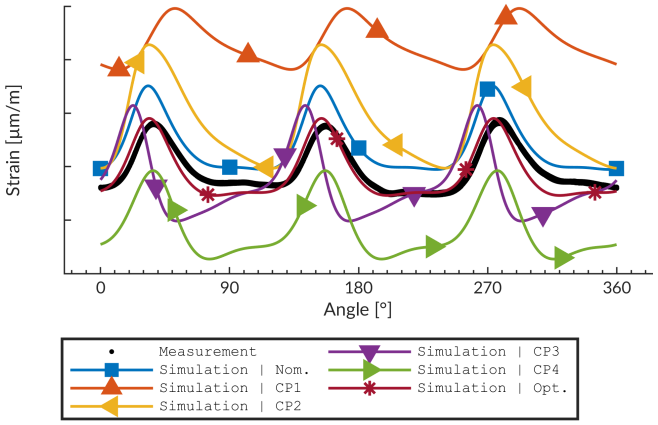
Next to these conclusions and the updated model, the approach also yields a ranking of the strain gauges in terms of MAE of simulated versus measured strain. This ranking can be used to select the sensor with the smallest modelling error in order to improved the virtual sensing performance.

5 | ESTIMATION ALGORITHM FOR VIRTUAL TORQUE SENSING

In this paper, the Augmented Extended Kalman Filter (AEKF) is used, which is a computationally efficient estimator capable of handling nonlinear systems with unknown inputs [27]. Before discussing the AEKF in section 5.2, we first introduce the required state space model structure in



(a) FE model of the gearbox with the patches of nodes used in the strain gauge position optimization highlighted.



(b) Correlation of simulated to measured strain, showing results for the nominal strain gauge position, optimized strain gauge position and Corner Points (CP) of the strain gauge patch. Numerical strain values have been redacted for reasons of confidentiality.

Figure 4 Correlation and model updating for strain gauges[19].

section 5.1. Next, the tuning of the covariance matrices for the AEKF is discussed in section 5.3, since these covariance matrices have a strong influence on the performance of the AEKF.

5.1 | Augmented state space model

The nonlinear, discrete-time state-space representation of the model discussed in section 4 is generally expressed as:

$$\mathbf{x}_k = \mathbf{f}\left(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{d}_{k-1}\right) + \mathbf{w}_{k-1} \tag{2}$$

$$\mathbf{y}_k = \mathbf{h}\left(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k\right) \qquad + \mathbf{v}_k, \tag{3}$$

where $\mathbf{x} \in \mathbb{R}^{N_x}$ is the state vector, $\mathbf{u} \in \mathbb{R}^{N_u}$ is the vector of known (control) inputs, $\mathbf{d} \in \mathbb{R}^{N_d}$ is a vector of unknown disturbances which should be estimated (in this paper the torque on main shaft and the high-speed shaft), and $\mathbf{y} \in \mathbb{R}^{N_y}$ is the measurement vector. All variables are defined at a discrete time step k, denoted in subscript. $\mathbf{f}$ and $\mathbf{h}$ are non-linear vector-valued functions which determine the state evolution, and the measurements, respectively. $\mathbf{w}$ and $\mathbf{v}$ are noise vectors that represent the mismatch between the model equations and the real system. These noise terms are assumed to be zero-mean Gaussian white noise processes, with (possibly time-variant) covariance matrices Q and R for $\mathbf{w}$ and $\mathbf{v}$, respectively.

The AEKF makes use of linear approximation of the system around the current operation point, involving the following Jacobian matrices:

$$\begin{aligned} A_k &= \left.\frac{\delta \mathbf{f}}{\delta \mathbf{x}}\right|_{\mathbf{x}_k} \quad B_{u_k} = \left.\frac{\delta \mathbf{f}}{\delta \mathbf{u}}\right|_{\mathbf{u}_k} \quad B_{d_k} = \left.\frac{\delta \mathbf{f}}{\delta \mathbf{d}}\right|_{\mathbf{d}_k} \\ C_k &= \left.\frac{\delta \mathbf{h}}{\delta \mathbf{x}}\right|_{\mathbf{x}_k} \quad D_{u_k} = \left.\frac{\delta \mathbf{h}}{\delta \mathbf{u}}\right|_{\mathbf{u}_k} \quad D_{d_k} = \left.\frac{\delta \mathbf{h}}{\delta \mathbf{d}}\right|_{\mathbf{d}_k}. \end{aligned} \tag{4}$$

In order to estimate the disturbance vector together with the states of the system, the disturbances are added to the state vector as augmented states. The state space model is then reformulated in terms of the augmented state vector:

$$\mathbf{x}_a = \begin{bmatrix} \mathbf{x} \\ \mathbf{d} \end{bmatrix} \tag{5}$$

$$\mathbf{x}_{a_k} = \underbrace{\begin{bmatrix} \mathbf{f}\left(\mathbf{x}_{a_{k-1}}, \mathbf{u}_{k-1}\right) \\ \mathbf{f}_{\mathbf{d}}\left(\mathbf{d}_{k-1}\right) \end{bmatrix}}_{\mathbf{f}_a(\ldots)} + \begin{bmatrix} \mathbf{w}_{k-1} \\ \mathbf{w}_{d_{k-1}} \end{bmatrix} \tag{6}$$

$$\mathbf{y}_k = \mathbf{h}\left(\mathbf{x}_{a_k}, \mathbf{u}_k\right) \qquad + \mathbf{v}_k, \tag{7}$$

where $\mathbf{f}_{\mathbf{d}}$ is an equation that describes the evolution the disturbances and $\mathbf{w}_{\mathbf{d}}$ is a zero-mean Gaussian white noise term with covariance matrix Q_d that represents deviations from this idealized equation. Generally, no prior information is available on the evolution of the disturbances, so a high covariance value is associated with $\mathbf{w}_{\mathbf{d}}$. As a result, the definition of $\mathbf{f}_{\mathbf{d}}$ only has a limited influence and a zero-order hold model is an appropriate choice: $\mathbf{f}_{\mathbf{d}}\left(\mathbf{d}_{k-1}\right) = \mathbf{d}_{k-1}$. This equation implies that on average, the disturbances remain constant. But, because of the noise vector $\mathbf{w}_{\mathbf{d}}$, the

AEKF can update the disturbance vector. $\mathbf{f}$ and $\mathbf{h}$ are the same vector-valued functions as in eq. (2), only the notation of the arguments is different. Generally, the different disturbances are considered uncorrelated, so Q_d is a diagonal matrix with $Q_{d_i}, i \in [1..N_d]$ on the main diagonal.

Finally, we also define the Jacobian matrices for the augmented system, and the augmented covariance matrix as:

$$\begin{aligned} A_{a_k} &= \left.\frac{\delta \mathbf{f}_a}{\delta \mathbf{x}_a}\right|_{\mathbf{x}_{a_k}} = \begin{bmatrix} A_k & B_{d_k} \\ 0 & I \end{bmatrix} \\ C_{a_k} &= \left.\frac{\delta \mathbf{h}_a}{\delta \mathbf{x}_a}\right|_{\mathbf{x}_{a_k}} = \begin{bmatrix} C_k & D_{d_k} \end{bmatrix} \\ Q_{a_k} &= \begin{bmatrix} Q_k & 0 \\ 0 & Q_{d_k} \end{bmatrix}. \end{aligned} \tag{8}$$

5.2 | The Augmented Extended Kalman Filter

The AEKF is a recursive stochastic estimator which processes measurements as they become available at each discrete time-step. Along with the mean value of the estimates, the AEKF also tracks their related uncertainty, represented by the estimation error covariance matrices $P_{\mathbf{x}_a}$ and $P_{\mathbf{y}}$, for the states and the measurements, respectively. These covariance matrices are used to determine optimized weighting factors between the information coming from the model and the measurements.

The AEKF is initialized by defining the initial state estimate $\hat{\mathbf{x}}_{a_0}^+$ and initial estimation error covariance $P_{\mathbf{x}_{a_0}}$. In these and future definitions, a quantity denoted by the hat symbol ($\hat{\bullet}$) indicates an estimated quantity. The estimation error covariance matrices are always estimated, hence no hat symbol is added to simplify notation. Ssuperscript $\bullet^-$ and $\bullet^+$ indicate results of the two estimation steps that form the main AEKF recursion, the a-priori and a-posteriori step, respectively.

The main AEKF recursion starts with the a-priori step, where the current estimates are propagated in time by the model (eq. (6)) to predict the state and measurements at the next time step:

$$\hat{\mathbf{x}}_{a_k}^- = \mathbf{f}_a\left(\hat{\mathbf{x}}_{a_{k-1}}^+, \mathbf{u}_{k-1}\right) \tag{9}$$

$$\hat{\mathbf{y}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{a_k}^-, \mathbf{u}_k\right). \tag{10}$$

The state and measurement covariance matrices and are also updated during the a-priori step, making use of the linearized state space model (eq. (8):

$$\begin{aligned} P_{\mathbf{x}_{a_k}}^- &= A_{a_{k-1}} P_{a_{k-1}}^+ A_{a_{k-1}}^T + Q_{a_{k-1}} \\ P_{\mathbf{y}_k}^- &= C_{a_k} P_{a_k}^- C_{a_k}^T + R_k. \end{aligned} \tag{11}$$

The second step of the AEKF recursion is the a-posteriori step, in which the a-priori estimate (which is entirely model-based) is updated based on the measurements recorded at time-step k. The correction is determined by the difference between the actual measurement and the prediction, a quantity which is known as the innovations $\mathbf{r}$. The innovation is weighted by the Kalman gain K_k to calculate the final correction:

$$K_k = P_{\mathbf{x}_{a_k}}^- C_{a_k}^T \left(P_{\mathbf{y}_k}^-\right)^{-1} \tag{12}$$

$$\hat{\mathbf{x}}_{a_k}^+ = \hat{\mathbf{x}}_{a_k}^- + K_k \underbrace{\left(\mathbf{y}_k - \hat{\mathbf{y}}_k^-\right)}_{\mathbf{r}}. \tag{13}$$

The Kalman gain matrix defined in eq. (12) can be interpreted as a relative comparison between the estimation error covariance matrices $P_{\mathbf{x}_{a_k}}^-$ and $P_{\mathbf{y}_k}^-$. If the elements of $P_{\mathbf{x}_{a_k}}^-$ are large compared to $P_{\mathbf{y}_k}^-$, K_k will increase, resulting in a significant update of $\hat{\mathbf{x}}_{a_k}^+$. Whether this situation arises depends on the properties of the system as expressed in the Jacobian matrices, the current state estimation error covariance matrix (which is itself the result of the previous steps of the AEKF) and the process and measurement noise covariance matrices. When K_k becomes large, this reflects a situation where more trust is given to the measurements, since the uncertainty (error covariance) of the measurements is lower. The reverse argument, in which case a small update $\hat{\mathbf{x}}_{a_k}^+$ is applied, also holds. With this definition of the Kalman gain, the AEKF minimizes the covariance of the estimation error for the linear approximation of the nonlinear system defined in eqs. (6) and (8)[27].

The a-posteriori step finished with an update of the state estimation covariance, which together with $\hat{\mathbf{x}}_{a_k}^+$ forms the input for the next estimation step:

$$P_{\mathbf{x}_{a_k}}^+ = \left(I - K_k C_{a_k}\right) P_{\mathbf{x}_{a_k}}^-. \tag{14}$$

5.3 | Covariance matrix tuning

As shown in eq. (11), the noise covariance matrices R, Q and Q_d influence the estimation error covariance matrices $P^-_{\mathbf{x}_{a_k}}$ and $P^-_{\mathbf{y}_k}$, which are used to define the Kalman gain (eq. (12)). As such, the setting of these matrices, often referred to as AEKF tuning, is crucial since it largely determines the accuracy of the AEKF. These matrices are typically tuned manually, requiring significant expertise and effort. Moreover, when comparing performance between different sensor sets, the performance depends on the tuning, making fair comparison difficult if the tuning is done manually. The next paragraphs will introduce automated, systematic tuning methods for each of the covariance matrices.

5.3.1 | Measurement noise covariance tuning

Measurement noise covers all deviations between the real measurement and the model-based prediction generated during the a-priori phase of the AEKF (eq. (10)). This includes noise caused by sensor or acquisition system imperfections such as electromagnetic interference or temperature variation, but also model inaccuracies. In this paper, an experimental characterisation of R is proposed, based on the errors observed during the model correlation and model update step discussed in section 4.2.

From the correlation study, the error between the model and the measurements is known over several revolutions of the gearbox. In fig. 5 this error is shown in the angular domain of the planet carrier for one representative strain gauge belonging to each group. These results illustrate that the error has a consistent periodic nature. This information can be embedded in a variable R matrix, which will ensure that the AEKF allocates a low weight (as expressed in the Kalman gain, eq. (12)) to a certain sensor when the prediction accuracy is poor, and vice versa.

The proposed approach splits the angular domain into bins, for each bin several observations the error between the simulated and the measured strain are available. These observations can be used to statistically characterize the error for each bin, leading to a mean error $\bar{v}_{i_j}$ and standard deviation $\sigma_{v_{i_j}}$ for strain gauge i and the bin j. Let us consider a zero-mean normal distribution, of which the probability density function envelops the observations for the bin. The standard deviation of this distribution can be defined as as:

$$\sigma_{i_j} = 2\left|\bar{v}_{i_j}\right| + \sigma_{v_{i_j}}. \tag{15}$$

This standard deviation can be used to build a variable R matrix:

$$R_j = \begin{bmatrix} \sigma^2_{1_j} & 0 & \dots & 0 \\ 0 & \sigma^2_{2_j} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma^2_{N_{y_j}} \end{bmatrix}. \tag{16}$$

However, in practice the definition in eq. (15) can be problematic for signals with intermittent behaviour like the strain gauges on the ring gear. During the strain peak, the mean error can become large, even though it is small relative to the actual strain gauge response. This would cause the estimator to ignore this sensor, exactly at the time it is giving its most informative measurement. In order to avoid this situation, a modified definition of σ_{i_j} is proposed, where the term related to the mean error is scaled depending on the mean strain registered for that bin:

$$\sigma^*_{i_j} = 2\left|\bar{v}_{i_j}\right| F_{i_j} + \sigma_{v_{i_j}}. \tag{17}$$

The scaling factor F_{i_j} is defined as:

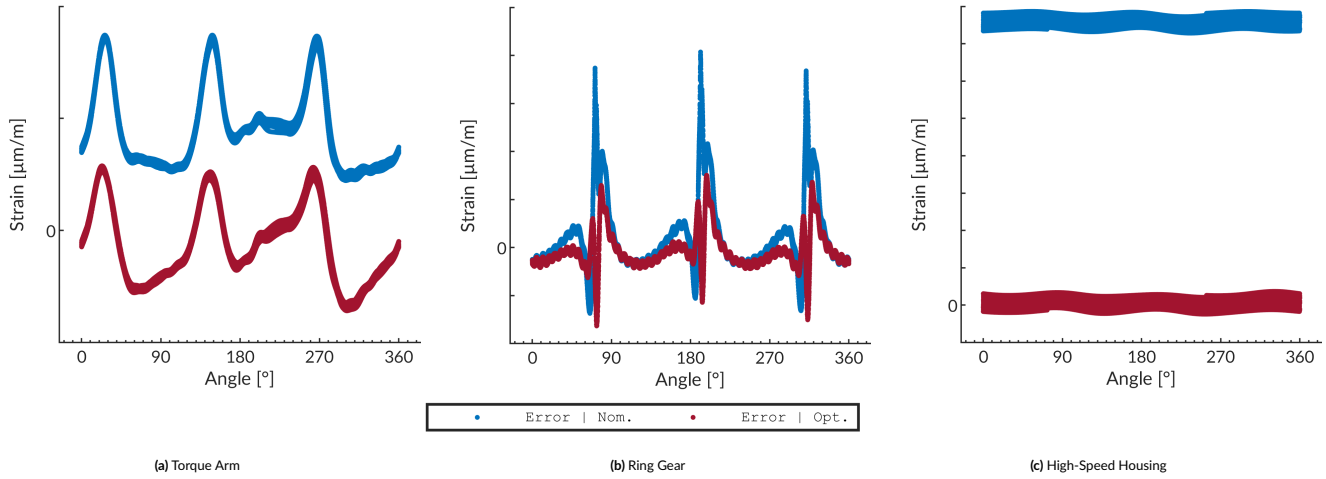
$$F_{i_j} = \frac{\epsilon_{m_i} - \left|\bar{\epsilon}_{i_j}\right|}{\epsilon_{m_i}}, \tag{18}$$

where $\bar{\epsilon}_{i_j}$ is the mean measured strain in bin j and ϵ_{m_i} is the maximum value of $\left|\bar{\epsilon}_{i_j}\right|$ across all bins. F_{i_j} will be close to zero when the strain is close to its maximum value, leading to a smaller $\sigma^*_{i_j}$. Conversely, F_{i_j} will become one when the strain is close to zero.

5.3.2 | State process noise covariance tuning

The process noise vector $\mathbf{w}$ describes all deviations between the model equations and the real evolution of the state vector (eq. (2)). Such deviation can originate from structural modelling errors (e.g. due to model assumptions), parametric errors, numerical discretization error, etc.

In this paper we use an approach to calculate the process noise covariance matrix Q based on parametric uncertainties [13]. The approach starts from a set of model parameters including drivetrain torsional inertia, viscous damper ratings, and gear and shaft stiffness values. For each model parameter, a normal distribution is assumed with the mean equal to the nominal parameter value p_{nom} and a certain standard deviation σ_p. A series of simulations is then performed, where the parameters are varied by sampling said normal distribution. For an individual simulation, the deviation of the results with respect to the results using the nominal parameters can be regarded as an observation of the random process error $\mathbf{w}$. When collected, these observations can be used to determine Q, by calculating the sample covariance matrix.



(a) Torque Arm (b) Ring Gear (c) High-Speed Housing

Figure 5 Error of simulated strain versus measured strain, given for the nominal strain gauge postition and the optimized strain gauge position. Numerical strain values have been redacted for reasons of confidentiality. All tick mark intervals are equivalent, allowing a qualitative comparison between the error values of different strain gauges.

This way, the problem of tuning the full covariance matrix (including off-diagonal elements) is transformed into the problem of tuning the probability distribution for the parameters. It is generally reasonable to assume that the parameters are uncorrelated. Hence, only the standard deviation of each parameter needs to be defined. As a further advantage, the physical meaning of the standard deviation of the parameters is easier to interpret than that of the elements of Q.

For each parameter, a standard deviation equal to $10\,\%$ of the nominal value is assumed. Although this tuning method likely underestimates the uncertainty because only parametric errors are taken into account, earlier results have shown that the resulting estimation accuracy when employing this tuning is comparable to the accuracy that can be reached through manual tuning [13]. Consequently, the manual tuning procedure is avoided while achieving a similar estimation accuracy.

5.3.3 | Disturbance process noise covariance tuning

Unlike the other noise vectors, the disturbance process noise vector $\mathbf{w_d}$ does not represent the output of a stochastic process. Instead, it serves as a mathematical tool to allow the AEKF to make updates to the estimated disturbances. Hence the covariance matrix Q_d is usually considered as a regularization factor [11]. High covariance values allow large changes to the estimated disturbance, creating a responsive estimator. The downside is that a highly responsive estimator will also attempt to track the process and measurement noise, leading to a noisy disturbance estimation. Some degree of noise rejection is a desirable property for the estimator, so a trade-off must be made. In extreme cases high covariance can also cause numerical instability in the AEKF.

When Q_d is considered as a regularization factor, standard regularization parameter estimation methods such as the L-curve can be used, as demonstrated in the work of Lourens [11]. In this approach, the so-called solution norm, defined as $n_{\mathbf{d}} = \sum_{k=1}^{N} \left\| \hat{\mathbf{d}}_k \right\|_2^2$, and the innovation norm $n_{\mathbf{r}} = \sum_{k=1}^{N} \|\hat{\mathbf{r}}_k\|_2^2$ are compared for different values of Q_d. A good regularization is characterized by a relatively small value for both norms, since a small value for $n_{\mathbf{d}}$ indicates that the estimated disturbance is smooth and does not contain extremely large values, and a small value for $n_{\mathbf{r}}$ indicates that the estimated and measured sensor response are in agreement. For standard regularization problems, an L-shaped curve is observed when $n_{\mathbf{d}}$ and $n_{\mathbf{r}}$ are plotted for different regularization parameter values, hence the name of the method. An optimal trade-off of both norms is found in the corner of the L-curve [28]. When applied to the AKEF tuning problem, it is noted that L-shape does not always occur [11]. In a related approach proposed by Petersen, only the innovation norm $n_{\mathbf{r}}$ is minimized [12]. Both approaches require performing the AEKF estimation multiple times with varying Q_d, which is time consuming. In this paper, the approach proposed by Petersen (minimum innovation norm) will be used as a reference method, to benchmark an alternative tuning method proposed in the next paragraphs.

A novel tuning method for Q_d is proposed that is based on the closed-loop Frequency Response Function (FRF) of the AEKF. Specifically, the FRF of the estimated disturbance vector ($\hat{\mathbf{d}}$) to the real (unknown) disturbance vector ($\hat{\mathbf{d}}$) is considered. The tuning method will modify Q_d so the FRF matches the desired functional shape as closely as possible. This method is formulated as an optimization problem, the precise definition of which will follow after the definition of the FRF itself.

The aforementioned FRF can be constructed using a linearized approximation of the state space model and the steady-state Kalman gain matrix K_∞ [13,27]. In the derivations that follow, the dependency of the system linearization (eq. (4) and eq. (8)) on the timestep k is omitted , i.e. the suffix k is dropped for simplicity of notation while discussing the tuning method. The approach is extended towards time variant systems in the last paragraph of this section. Note that the the dependency on the timestep is not caused by true time-variant behaviour, but rather by the non-linear nature of the system equations (eq. (2)). Since the state of the gearbox is constantly changing, different matrices are obtained when the system is linearized at different operating points.

The calculation of virtual sensor FRF starts from the calculation of K_∞, which is the value that the Kalman gain converges to, assuming the system is time-invariant and the input and noise terms are statistically stationary processes. This matrix can be used as an approximation of the Kalman gain, which is available without actually running the AEKF. As such, the dynamic behaviour of the AEKF can be predicted efficiently. The steady-state augmented state covariance matrix P_{a_∞} is required to calculate K_∞, it is found by solving the Discrete Algebraic Ricatti Equation (DARE)[27]:

$$P_{a_\infty} = A_a P_{a_\infty} A_a^T - A_a P_{a_\infty} C_a^T \left[C_a P_{a_\infty} C_a^T + R \right]^{-1} C_a P_\infty A_a^T + Q_a. \tag{19}$$

The solution to eq. (19) is found using an implicit solver. K_∞ is found by inserting P_{a_∞} in eq. (12) for the Kalman gain:

$$K_\infty = P_{a_\infty} C_a^T \left[C_a P_{a_\infty} C_a^T + R \right]^{-1}. \tag{20}$$

Next, the actual FRFs are constructed. Figure 6 shows the block diagram of the state-space system and Kalman Filter in steady-state, which is used in this derivation. In the diagram, as well as in the following derivations, the capitalized vectors $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{U}$, and $\mathbf{D}$ represent z-transform of the state, measurement, input and disturbance vector, respectively. First, the characteristics of the system are considered, in the form of a Transfer Function (TF) from the measurements $\mathbf{Y}(z)$ to the real disturbances $\mathbf{D}(z)$ which is given by:

$$H_{\mathbf{dy}}(z) = C\left[zI - A\right]^{-1} B_d + D_d. \tag{21}$$

Next, the TF from the measurements to the estimated disturbance $\hat{\mathbf{D}}(z)$ is calculated:

$$H_{\mathbf{y\hat{d}}}(z) = S_d \left[I - (I - K_\infty C_a) A_a z^{-1} \right]^{-1} K_\infty, \tag{22}$$

where S_d is a selection matrix to extract the unknown loads from the augmented state vector. Finally, the TFs in eqs. (21) and (22) are combined and evaluated on the unit circle (i.e. $z = e^{j2\pi f}$) to derive the FRF from the real input loads to the estimated loads: $H_{\mathbf{d\hat{d}}}\left(e^{j2\pi f}\right) = H_{\mathbf{y\hat{d}}}\left(e^{j2\pi f}\right) H_{\mathbf{dy}}\left(e^{j2\pi f}\right)$. Ideally, this FRF should be equal to one in the frequency range of interest, implying that the estimated disturbance equals the real value. Outside the frequency range of interest, the FRF should drop off in order to ignore any noise that might be present in this frequency range. This can be formulated as an optimization problem as:

$$\min_{Q_d} 0.5 \int_0^{f_c} \left| 1 - \left| H_{\mathbf{d\hat{d}}}\left(e^{j2\pi f}\right) \right| \right| df + 0.5 \int_{f_c}^{f_s/2} \left| H_{\mathbf{d\hat{d}}}\left(e^{j2\pi f}\right) \right| df, \tag{23}$$

where f_c is the cut-off frequency for the target FRF and f_s is the sampling frequency. f_c can be set based on the required bandwidth of the virtual load sensors (e.g. the $10\,\mathrm{Hz}$ target considered in this paper) or based on prior knowledge of the expected frequency content of $\mathbf{d}$. A standard gradient-free optimizer (pattern search) is used to solve eq. (23).

For the model discussed in section 4, the linearized matrices are not constant. In order to cope with this, the tuning of Q_d could be repeated at each iteration. However, this would lead to a prohibitively high computation cost. Fortunately, practical experience shows that the resulting tuning does not change significantly over typical state trajectories. Furthermore, it is commonly observed that the AEKF performance is relatively insensitive to the precise value of Q_d near its optimal value [13,11]. Hence, a constant Q_d matrix is adopted, equal to the median of the Q_d matrices achieved over a number of operating states.

The benefit of the proposed method is that there is no need to perform the AEKF estimation in order to tune Q_d. All that is required is a number of realistic operating states, at which the linear system approximation, K_∞ and FRFs can be computed. An open-loop simulation or measurements can be used to attain the required operating states. Furthermore, the cut-off frequency for the target FRF f_c is a tuning parameter for the method with a direct physical interpretation. The definition of the cost function in eq. (23) can be adapted based on the system properties, giving flexibility to the method.

6 | EXPERIMENTAL VALIDATION OF VIRTUAL TORQUE SENSING

This section validates the virtual torque sensing approach on the start-stop sequence in section 2, using the sensors, models, estimation algorithm, and tuning methods introduced in previous sections. Section 6.1 will start off with a method to select the active channels from among the 101 strain gauge channels available (see table 1). Next, section 6.2 demonstrates the performance of the virtual torque sensing approach for different

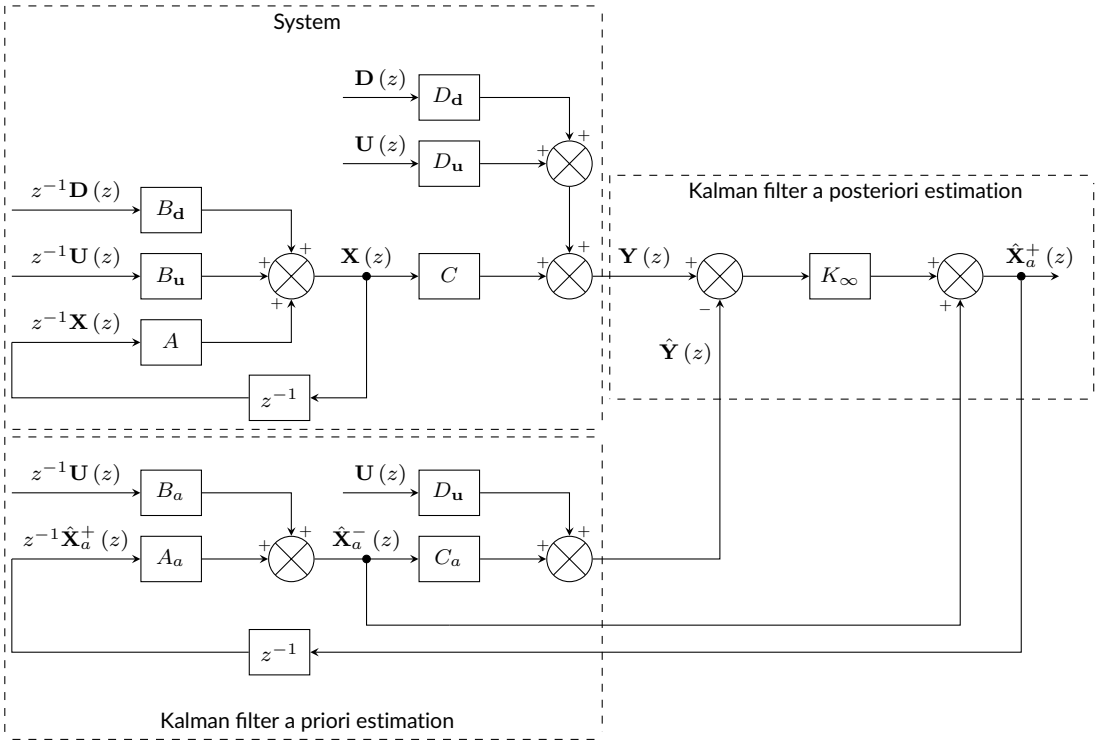


Figure 6 Closed loop block diagram of the system and Kalman Filter in steady-state.

sensor sets picked from the three strain gauge groups discussed in section 3.3.1. The influence of the amount of strain gauges included is also discussed. Finally, section 6.3 discusses the influence of the tuning method for the disturbance process noise Q_d on the virtual torque sensing results. A comparison between the reference tuning method (minimum innovation norm) and the proposed FRF-based method is made, and the influence of the cut-off frequency on the results of the FRF-based method is investigated.

6.1 | Active channel selection: Optimizing sensor response and model prediction error

Earlier sections presented two metrics to rank sensors based on two properties desirable for virtual torque sensing: in section 3.4, sensors were ranked based on their response as quantified by the linear fit parameter f_T; and in section 4.2, the MAE between measured and simulated response is calculated (as part of the model updating), the MAE can also be used to rank sensors according to predictability.

Since there is no clear priority between both metrics and since no obvious reasonable weighting between the metrics is apparent, the active channel selection is approached similar to a multi-objective objective optimization problem. Figure 7 presents the active channel selection approach visually by representing each strain gauge as a marker characterized by its response magnitude and MAE. In the results that follow, the active strain gauges used will be found by selecting gauges on or near one of the Pareto fronts.

6.2 | Virtual torque sensing with different sensor sets

In this section, virtual torque sensing results will be presented using the three different strain gauge groups. For each group, the best sensors are selected on or near the respective Pareto front considering only the strain gauges in that group (see fig. 7. In the selection, all strain gauges for which the MAE between measurements and simulations exceeds $5\,\frac{\mu\mathrm{m}}{\mathrm{m}}$ are disregarded. All results are generated using the automated AEKF tuning methods discussed in section 5.3. Specifically, the FRF-based tuning method for the disturbance noise covariance was used, using a cut-off frequency of $10\,\mathrm{Hz}$, equal to the limit of the frequency range of interest.

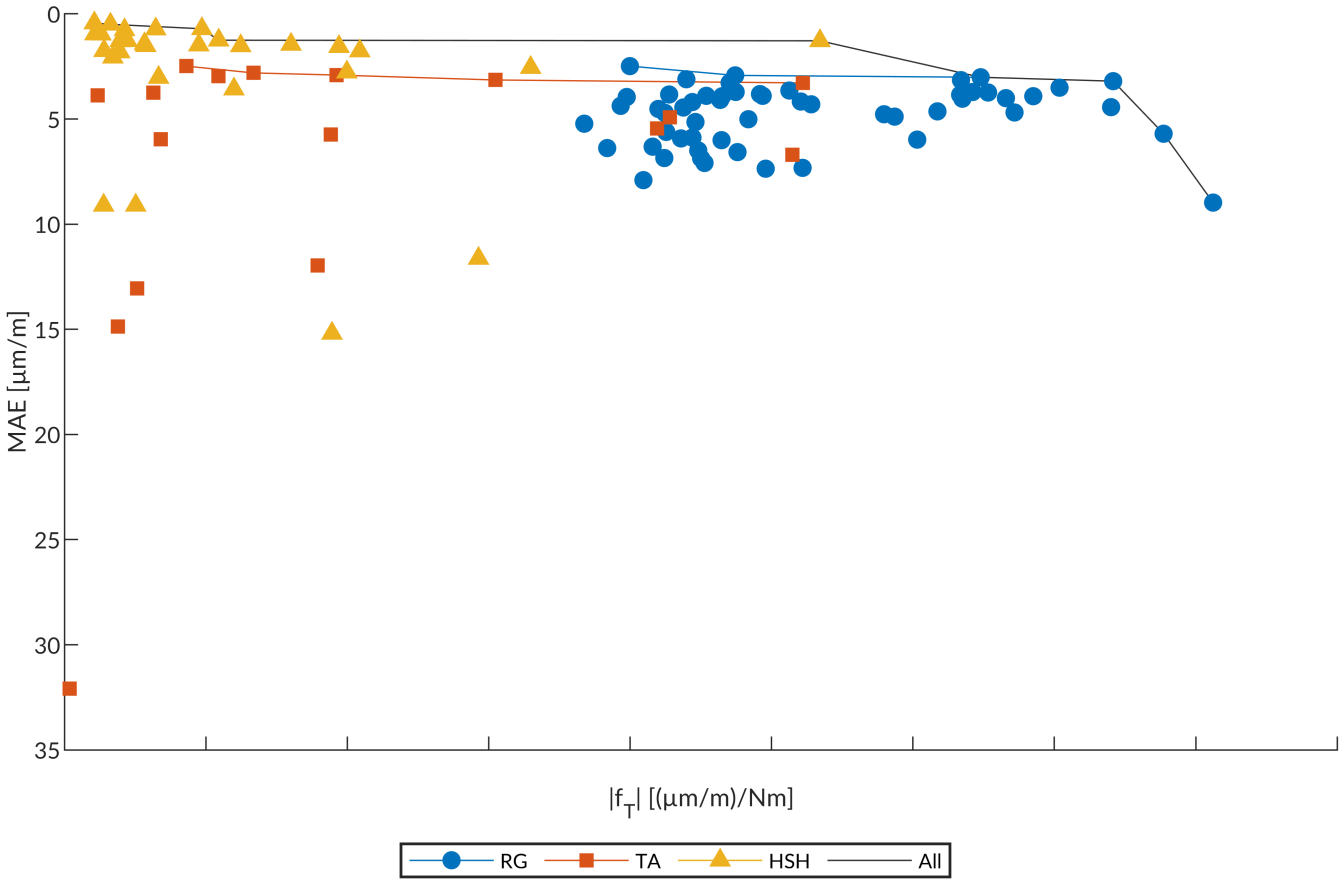


Figure 7 Visual representation of the active channel selection approach. Each marker represents a strain gauge, characterized by its response magnitude and MAE of simulated versus measured response. Since high response and low MAE are desired, the optimal selection is found in the upper right corner (note that the y-axis has been reversed to enhance visualisation). Four different Pareto fronts are displayed, representing the optimal selection of sensors for each of the strain gauge groups as well as the global Pareto front when all strain gauges are considered simultaneously. Numerical strain sensitivity values have been redacted for reasons of confidentiality.

6.2.1 | Torque arm strain gauges

Figure 8 shows the virtual torque sensing result using the angular encoder and 3 strain gauges on the torque arm. The estimated torque is compared to the torque that was measured by the validation sensor. Over the entire time window shown in fig. 8, the virtual torque sensor shows very good estimation results, with a Normalized Mean Absolute Error (NMAE) of $3.41\,\%$, when normalized by the nominal gearbox torque level.

During the first $20\,\mathrm{s}$, the torque and speed are ramped up to the nominal values. During this interval, the NMAE is slightly higher than the overall mean, with $3.99\,\%$. From $20\,\mathrm{s}$ to $60\,\mathrm{s}$, the torque and speed are kept constant. The virtual torque sensor shows excellent tracking during this period, with an NMAE of $3.10\,\%$. A slight ripple is visible on the estimated torque during this period which can be associated with a small mismatch between the predicted and the real measurements, which persists despite the model updating performed in section 4.2. It is a commonly observed phenomenon when using the AEKF that such a mismatch appears in the estimated torque [13]. Shortly after the $60\,\mathrm{s}$ mark, the torque and speed are ramped back down. During this ramp we again see good tracking. Even the negative torque spike shortly after the $70\,\mathrm{s}$ mark is captured relatively well. This is remarkable because the backlash that will be caused in the drivetrain due to this negative torque is currently not included in the models. The NMAE in the interval from $60\,\mathrm{s}$ to the end of the time window is $3.45\,\%$.

The virtual torque sensor performance with fewer and more strain gauges is summarized in table 2. Even with a single strain gauge the virtual torque sensor is already achieves an NMAE of $3.57\,\%$. Adding more sensors improves the performance, but with diminishing returns.

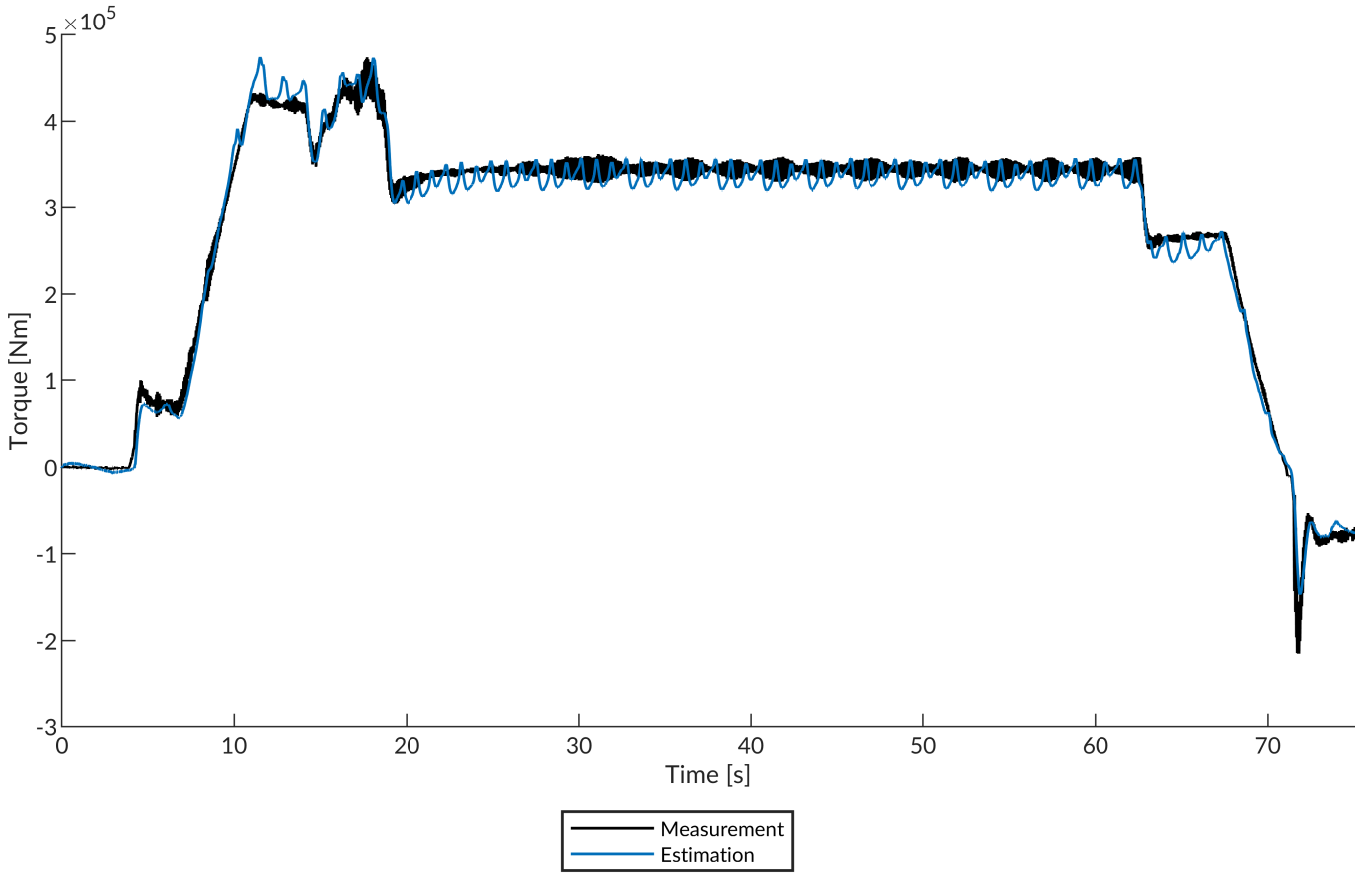


Figure 8 Virtual torque sensing results using an angular encoder and 3 strain gauges on the torque arm.

6.2.2 | Ring gear strain gauges

Figure 9 shows the virtual torque sensing result using the angular encoder and 3 strain gauges on the ring gear. Over entire time window, the performance of the virtual torque sensor is slightly worse with this sensor set, as compared to the previous results, with NMAE of $4.59\,\%$. This is primarily caused by worse performance during the torque ramps, during which a slight staircase pattern can be seen in the estimated torque. This effect is caused by the intermittent nature of the sensor response in this location. During the peaks in the strain, the estimator is able to update the torque to the correct value while in between the peaks, the torque estimate stays roughly constant. As a result the NMAE is higher during the upwards torque and speed ramp ($4.65\,\%$). During the dwell, the NMAE is $3.11\,\%$, which is comparable to the result attained with the torque arm strain gauges.

During the downward ramp, the same staircase pattern can be seen. The negative torque spike shortly after the $70\,\mathrm{s}$ mark is not captured well. Unlike the torque arm strain gauges, the strain on the ring gear is strongly affected by the backlash in the gearbox as the torque changes direction, which is not included in the current model. Furthermore, through coincidence the negative torque spike occurs when none of the planet gears are close to a strain gauge, leading to low response in the sensors. The low speed of rotation exacerbates the latter issue. The NMAE in this final interval is $8.47\,\%$.

Also for this group of sensors the performance with fewer and more strain gauges is summarized in table 2. Observe that with a single strain gauge the virtual torque sensor performance is much worse, this is due to an increase of the staircase effect during the torque ramp (NMAE $7.27\,\%$). Adding more sensors again improves the performance with diminishing returns.

6.2.3 | High-speed housing strain gauges

Figure 10 shows the virtual torque sensing result using the angular encoder and 3 strain gauges on the high-speed housing. Over entire time window, the performance of the virtual torque sensor is worse with this sensor set, as compared to the previous results, with NMAE of $7.47\,\%$. Again, the difference in performance is most apparent during the torque ramps. However, even during the dwell, a slight drift of the estimated torque can be seen. This drift is caused by a drift in the strain signals, which show a strain response increasing by approximately $5\,\frac{\mu\mathrm{m}}{\mathrm{m}}$ even though

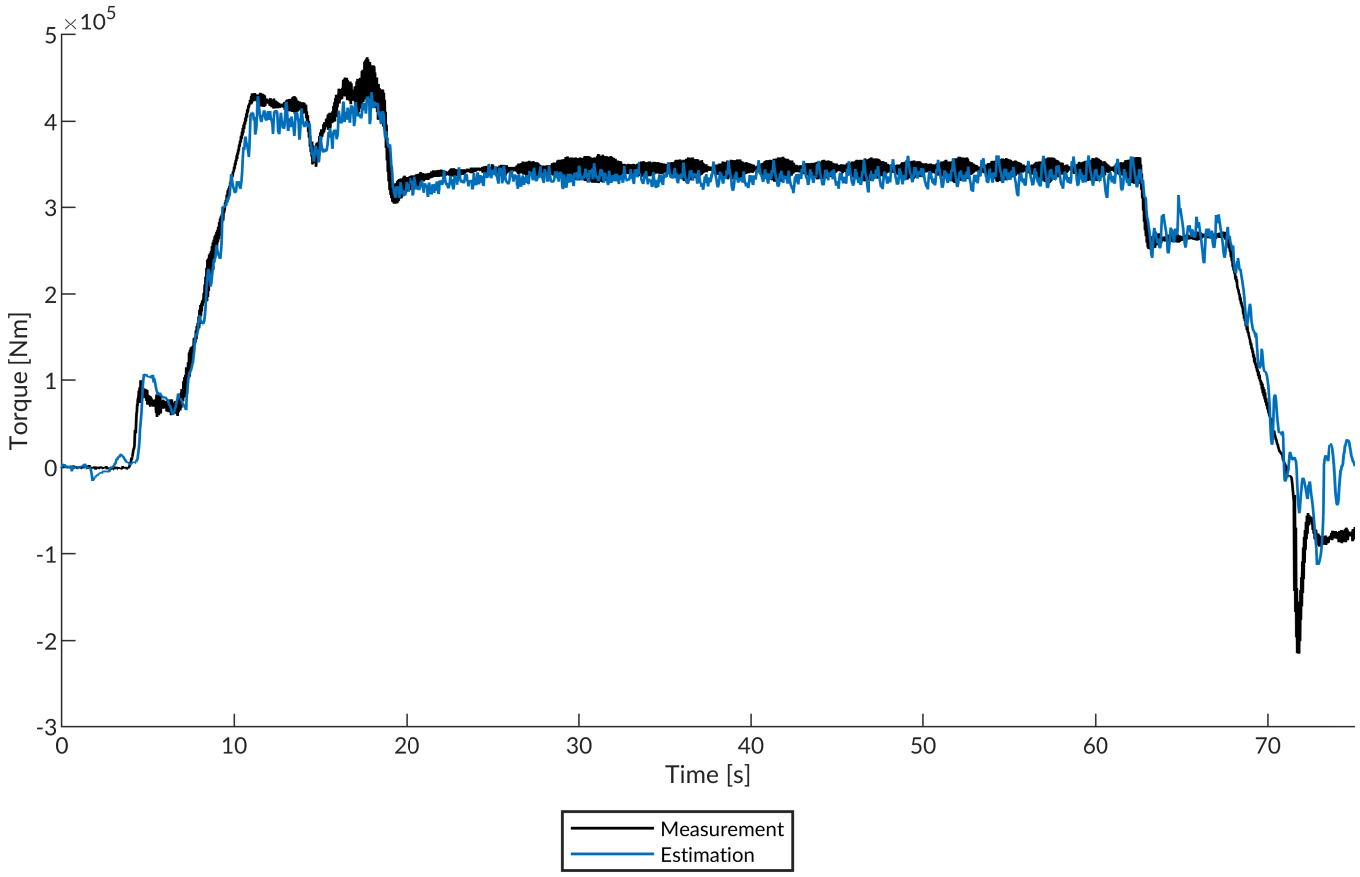


Figure 9 Virtual torque sensing results using an angular encoder and 3 strain gauges on the ring gear.

Table 2 Overview of virtual torque sensing results utilizing different sensor sets.

Strain gauge group	# sensors	NMAE [%]			
		Full	Ramp ↑	Dwell	Ramp ↓
Torque arm	1	3.57	3.80	3.46	3.57
	3	3.43	3.43	3.41	3.49
	5	3.41	3.99	3.10	3.45
Ring gear	1	7.27	5.77	6.46	11.46
	3	4.59	4.65	3.11	8.47
	5	3.90	3.92	2.67	7.15
High-speed housing	1	6.81	8.70	2.44	15.96
	3	7.47	6.92	3.96	17.56

the torque remains constant. All strain gauges on the high-speed housing show this behaviour and the drift is similar for all channels, hence it cannot be attributed to a single faulty sensor. Temperature variation can explain part of the strain variation, but the recorded temperature variation (approximately $1\,^{\circ}\mathrm{C}$) is too small to offer a complete explanation. Non-linear effects such as creep in the many bolted connections of the high-speed housing could also explain the observed variation, but this could not be proven conclusively.

From the summary given in table 2, we see that adding additional sensors on the high-speed housing counter-intuitively reduces the performance of this sensor set. This happens because the active channel selection first selects the strain gauges with the highest response, for the sensor with the highest response the drift described above is less significant relative to the total sensor response.

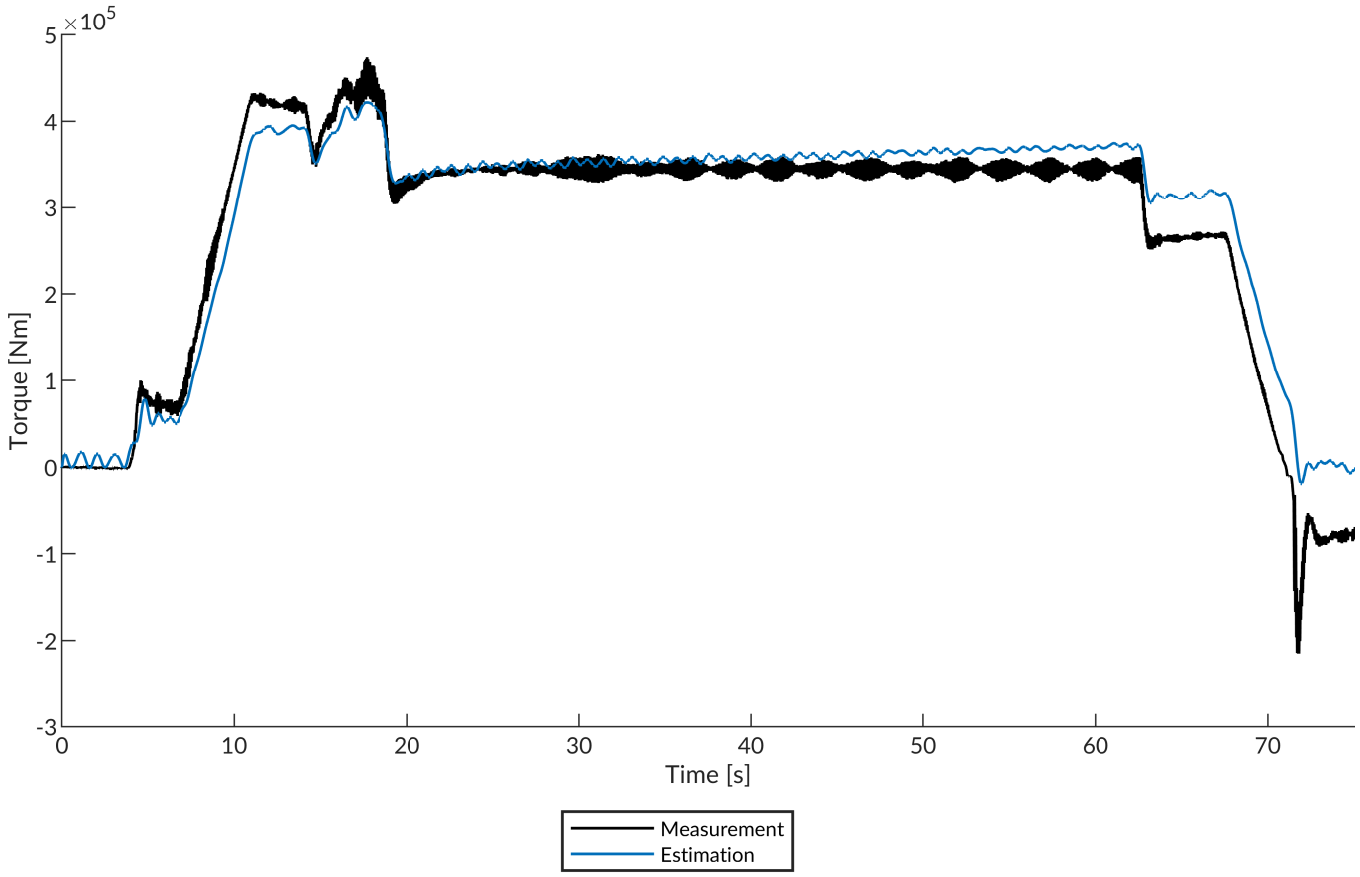


Figure 10 Virtual torque sensing results using an angular encoder and 3 strain gauges on the high-speed housing.

6.3 | Influence of the disturbance process noise tuning method on virtual torque sensing results

The influence of the disturbance process tuning method is investigated by comparing the proposed FRF-based method to the reference method based on minimizing the innovation norm $n_{\mathbf{r}}$. The influence of the cut-off frequency f_c is also investigated by tuning the AEKF with several different frequencies. The results are summarized in table 3, for the same load case as considered in the previous section. For each strain gauge group, 3 active channels are used, using same selection as before.

From table 3 we see that generally the reference tuning method results in a higher value for the elements of Q_d. This in turn results in a more responsive estimate, but also in a noisier estimate. For all considered sensor sets, the proposed FRF-based tuning with $f_c = 10\,\mathrm{Hz}$ outperforms the reference method, indicating that FRF-based tuning introduces a more appropriate level of noise rejection while still retaining sufficient responsiveness.

When the cut-off frequency for the FRF-based method is lowered to $1\,\mathrm{Hz}$, the values in Q_d are lowered, leading to lower responsiveness and improved noise rejection. This behaviour also reflected in the changes to the NMAE during the different segments of the measurement. For all sensor sets, NMAE is increased during the torque ramps and decreased during the torque dwell, since responsiveness is not relevant during the dwell. The overall NMAE is slightly higher with $f_c = 1\,\mathrm{Hz}$, but this conclusions cannot be generalized towards other operating conditions which may contain a different mix of ramps, dwells and other dynamic torque and speed changes.

When the cuf-off frequency is increased to $100\,\mathrm{Hz}$, the value for Q_{d_1} (for the torque on the main shaft) remains roughly constant while the value for the value for Q_{d_2} (for the torque on the high-speed shaft) still increases. This indicates that the FRF for the main shaft torque is insensitive to further increases of Q_{d_1}. The estimation performance is similar to the performance with $f_c = 10\,\mathrm{Hz}$.

Table 3 Overview of virtual torque sensing results utilizing different tuning methods for the disturbance noise covariance Q_d. All results pertain to a sensor set containing an angular encoder and 3 strain gauges from the strain gauge group indicated in the first column.

Strain gauge group	Tuning method	$\log_{10} Q_{d_1}$	$\log_{10} Q_{d_2}$	NMAE [%]			
				Full	Ramp ↑	Dwell	Ramp ↓
Torque arm	$\min_{Q_d} n_{\mathbf{r}}$	8.0	4.0	3.91	3.71	3.98	4.01
	FRF, $f_c = 1\,\mathrm{Hz}$	4.3	0.8	3.80	5.36	2.71	4.62
	FRF, $f_c = 10\,\mathrm{Hz}$	5.6	2.3	3.43	3.43	3.41	3.49
	FRF, $f_c = 100\,\mathrm{Hz}$	5.4	5.1	3.42	3.44	3.39	3.46
Ring gear	$\min_{Q_d} n_{\mathbf{r}}$	8.0	6.0	11.49	8.94	11.37	15.19
	FRF, $f_c = 1\,\mathrm{Hz}$	4.8	1.1	6.11	8.19	2.36	13.35
	FRF, $f_c = 10\,\mathrm{Hz}$	6.4	2.5	4.59	4.65	3.11	8.47
	FRF, $f_c = 100\,\mathrm{Hz}$	6.4	5.3	5.28	4.87	4.07	9.05
High-speed housing	$\min_{Q_d} n_{\mathbf{r}}$	8.0	6.0	9.93	6.03	7.51	18.54
	FRF, $f_c = 1\,\mathrm{Hz}$	3.6	−0.6	7.57	7.07	3.88	18.08
	FRF, $f_c = 10\,\mathrm{Hz}$	5.1	1.2	7.47	6.92	3.96	17.56
	FRF, $f_c = 100\,\mathrm{Hz}$	5.6	2.5	7.50	6.92	4.03	17.55

7 | CONCLUSIONS

In this paper, a virtual torque sensing approach for wind turbine gearboxes has been introduced and verified experimentally. The components that make up the virtual load sensor are: a set of non-intrusive sensors installed on the gearbox, a physics-based model capable of predicting the response of aforementioned sensors, and an estimation algorithm that combines the measured response with predictions from the model to infer the gearbox input load.

The non-intrusive sensor set installed on the gearbox consists of an angular encoder and strain gauges mounted on the external surface of the gearbox housing. The strain gauges can be divided into 3 main groups based on the component on which they are installed, being the Torque Arm (TA), Ring Gear (RG), and High-Speed Housing (HSH). A sensor response sensitivity study was performed in order to rank all sensors according to their response to torque. This study concluded that the strain gauges on the RG showed the highest response, although the response is intermittent, showing a large strain peak when a planet gear meshes with the RG close to the sensor and almost no response in between the peaks. The strain gauges on the TA show the second highest response and the strain gauges on the HSH have the lowest response of the sensors considered here. However, these last two strain gauge groups have the advantage of a more constant response to torque when compared to the strain gauges on the RG. Despite the differences in sensitivity, all strain gauges have an adequate signal to noise ratio.

The physics-based model used in this paper strikes a balance between accuracy and computational cost through model-order reduction and modelling assumptions tailored to the virtual torque sensing application. The resulting model is validated and updated to ensure it accurately predicts the strain measured on the housing. Based on the model validation and updating procedure, it is also possible to rank the sensors in function of the error between the measured and the simulated response.

The Augmented Extended Kalman Filter (AEKF) is used as estimation algorithm in this paper. The AEKF is a computationally efficient estimator capable of handling non-linear systems with unknown inputs. A key factor determining the performance of the AEKF is the tuning of the covariance matrices for the process, measurement, and disturbance noise. In this paper, methods were introduced to systematically tune the covariance matrices. With these methods, time-consuming manual tuning is avoided. Furthermore, estimation results with different sensor sets can be compared fairly, since a uniform tuning method is used for all sensors. Particular focus is given to the tuning of the disturbance noise covariance matrix since it has a strong influence on the trade-off between responsiveness and noise rejection of the virtual load sensor. A novel tuning method is developed for this covariance matrix, based on the theoretical Frequency Response Function (FRF) of the virtual load sensor. The FRF-based method is shown to outperform a reference method from the state of the art.

Virtual torque sensing results are demonstrated with each of the aforementioned strain gauge groups. From each strain gauge group, the active channels are selected by maximizing the sensor response magnitude and minimizing the error between the measured and the simulated response. Since there is no clear priority between both metrics and since no obvious reasonable weighting between the metrics is apparent, the active channel selection is approached similar to a multi-objective objective optimization problem. This allows a Pareto front to be constructed, identifying the sensors with the best performance for both objectives. Sensors on or near the Pareto front are selected for virtual torque sensing.

The virtual torque sensor shows promising results, with a Normalized Mean Absolute Error (NMAE) below $8\,\%$ achieved for all sensor sets, when normalized with the nominal gearbox torque level. The best performance was achieved with the strain gauges on the torque arm, achieving a NMAE of $3.41\,\%$ using 5 strain gauges. The strain gauges on the ring gear give the second best performance, with an NMAE of $3.90\,\%$, again using 5 strain gauges. When using the strain gauges on the HSH, the performance is worse, with NMAE of $6.81\,\%$, using a single strain gauge. This poorer performance is caused by a drift in the measured strain on the HSH, which could be caused by a temperature variation, or by non-linear effects such as creep in the bolted connections. The precise cause of the drift in the measured strain on the HSH could not be established, and is a subject for future investigation.

The effect of the amount of sensors on the performance is also studied. For the TA and the RG, increasing the amount of sensors improves the performance, but a diminishing return is observed for larger sensor sets. The increase in performance is more significant when using the sensors on the RG, because of the intermittent nature of the measured signals. When using the strain gauges on the HSH, adding additional sensors actually decreases the performance. This counter-intuitive result is again explained by the drift in the measured strain.

Based on these results, we conclude that the virtual sensing approach is a promising technique to measure load on a wind turbine gearbox indirectly. This technique allows monitoring the crucial but hard to measure gearbox torque, paving the way for more advanced condition monitoring strategies, for example by calculating the Remaining Useful Life (RUL) of different gearbox components. Ultimately such strategies can lead to improved maintenance planning and lifetime extension. Lessons learned from such a monitoring solution could also be applied in future designs to avoid over and under-designing component.

Further research is also necessary to validate the proposed methodology in field conditions under more variable load conditions. Alternative sensing technologies could also be considered to replace strain gauges, since strain gauges have a limited lifetime and a labour-intensive installation procedure. Candidates include fiber-optic strain gauges, bolt-on strain gauges, and MEMS accelerometers. An extension towards measuring non-torque loads (e.g. bending moment on the main shaft caused by asymmetric aerodynamic load on the blades) with a similar technique is also being investigated.

ACKNOWLEDGMENTS

This research is supported by the Flanders Innovation & Entrepreneurship Agency within the V6LoadS project in cooperation with the Center for Wind Power Drives and ZF Wind Power. The Center for Wind Power Drives is acknowledged for providing the measurement data considered in this paper, in the context of the joint V6LoadS project. Furthermore, the Internal Funds KU Leuven are gratefully acknowledged for their support.

Conflict of interest

The authors declare no potential conflict of interests.

References

1. IRENA . Renewable Power Generation Costs in 2021. tech. rep., IRENA; Abu Dhabi: 2022. https://www.irena.org/Publications/2022/Jul/Renewable-Power-Generation-Costs-in-2021, Accessed: 2023-04-16.

2. IEA . Renewables 2022. tech. rep., IEA; Paris: 2022. https://www.iea.org/reports/renewables-2022, Accessed: 2023-04-16.

3. Nejad AR, Keller J, Guo Y, et al. Wind turbine drivetrains: State-of-the-art technologies and future development trends. *Wind Energy Science* 2022; 7(1): 387–411. doi: 10.5194/wes-7-387-2022

4. Peeters J, Schmidt S, Verellen M. Empowering a sustainable future with continuous innovation in wind turbine powertrains. In: Desmet W, Pluymers B, Moens D, Neeckx S., eds. *Proceedings of ISMA2022 International Conference on Noise and Vibration Engineering* . KU Leuven; 2022; Leuven: 27–39.

5. Liserre M, Cardenas R, Molinas M, Rodriguez J. Overview of Multi-MW Wind Turbines and Wind Parks. *IEEE Transactions on Industrial Electronics* 2011; 58(4): 1081–1095. doi: 10.1109/TIE.2010.2103910

6. Turnbull A, McKinnon C, Carrol J, McDonald A. On the Development of Offshore Wind Turbine Technology: An Assessment of Reliability Rates and Fault Detection Methods in a Changing Market. *Energies* 2022; 15(9). doi: 10.3390/en15093180

7. Carroll J, McDonald A, McMillan D. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. *Wind Energy* 2016; 19: 1107–1119. doi: 10.1002/we.1887

8. HBM . T12HT - Very High-Capacity Torque Sensor. https://www.hbm.com/en/6337/t12ht-high-capacity-torque-sensor; . Accessed: 2022-10-17.

9. Bilen E, Azzam B, Schelenz R, Runkel T, Raddatz M, Jacobs G. From Strain to Loads: Development of a Measurement Solution for Wind Turbine Transmission Input Loads during Drivetrain Testing. *Sensors* 2023; 23(4): 1824. doi: 10.3390/s23041824

10. Croes J. *Virtual sensing in mechatronic systems State estimation using system level models*. PhD thesis. KU Leuven, ; 2017.

11. Lourens E, Reynders E, De Roeck G, Degrande G, Lombaert G. An augmented Kalman filter for force identification in structural dynamics. *Mechanical Systems and Signal Processing* 2012; 27: 446–460. doi: 10.1016/j.ymssp.2011.09.025

12. Petersen W, Øiseth O, Nord TS, Lourens E. Estimation of the full-field dynamic response of a floating bridge using Kalman-type filtering algorithms. *Mechanical Systems and Signal Processing* 2018; 107: 12–28. doi: 10.1016/j.ymssp.2018.01.022

13. Forrier B, Naets F, Desmet W. Broadband Load Torque Estimation in Mechatronic Powertrains Using Nonlinear Kalman Filtering. *IEEE Transactions on Industrial Electronics* 2018; 65(3): 2378–2387. doi: 10.1109/TIE.2017.2739709

14. Risaliti E, Tamarozzi T, Vermaut M, Cornelis B, Desmet W. Multibody model based estimation of multiple loads and strain field on a vehicle suspension system. *Mechanical Systems and Signal Processing* 2019; 123: 1–25. doi: 10.1016/j.ymssp.2018.12.024

15. Perišić N, Kirkegaard PH, Pedersen BJ. Cost-effective shaft torque observer for condition monitoring of wind turbines. *Wind Energy* 2013; 18: n/a–n/a. doi: 10.1002/we.1678

16. Branlard E, Giardina D, S. D. Brown C. Augmented Kalman filter with a reduced mechanical model to estimate tower loads on a land-based wind turbine: A step towards digital-Twin simulations. *Wind Energy Science* 2020; 5(3): 1155–1167. doi: 10.5194/wes-5-1155-2020

17. Cappelle C, Cattebeke M, Bosmans J, Kirchner M, Croes J, Desmet W. Sensor selection for cost-effective virtual torque measurements on a wind turbine gearbox. *Forschung im Ingenieurwesen/Engineering Research* 2021; 85(2): 325–334. doi: 10.1007/s10010-021-00464-z

18. Mehlan FC, Nejad AR, Gao Z. Digital twin based virtual sensor for online fatigue damage monitoring in offshore wind turbine drivetrains. *Journal of Offshore Mechanics and Arctic Engineering* 2022; 144(December): 1–9. doi: 10.1115/1.4055551

19. Bosmans J, Kirchner M, Croes J, Desmet W. Validation of a wind turbine gearbox strain simulation model in service to virtual sensing. *Forschung im Ingenieurwesen* 2023; 87(1): 107–117. doi: 10.1007/s10010-023-00635-0

20. CWD . 4 MW-WT-system test rig for On-Shore WEA. https://www.cwd.rwth-aachen.de/cms/CWD/Das-Center/Infrastruktur/~mqjnf/4-MW-Systempruefstand/lidx/1/; . Accessed: 2023-06-29.

21. Averous NR, Stieneker M, Kock S, et al. Development of a 4 MW Full-Size Wind-Turbine Test Bench. *IEEE Journal of Emerging and Selected Topics in Power Electronics* 2017; 5(2): 600–609. doi: 10.1109/JESTPE.2017.2667399

22. Gnauert J, Krause A, Jacobs G, Bosse D. Development of a Methodology for a WT Gearbox Robustness Test. *Journal of Physics: Conference Series* 2022; 2362(1): 012014. doi: 10.1088/1742-6596/2362/1/012014

23. Azzam B, Schelenz R, Cardaun M, Jacobs G. From Simulations to Accelerated Testing: Design of Experiments for Accelerated Load Testing of a Wind Turbine Drivetrain Based on Aeroelastic Multibody Simulation Data. *Applied Sciences* 2023; 13(1): 356. doi: 10.3390/app13010356

24. Noll MC, Godfrey JW, Schelenz R, Jacobs G. Analysis of time-domain signals of piezoelectric strain sensors on slow spinning planetary gearboxes. *Mechanical Systems and Signal Processing* 2016; 72-73: 727–744. doi: 10.1016/j.ymssp.2015.10.028

25. Peeters J. *Simulation of dynamic drive train loads in a wind turbine*. PhD thesis. KU Leuven, ; 2006.

26. Bosmans J, Gallas S, Rocha de Melo Filho NG, Kirchner M, Desmet W. Validation of wind turbine gearbox structural FE models through modal analysis at component and (sub-)assembly level. In: Desmet W, Pluymes B, Moens D, Neeckx S., eds. *Proceedings of ISMA2022 International Conference on Noise and Vibration Engineering* . KU Leuven; 2022; Leuven: 4551–4565.

27. Simon D. *Optimal state estimation: Kalman, H [infinity] and nonlinear approaches*. Hoboken, N.J: Wiley-Interscience . 2006.

28. Hansen PC. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve. *SIAM Review* 1992; 34(4): pp. 561–580. doi: 10.1137/1034115

How to cite this article: J. Bosmans, S. Gallas, V. Smeets, M. Kirchner L. Geens J. Croes, and W. Desnet (2023), Experimental validation of virtual torque sensing for wind turbine gearboxes based on strain measurements, *Wind Energy*, .