Journal of Sound and Vibration 612 (2025) 119153

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

# A sub-structuring approach to overcome model limitations for input-state estimation of offshore wind turbines

Harry A. Simpson [a,*], Eleni N. Chatzi [b], Manolis N. Chatzis [a]

[a] Department of Engineering Science, University of Oxford, Parks Road, OX1 3PJ, Oxford, UK
[b] Department of Civil, Environmental and Geomatic Engineering, ETH Zürich, Stefano-Franscini-Platz 5, 8093, Zürich, Switzerland

ARTICLE INFO

*Keywords:*
Virtual sensing
Input-state estimation
Offshore wind turbines
Modelling challenges
Sub-structuring

ABSTRACT

The Augmented Kalman Filter (AKF) has been applied previously for input-state estimation of offshore wind turbines (OWT). However, the accuracy of the estimated results depend on the chosen model, for which various complexities exist, making this a challenging task. Two of which are the lack of information required to model the Rotor-Nacelle Assembly (RNA), and the high uncertainty associated with the soil–structure-interaction (SSI). Therefore, the primary focus of this work is to avoid these limitations by considering a suitable substructure which eliminates the need to model the RNA and the SSI, thus significantly reducing uncertainties. The substructure is obtained by 'cutting' the OWT at the top of the tower and at the ground level. To define the model, the resulting substructure then only requires geometries and material properties for the monopile and tower; information which is often known with greater certainty. A numerical case study is presented to investigate the accuracy of the proposed approach for input-state estimation of a 15 MW OWT. A series of commonly used setups involving accelerometers and inclinometers are used and the effects on the predicted fatigue life of the structure are discussed. Additionally, a simple approximation of the wave loading is considered to estimate and account for its contribution to the dynamics of the substructure. The proposed approach is shown to be an effective solution for input-state estimation of OWTs when the RNA or SSI are unknown or associated with significant uncertainty.

## 1. Introduction

The increased urgency to reduce carbon emissions and maximise clean energy production has led many nations to turn to offshore wind energy as a crucial player in their clean energy strategies. For example, the UK has set an ambitious target of reaching 50 GW of installed offshore wind by 2030 [1]. However, the operational availability for offshore wind turbines can be considerably lower than their onshore counterparts [2], primarily due to location based challenges. As the dependence on offshore wind energy grows, addressing these operational challenges becomes increasingly important.

To help mitigate the risk of downtime, measured data from installed sensors and structural health monitoring (SHM) techniques can be utilised to provide ongoing assessment of an offshore wind turbine's condition. A SHM technique, which can provide insight into the condition of a structure, beyond the location of typically sparsely deployed sensors, is virtual sensing [3]. Here, the idea is to make use of observed data, complemented with a system model, which supplies sufficient information for estimating the response at unmeasured locations [4,5]. This is particularly useful for offshore structures, since installing sensors at submerged or

* Corresponding author.
*E-mail addresses:* henry.simpson@eng.ox.ac.uk (H.A. Simpson), chatzi@ibk.baug.ethz.ch (E.N. Chatzi), manolis.chatzis@eng.ox.ac.uk (M.N. Chatzis).

https://doi.org/10.1016/j.jsv.2025.119153
Received 18 December 2024; Received in revised form 4 April 2025; Accepted 23 April 2025
Available online 5 May 2025

foundation-embedded locations can be both costly and lead to poor data quality. Therefore, virtual sensing provides a means to estimate the response at these challenging locations based on a limited set of measurements.

Various approaches have been explored within the domain of virtual sensing, which can be broadly categorised as deterministic and probabilistic. Amongst the deterministic approaches, modal decomposition and expansion (MD&E) is one of the most popular schemes [6]. In this approach, the response at unmeasured locations is estimated based on a system model and a decomposition of the measured data into the contribution of each mode. Examples of MD&E being applied to offshore structures include the work by Song et al. [7], where it has been used to predict the strain response of an offshore platform and the works by Henkel [8] for estimating the subsoil and submerged response of offshore wind turbines. Additionally, Nabiyan et al. [9] successfully applied an MD&E approach to accurately estimate the moment time history at the mudline of a 2 MW offshore wind turbine.

However, since MD&E is a deterministic approach, it does not consider any measurement or modelling uncertainty. This can be a significant limitation for offshore wind turbine structures, since there is often major modelling uncertainty due to the necessary simplifying assumptions. To overcome this limitation, focus has shifted towards probabilistic approaches. One such option is found in Bayesian schemes, where the model and measurement uncertainties are incorporated into the framework. If the target is to achieve online estimation (real-time or near real-time), such schemes are often implemented in variants of the Kalman filter [10]. In this framework, the prior knowledge of a state is predicted in a time update step, before computing the innovation (discrepancy) between measurements and estimated quantities, to estimate the posterior of the entire state vector. Alternative Bayesian estimators have also been proposed. For example, Ebrahimian et al. [11] introduced a Bayesian inference-based framework for output-only system identification of civil structures, which can estimate unknown model parameters simultaneously with the unmeasured response.

The classical Kalman filter has been widely used for state estimation in various disciplines. However, for large scale engineering structures, the inputs which excite the system are often unknown. These unknown inputs can be accounted for within the measurement and modelling uncertainty terms, which are often (although not necessarily) assumed to be Gaussian noise sources but are not estimated. Recent works have tackled the theoretical problem of estimating the observability of systems under the presence of unmeasured inputs [12,13], for deducing sensor setups that would allow estimating not only the states and parameters of the systems but also the unmeasured inputs or the corresponding Lie symmetries [14]. Since such knowledge can be valuable, the targeted estimation of inputs can been dealt with under joint input-state estimation frameworks. For this task, many variations of the classical Kalman filter have been developed, a selection of them include the Augmented Kalman Filter (AKF) [15], the Dual Kalman Filter (DKF) [16] and the Gilligns De Moor Filter (GDF) [17,18]. Additionally, in [19], a scheme is introduced which implements modal expansion in an augmented Kalman Filter (ME-AKF).

Each approach has its own advantages and the suitability often depends on the application at hand. For example, the GDF offers the advantage of not placing any assumptions on the input(s), since the filter satisfies the conditions of minimum variance and unbiasedness. On the other hand, the AKF and DKF often assume a random walk model for the to-be-estimated inputs. Nevertheless, in the literature, the random walk model has been shown to perform well. For example, Tatsis and Lourens [20] compared the DKF and GDF in a numerical investigation for the state estimation of an offshore wind turbine supported by a jacket foundation. They observed similar performance between the DKF and the GDF, but conclude that the GDF slightly outperforms the DKF due to the DKF's need for a regularisation parameter.

One common challenge across many of these filtering methods is the need to tune the process and measurement noise covariance matrices. This tuning process can significantly influence the filter's performance, particularly when estimating a large number of inputs. A few methodologies have been developed to address this issue. For example the L-curve [21] has been used in many works for virtual sensing [20,22,23], which aids the user by providing information about the regularisation properties. However, the L-curve becomes more complicated when multiple parameters require tuning. Moreover, in practice it has been shown to produce curves which do not exhibit a clear L shape, rendering the selection of optimal parameter values difficult [15,24]. This shortcoming led Vettori et al. [25] to develop the adaptive-AKF, in which the process noise is adaptively tuned based on a set of optimal parameters for a given batch of data. This approach also has the advantage of tuning in an online manner, in contrast to the L-curve, which has to be carried out offline in a preliminary step. While these methods can help to alleviate the challenges associated with tuning, they still have their own limitations. As a result, a trial-and-error approach to tuning can often achieve satisfactory results in less time.

In the context of wind turbines, Tatsis et al. [26] applied the AKF for estimating the response of an onshore wind turbine, which they use to provide a probabilistic estimate of accumulated fatigue damage. Branlard et al. [27] also used the AKF to estimate tower loads on an onshore wind turbine from simulated data. Maes et al. [28,29] compared a Kalman filter, GDF, and MD&E methods to estimate the dynamic strain of a real world 3 MW offshore wind turbine with a monopile foundation. They conclude that in general, the three methods perform competitively for this task. The AKF has also been applied for estimating the response of an experimental wind turbine blade by Vettori et al. [25]. Mehrjoo et al. [30] introduced a Bayesian method to recursively estimate the wind load on an offshore wind turbine, validating it both numerically and experimentally.

The literature highlights the potential benefits that virtual sensing can have on the wind industry. However, its performance can strongly depend on the accuracy of the system model, emphasising the importance of employing a model that accurately captures the underlying physics of the structure of interest. This task is particularly challenging for offshore wind turbines due to various inherent complexities. One such complexity exists due to the lack of detailed modelling information provided by manufacturers regarding the rotor-nacelle assembly (RNA). This typically leads to the use of simplified models, where the RNA is considered as a lumped mass applied to the top of the tower. For example, the works by the OWI-lab [6,31,32] applied this approach for fatigue and input estimation on a real world offshore wind turbine. However, these simplified models risk introducing errors as they do not consider the flexibility of the blades, thus failing to capture the coupled tower-blade dynamics.

Tatsis et al. [22] presented a solution to this challenge by developing a general substructuring framework for input-state estimation. Their work includes a numerical case study using an onshore wind turbine, where a 'cut' is applied at the top of the tower and the AKF is used to estimate the tower substructure's response. The advantage of this approach is that the RNA does not need to be modelled, thus requiring no design details from the manufacturer. However, it does introduce the task of estimating the interface inputs instead of directly modelling the RNA.

For offshore wind turbines, a second complexity stems from the uncertainty surrounding the soil–structure interaction (SSI). For turbines supported by monopile foundations, the PISA design method [33] is the current state-of-the-art for modelling SSI. In this method, the SSI is represented by four soil reaction components: a distributed lateral load, a distributed moment, a lateral base force and a base moment. Compared to traditional p-y design methods [34], originally developed for the oil and gas industry with slender monopiles, the PISA method has been shown to perform better for offshore wind turbines.

Although the PISA method offers a more suitable model for the offshore wind industry, the modelling of SSI still remains inherently challenging, often requiring the use of large safety factors. Furthermore, additional research is needed to address the PISA design method for cyclic loading [35] and to accommodate layered soil profiles [36]. Moreover, offshore site investigations are both challenging and costly [37], making it impractical to obtain accurate and reliable soil profiles for every turbine within a wind farm. These factors collectively contribute to significant uncertainty regarding SSI models, inevitably resulting in large modelling errors and potentially leading to inaccurate response estimates.

This paper will aim to address the aforementioned challenges by employing a substructuring approach for input-state estimation that alleviates the need to model the soil–structure interaction and the RNA. To obtain an appropriate model, the structure is 'cut' at two locations: one at the top of the tower and another at the ground level, as illustrated in Fig. 1. For situations where wave loading is present, the unmeasured wave excitations on the tower are also considered as inputs to be identified. The resulting substructure comprises the non-embedded portion of the monopile and the tower. However, for simplicity, it will be referred to as the 'tower substructure' in this work. Since the considered substructure requires fewer modelling assumptions, it is expected to be associated with significantly reduced model uncertainty, thus offering an advantage over modelling the entire wind turbine system.

To the best of our knowledge, this approach has not been previously attempted. However, it is worth noting that while virtual sensing was successfully applied to the monopile-tower substructure of a 3.3 MW offshore wind turbine in [38], this method requires either a numerical model for the soil–structure interaction or measurements at submerged locations—both of which are assumed to be unavailable in this work.

This paper is structured as follows. First, in Section 2, the mathematical model for a substructure is presented, followed by an overview of the AKF algorithm in Section 3. In Section 4, the nonlinear offshore wind turbine model used to generate synthetic data is described. Section 5 then presents two numerical case studies involving a 15 MW offshore wind turbine with a monopile foundation to evaluate the proposed approach. The first considers a case where only wind loading is applied, and the second considers a case study with combined wind and wave loading. Finally, conclusions are drawn in Section 6.

## 2. Mathematical model for a substructure

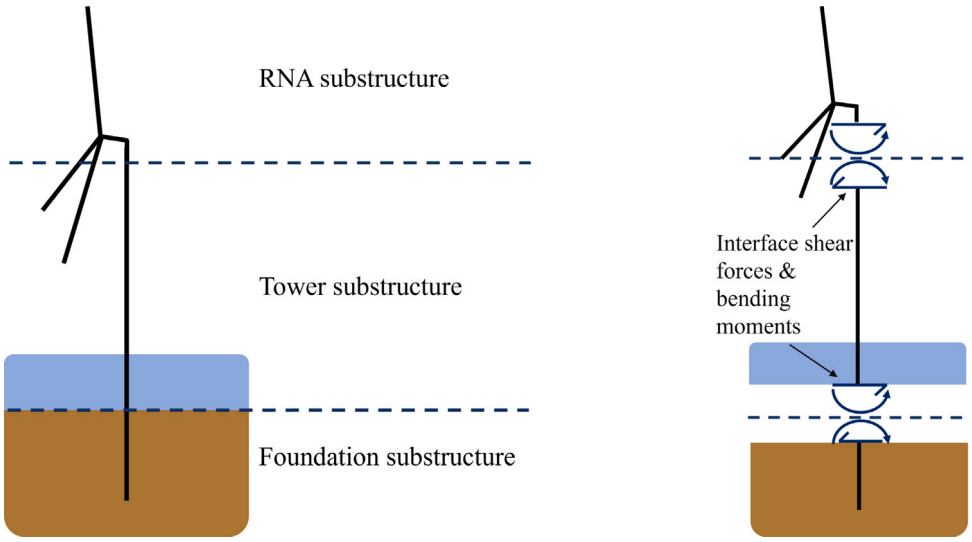
Offshore wind turbines are complex dynamical systems, with inputs arising from aerodynamics, hydrodynamics, and the controller that govern their response. Despite this complexity, the system can be divided into various substructures to alleviate specific challenges. For example, Fig. 1(a) shows an offshore wind turbine separated into three substructures: the RNA substructure, the tower substructure and the foundation substructure. By focussing only on the tower substructure, beam-type finite elements can be used, which require no detailed knowledge of the rest of the system. However, the coupled interaction with the adjacent substructures must be taken into account in the analysis to accurately capture the overall system's behaviour. This section outlines the mathematical framework for developing a reduced order model for the tower substructure. The presented theory can be applied generally to other systems when substructuring is of interest. Furthermore, since this substructure is unconstrained, rigid body modes are present, and a method for handling these modes is also described in this section.

The second order equation of motion for a linear time invariant system with two substructures, denoted with superscripts $\alpha$ and $\beta$, can be represented in the following form:

$$\begin{bmatrix}\mathbf{M}^{\alpha} & \mathbf{0}\\ \mathbf{0} & \mathbf{M}^{\beta}\end{bmatrix}\begin{bmatrix}\ddot{\mathbf{u}}^{\alpha}(t)\\ \ddot{\mathbf{u}}^{\beta}(t)\end{bmatrix}+\begin{bmatrix}\mathbf{C}^{\alpha} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}^{\beta}\end{bmatrix}\begin{bmatrix}\dot{\mathbf{u}}^{\alpha}(t)\\ \dot{\mathbf{u}}^{\beta}(t)\end{bmatrix}+\begin{bmatrix}\mathbf{K}^{\alpha} & \mathbf{0}\\ \mathbf{0} & \mathbf{K}^{\beta}\end{bmatrix}\begin{bmatrix}\mathbf{u}^{\alpha}(t)\\ \mathbf{u}^{\beta}(t)\end{bmatrix}=\begin{bmatrix}\mathbf{S}_{p}^{\alpha} & \mathbf{S}_{p}^{\beta}\end{bmatrix}\begin{bmatrix}\mathbf{p}^{\alpha}(t)\\ \mathbf{p}^{\beta}(t)\end{bmatrix}, \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are the mass, damping and stiffness matrices respectively and $\mathbf{u}(t)$ is the displacement vector at time $t$. The inputs applied to the system are denoted by $\mathbf{p}(t)$ and $\mathbf{S}_p$ is a selection matrix, which maps the input signal(s) to the degrees of freedom (DOFs) at which they act. Eq. (1) represents the uncoupled equation of motion, where each substructure contains the DOFs relevant to that substructure, including the shared DOFs at the interface between adjacent substructures. Note that in the form presented here, $\mathbf{p}(t)$ contains both the externally applied loads and the internal forces. The interface forces are required so that compatibility between the two components is ensured [39].

To relate the notation within the context of this study, $\alpha$ may denote the primary substructure of interest (e.g. the tower substructure, as shown in Fig. 1), while $\beta$ represents the substructure(s) that are to be disregarded (e.g. the RNA and foundation substructures, as shown in Fig. 1). Since only substructure $\alpha$ is being modelled, the superscripts $\alpha$ and $\beta$ in Eq. (1) will be omitted going forward to simplify the notation.




(a) Offshore wind turbine substructures (b) Individual substructures and interface forces.

**Fig. 1.** Overview of substructuring method for an offshore wind turbine. The structure is 'cut' at the dashed lines, which introduces three substructures: the Rotor-Nacelle Assembly (RNA), the tower, and the foundation. This figure provides a simplified 2D representation, whereas the model used in this study is 3D.

### 2.1. Reduced order model

Eq. (1) presents the equation of motion in physical coordinates. In this form, the model can be associated with a large number of system states, making input-state estimation a challenging task. However, in practice, the system's response is often dominated by only a limited number of modes. Therefore, it is often advantageous to apply model reduction techniques to approximate the original full-order model. While several methods exist for model reduction [40], this work uses a free-interface approach from the class of component mode synthesis techniques [41,42], which allows for the reduction to be applied at the substructure level.

To apply this reduction, a transformation is introduced to convert Eq. (1) from physical coordinates to modal coordinates:

$$\mathbf{u}(t) = \mathbf{\Phi}\mathbf{z}(t), \tag{2}$$

where $\mathbf{z}(t) \in \mathbb{R}^{n_m \times 1}$ is the displacement vector in modal coordinates and $\mathbf{\Phi} \in \mathbb{R}^{n_{dof} \times n_m}$ is the reduction basis. $n_m$ and $n_{dof}$ are the number of modes in the reduction basis and the number of degrees of freedom for the considered substructure, respectively.

After substituting Eq. (2) into Eq. (1), the following equation of motion in modal coordinates can be obtained:

$$\tilde{\mathbf{M}}\ddot{\mathbf{z}}(t) + \tilde{\mathbf{C}}\dot{\mathbf{z}}(t) + \tilde{\mathbf{K}}\mathbf{z}(t) = \mathbf{\Phi}^T\mathbf{S}_p\mathbf{p}(t), \tag{3}$$

where the reduced order system matrices are computed as: $\tilde{\mathbf{M}} = \mathbf{\Phi}^T\mathbf{M}\mathbf{\Phi}$, $\tilde{\mathbf{C}} = \mathbf{\Phi}^T\mathbf{C}\mathbf{\Phi}$ and $\tilde{\mathbf{K}} = \mathbf{\Phi}^T\mathbf{K}\mathbf{\Phi}$.

#### 2.1.1. Reduction basis: $\mathbf{\Phi}$

In this work, the reduction basis, $\mathbf{\Phi}$, consists of the retained free-interface normal modes, $\boldsymbol{\phi}_n \in \mathbb{R}^{n_{dof} \times n_r}$, the attachment modes, $\boldsymbol{\phi}_a \in \mathbb{R}^{n_{dof} \times n_a}$, and rigid body modes, $\boldsymbol{\phi}_{rb} \in \mathbb{R}^{n_{dof} \times n_{rb}}$, where $n_r$, $n_a$, and $n_{rb}$ are the number of retained normal modes, attachment modes and rigid body modes, respectively. Consequently, the size of the reduction basis can be calculated as $n_m = n_r + n_a + n_{rb}$ and:

$$\mathbf{\Phi} = \begin{bmatrix} \boldsymbol{\phi}_{rb} & \boldsymbol{\phi}_n & \boldsymbol{\phi}_a \end{bmatrix}. \tag{4}$$

The retained free-interface normal modes, $\boldsymbol{\phi}_n$, attachment modes, $\boldsymbol{\phi}_a$, and rigid body modes, $\boldsymbol{\phi}_{rb}$ are defined as follows.

The rigid body modes, $\boldsymbol{\phi}_{rb}$, represent the substructure's motion without deformation [41], and their number correspond to the unconstrained DOFs of the substructure. In free-interface methods, the normal modes are obtained by placing no constraints at the interface DOFs. These modes capture the internal dynamics of the substructure, and by retaining only those that correspond to the dominant frequency range, the model size can be reduced without compromising accuracy [40]. To compute $\boldsymbol{\phi}_{rb}$ and $\boldsymbol{\phi}_n$, a free-free model of the substructure is first assembled, followed by an eigenvalue analysis:

$$\mathbf{K}\mathbf{\Psi} = \mathbf{M}\mathbf{\Psi}\mathbf{\Sigma}, \tag{5}$$

where $\mathbf{\Psi}$ is a matrix whose columns correspond to the eigenvectors and $\mathbf{\Sigma}$ is a diagonal matrix containing the associated eigenvalues. The matrix $\mathbf{\Psi}$ includes both the rigid body modes and all free-interface normal modes. Therefore, $\boldsymbol{\phi}_{rb}$ is defined as the eigenvectors associated with zero eigenvalues and $\boldsymbol{\phi}_n$ comprises the retained subset of remaining eigenvectors.

The attachment modes, $\boldsymbol{\phi}_a$, are defined as the static displacement response to an imposed unit force at an interface DOF, with zero force applied elsewhere. These are obtained by solving the static equilibrium equation:

$$\mathbf{K}\boldsymbol{\phi}_a = \begin{bmatrix}\mathbf{0}\\ \mathbf{I}_s\end{bmatrix}, \tag{6}$$

where $\mathbf{I}_s \in \mathbb{R}^{n_a \times n_a}$ is an identity matrix corresponding to the interface DOFs, and the zero matrix is of size $(n_{dof} - n_a) \times n_a$ to ensure compatibility.

These modes complement $\boldsymbol{\phi}_{rb}$ and $\boldsymbol{\phi}_n$ by capturing the static displacement induced by unit interface forces between coupled substructures [42,43]. By including attachment modes in the reduction basis, the reduced-order model preserves the effect of interface forces, ensuring an accurate representation of the dynamic coupling between adjacent substructures.

To obtain $\boldsymbol{\phi}_a$, the stiffness matrix $\mathbf{K}$ needs to be inverted and therefore must be non-singular. However, in the case of unconstrained components, such as the tower substructure considered here, $\mathbf{K}$ will be a singular matrix and cannot be inverted.

To overcome this difficulty, the Residual Inertia-Relief Attachment (RIRA) modes $\boldsymbol{\phi}_{RIRA}$ can be used instead. A brief overview of their computation is presented next. However, for further information, the interested reader is referred to the review of substructuring methods in [44] or the paper by Vettori et al. [45], in which the following methodology is applied to estimate the response of a box assembly with a removable component.

1. Obtain the rigid body modes $\boldsymbol{\phi}_{rb}$ and the retained free-interface normal modes $\boldsymbol{\phi}_n$;
2. Calculate the reduced order stiffness matrix: $\boldsymbol{\Lambda} = \boldsymbol{\phi}_n^T \mathbf{K} \boldsymbol{\phi}_n$;
3. Calculate the inertia-relief projector: $\mathbf{P}_r = \mathbf{I} - \mathbf{M}\boldsymbol{\phi}_{rb}\boldsymbol{\phi}_{rb}^T$;
4. Arbitrarily choose $n_a$ DOFs to constrain, so that $\mathbf{K}$ becomes non-singular. Then compute the Inertia-Relief Attachment modes $\boldsymbol{\phi}_{IRA}$:

$$\mathbf{K}\boldsymbol{\phi}_{IRA} = \mathbf{P}_r \begin{bmatrix}\mathbf{0}\\ \mathbf{I}_s\end{bmatrix} \tag{7}$$

5. Finally, calculate $\boldsymbol{\phi}_{RIRA}$:

$$\boldsymbol{\phi}_{RIRA} = \mathbf{P}_r^T \boldsymbol{\phi}_{IRA} - \boldsymbol{\phi}_n \boldsymbol{\Lambda}^{-1} \boldsymbol{\phi}_n^T \begin{bmatrix}\mathbf{0}\\ \mathbf{I}_s\end{bmatrix} \tag{8}$$

### 2.2. State space model

A state space model can be used to replace the second order differential equation of motion (Eq. (3)) with 2 first order differential equations, so that the mathematics are compatible with the Augmented Kalman Filter. This results in a system described by two equations:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{p}(t) \tag{9}$$

$$\mathbf{y}(t) = \mathbf{G}\mathbf{x}(t) + \mathbf{J}\mathbf{p}(t). \tag{10}$$

Eq. (9) is the state equation, with state vector $\mathbf{x}(t) \in \mathbb{R}^{n_s \times 1}$ and the number of states $n_s$. $\mathbf{A} \in \mathbb{R}^{n_s \times n_s}$ is known as the state matrix and $\mathbf{B} \in \mathbb{R}^{n_s \times n_p}$ is the input matrix, with $n_p$ being the number of inputs to the system. After defining the state vector as $\mathbf{x}(t) = \begin{bmatrix}\mathbf{z}(t) & \dot{\mathbf{z}}(t)\end{bmatrix}^T$, Eq. (3) can be rearranged so that $\mathbf{A}$ and $\mathbf{B}$ take the following form:

$$\mathbf{A} = \begin{bmatrix}\mathbf{0} & \mathbf{I}\\ -\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{K}} & -\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{C}}\end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix}\mathbf{0}\\ \tilde{\mathbf{M}}^{-1}\boldsymbol{\Phi}^T\mathbf{S}_p\end{bmatrix}. \tag{11}$$

Eq. (10) is known as the output equation, where $\mathbf{y}(t) \in \mathbb{R}^{n_o \times 1}$ is the measurement vector and $n_o$ is the number of measured quantities:

$$\mathbf{y}(t) = \begin{bmatrix}\mathbf{S}_d & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}_v & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{S}_a\end{bmatrix}\begin{bmatrix}\mathbf{u}(t)\\ \dot{\mathbf{u}}(t)\\ \ddot{\mathbf{u}}(t)\end{bmatrix}. \tag{12}$$

$\mathbf{S}_d$, $\mathbf{S}_v$ and $\mathbf{S}_a$ are selection matrices for the displacements, velocities and accelerations respectively, and are used to map the physical degrees of freedom of the system to the measured degrees of freedom. $\mathbf{G} \in \mathbb{R}^{n_o \times n_s}$ is the output matrix and $\mathbf{J} \in \mathbb{R}^{n_o \times n_p}$ is the feedthrough matrix. After substituting Eq. (12) and rearranging Eq. (3), it can be shown that $\mathbf{G}$ and $\mathbf{J}$ take the following form:

$$\mathbf{G} = \begin{bmatrix}\mathbf{S}_d\boldsymbol{\Phi} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}_v\boldsymbol{\Phi}\\ -\mathbf{S}_a\boldsymbol{\Phi}\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{K}} & -\mathbf{S}_a\boldsymbol{\Phi}\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{C}}\end{bmatrix}, \quad \mathbf{J} = \begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{S}_a\boldsymbol{\Phi}\tilde{\mathbf{M}}^{-1}\boldsymbol{\Phi}^T\mathbf{S}_p\end{bmatrix}. \tag{13}$$

Eqs. (9) & (10) represent the continuous time state space model. However, because the measurements are typically only available at discrete times, the continuous system needs to be transformed into its equivalent discrete-time representation:

$$\mathbf{x}_k = \mathbf{A}_d\mathbf{x}_{k-1} + \mathbf{B}_d\mathbf{p}_{k-1} \tag{14}$$

$$\mathbf{y}_k = \mathbf{G}\mathbf{x}_k + \mathbf{J}\mathbf{p}_k, \tag{15}$$

where $\mathbf{A}_d$ and $\mathbf{B}_d$ are the discrete versions of $\mathbf{A}$ and $\mathbf{B}$, respectively. $k$ denotes the current time step, such that $t = k/f_s$ for a sampling frequency of $f_s$. If a zero-order hold scheme is used for the discretisation, then $\mathbf{A}_d$ and $\mathbf{B}_d$ can be computed as:

$$\mathbf{A}_d = \exp\left(\mathbf{A}/f_s\right) \quad \text{and} \quad \mathbf{B}_d = \mathbf{A}^{-1}\left[\mathbf{A}_d - \mathbf{I}\right]\mathbf{B}. \tag{16}$$

Note that exp in this equation denotes the matrix exponential.

## 3. Augmented Kalman filter for input-state estimation

The Augmented Kalman Filter (AKF) [15] is a recursive Bayesian filter that allows for the joint estimation of inputs and states. In the AKF, the unknown inputs $\mathbf{p}_k$ are included in an augmented state vector $\mathbf{x}_k^a$:

$$\mathbf{x}_k^a = \begin{bmatrix}\mathbf{x}_k \\ \mathbf{p}_k\end{bmatrix}. \tag{17}$$

Having defined Eq. (17), Eqs. (14)–(15) can be developed to take the following stochastic form:

$$\dot{\mathbf{x}}_k^a = \mathbf{A}_d^a \mathbf{x}_{k-1}^a + \mathbf{w}_{k-1} \tag{18}$$

$$\mathbf{y}_k = \mathbf{G}^a \mathbf{x}_k^a + \mathbf{v}_k, \tag{19}$$

where $\mathbf{w}_k$ and $\mathbf{v}_k$ are independent zero-mean random variables. They represent the process noise and measurement noise, respectively, with normal probability distributions:

$$p(\mathbf{w}) \sim \mathcal{N}(0, \mathbf{Q}^a) \tag{20}$$

$$p(\mathbf{v}) \sim \mathcal{N}(0, \mathbf{R}), \tag{21}$$

where $\mathbf{R} \in \mathbb{R}^{n_o \times n_o}$ is the measurement covariance matrix and $\mathbf{Q}^a \in \mathbb{R}^{(n_s+n_p)\times(n_s+n_p)}$ is the augmented process noise covariance matrix. The augmented state matrix $\mathbf{A}_d^a$ and output matrix $\mathbf{G}^a$ then become:

$$\mathbf{A}_d^a = \begin{bmatrix}\mathbf{A}_d & \mathbf{B}_d \\ \mathbf{0} & \mathbf{I}\end{bmatrix}, \quad \mathbf{G}^a = \begin{bmatrix}\mathbf{G} & \mathbf{J}\end{bmatrix}. \tag{22}$$

It should be noted that Eq. (22) assumes a discrete-time random walk model for the unknown dynamics of the inputs, where:

$$\mathbf{p}_k = \mathbf{p}_{k-1} + \mathbf{w}_{k-1}^{\mathbf{p}}, \tag{23}$$

with superscript $^{\mathbf{p}}$ denoting the components of $\mathbf{w}_{k-1}$ which relate to the inputs. While this random walk assumption was found to be sufficient in this work, it could be replaced by a Gaussian Process Latent Force Model (GPLFM), which allows for more advanced modelling of the input dynamics [46].

$\mathbf{R}$ reflects the uncertainty in the measurements and is often assumed to be a diagonal matrix, where each diagonal element represents the variance in the corresponding sensor. As a starting point, this is often defined based on the known uncertainty provided by the sensor specification sheet.

$\mathbf{Q}^a$ reflects the uncertainty in both the unknown states and inputs. It is usually assumed that the uncertainty in the states and inputs are independent, so that the following form can be adopted:

$$\mathbf{Q}^a = \begin{bmatrix}\mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}\end{bmatrix}, \tag{24}$$

where $\mathbf{Q} \in \mathbb{R}^{n_s \times n_s}$ is the covariance matrix for the unknown states and $\mathbf{S} \in \mathbb{R}^{n_p \times n_p}$ is the covariance matrix for unknown inputs. In contrast with $\mathbf{R}$, there is typically very little information available about $\mathbf{Q}$ and $\mathbf{S}$. As a result, these parameters tend to require tuning when using the AKF for input-state estimation. Although more sophisticated methods, discussed in Section 1, could be applied for this task, a trial and error method was chosen in this work. This decision was made because satisfactory results were achieved in a short time frame, making the more complex tuning methods unnecessary.

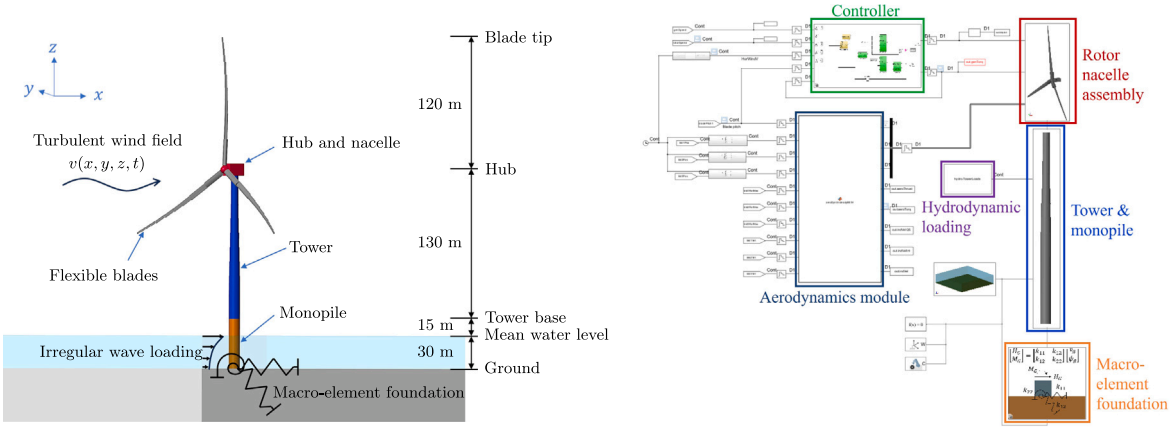
With the augmented state space model described as above, the AKF algorithm (see Alg. 1) can be used to recursively estimate the mean augmented state vector $\hat{\mathbf{x}}_k^a$ and its associated covariance matrix $\mathbf{P}_k$. This algorithm iteratively updates $\hat{\mathbf{x}}_k^a$ and $\mathbf{P}_k$ in two stages: a time update and a measurement update. During the time update stage, the process model (Eq. (14)) is employed to predict the state vector mean and covariance matrix for the current time step, based on the updated values from the previous time step. These predicted values form the *a priori* estimate of the state vector and covariance matrix, which are then passed to the measurement update stage. The measurement update stage utilises the output equation (Eq. (15)) to compute the *a posteriori* estimate of the state vector and covariance matrix. In summary, the time update stage provides a prior belief for the unknown states, which are then updated using the measurements, to obtain a posterior belief.

**Algorithm 1** AKF

**Given:** $\mathbf{A}_d^a,\ \mathbf{G}^a,\ \mathbf{Q}^a,\ \mathbf{R}$
**Initialise:** $\hat{\mathbf{x}}_0^a,\ \mathbf{P}_0$
**for** $k = 1, 2, \ldots$ **do**
(1) Time Update:
$\hat{\mathbf{x}}_k^{a-} = \mathbf{A}_d^a \hat{\mathbf{x}}_{k-1}^a$ ▷ Predict state vector
$\mathbf{P}_k^- = \mathbf{A}_d^a \mathbf{P}_{k-1} \mathbf{A}_a^{aT} + \mathbf{Q}^a$ ▷ Predict state covariance matrix

(2) Measurement Update:
$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{G}^{aT} (\mathbf{G}^a \mathbf{P}_k^- \mathbf{G}^{aT} + \mathbf{R})^{-1}$ ▷ Compute Kalman gain $\mathbf{K}_k$
$\hat{\mathbf{x}}_k^a = \hat{\mathbf{x}}_k^{a-} + \mathbf{K}_k (\mathbf{y}_k - \mathbf{G}^a \hat{\mathbf{x}}_k^{a-})$ ▷ Update state vector
$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{G}^a) \mathbf{P}_k^{-1}$ ▷ Update state covariance matrix
**end for**

**Result:** Augmented state mean $\hat{\mathbf{x}}^a$ and covariance $\mathbf{P}$




(a) 3D OWT Forward Simulator (b) Simulink user interface

**Fig. 2.** Overview of the coupled servo-hydro-aero-elastic Simulink model used for generating synthetic data.

## 4. Forward simulator

This section provides an overview of the numerical model used to generate synthetic data for the case studies presented in Section 5. To ensure a robust numerical investigation, realistic measurements that capture the coupled interaction between the foundation, tower, and flexible rotor-nacelle assembly are essential. To achieve this, a 3D coupled servo-hydro-aero-elastic model of the entire wind turbine system was developed in the Matlab/Simulink environment [47,48]. This model accounts for the full dynamics of the system, incorporating hydrodynamics, aerodynamics and controller dynamics to generate synthetic data that accurately reflects the complex behaviour of the wind turbine structure.

The wind turbine used to generate realistic measurements is based on NREL's 15 MW reference offshore wind turbine with a monopile foundation [49]. The main features of this model are depicted in Fig. 2, along with the Simulink user interface.

The tower and monopile are modelled using flexible cylindrical Euler–Bernoulli beam elements, made from steel with an elastic modulus of 210 GPa and Poisson's ratio of 0.3. The diameter and wall thickness distributions for both the tower and monopile can be found in the reference document [49]. The monopile extends 15 m above the mean water level (MWL), which is set at 30 m, with a material density of 7850 $\mathrm{kg/m^3}$. The tower sits on top of the monopile and extends to a height of 145 m above the mean water level, with a material density of 8500 $\mathrm{kg/m^3}$.

To model the soil–structure interaction, a macro-stiffness element at the ground level was used. This type of foundation comprises three coupled linear springs: a lateral spring $k_{11}$, a rotational spring $k_{22}$, and a cross-coupling spring $k_{12}$. These springs relate the ground level lateral force $H_{gl}$ and moment $M_{gl}$, to the ground level displacement $x_{gl}$ and rotation $\theta_{gl}$:

$$\begin{bmatrix} H_{gl} \\ M_{gl} \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{bmatrix} \begin{bmatrix} x_{gl} \\ \theta_{gl} \end{bmatrix}. \tag{25}$$

This SSI model replaces the embedded foundation model with statically equivalent coupled linear springs. It assumes that the mass and inertia of the embedded foundation are negligible and that the torsional and axial responses can be ignored. The stiffness

**Table 1**
List of user-defined inputs required to model the tower substructure.

| Model input | Value/reference for 15 MW wind turbine |
|---|---|
| Diameter profile | Provided in the reference document [49] |
| Wall thickness profile | Provided in the reference document [49] |
| Young's modulus | 210 GPa |
| Poisson's ratio | 0.3 |
| Material density | 8500 kg/m$^3$ for tower elements & 7850 kg/m$^3$ for monopile elements |

parameters $k_{11}$, $k_{22}$, and $k_{12}$, were calculated by applying the PISA design method to obtain a depthwise stiffness matrix for a monopile in Cowden clay [33], with an embedded length of 22.5 m and a diameter of 10 m. Once the depthwise stiffness matrix was obtained, the ground-level macro-element stiffness matrix, $\boldsymbol{K}_0$, was calculated as the set of statically equivalent springs:

$$\boldsymbol{K}_0 = \begin{bmatrix} 0.0532 & 0.3224 \\ 0.3224 & 5.1368 \end{bmatrix} \times 10^{11}, \tag{26}$$

where the units of $k_{11}$, $k_{22}$ and $k_{12}$ are N/m, N m/rad and N m/m, respectively. As the soil conditions in this study are homogeneous, and due to the symmetry of the monopile, the macro-element stiffness matrix was assumed to be axisymmetric about the vertical axis.

For the RNA, the hub and nacelle were modelled as lumped masses with inertia. However, the blades were modelled using flexible beam elements. Based on the blade geometry, mass, and stiffness distributions defined in [49], 1D Euler–Bernoulli beam theory was used to assemble stiffness and mass matrices for each blade element. To reduce the computational complexity, Craig–Bampton reduction [50] was then applied to calculate the reduced order mass and stiffness matrices. Each blade was divided into 10 reduced order elements, each containing 2 external nodes and retaining 2 internal modes. Using multiple reduced-order elements for each blade allows the model to capture potential nonlinear deflections in the blades.

It is worth emphasising that the SSI and RNA modelling presented here is solely for generating data for the numerical case studies in Section 5. For the reverse problem, a separate model—introduced later—is used, which only requires the tower and monopile to be modelled.

In Section 5, two case studies are considered: one with wind loading only, and a second with both wind and wave loading. To generate data with wind loading, the aerodynamic forces acting on the blades are computed based on unsteady blade element momentum theory [51]. The calculation of these forces requires a 3-D velocity wind field covering the rotor-plane, which was generated using Turbsim [52] in this work.

For the case study with wind and wave loading, the Morison equation [53] was used to calculate the forces on submerged elements. Linear wave theory was used to generate unidirectional irregular wave kinematics, based on a JONSWAP spectrum [54].

The hyperparameters for calculating the turbulent wind and irregular wave loading are defined later, when introducing each case study. However, both case studies consider the turbine in an operational state, where the wind turbine controller regulates the generator torque and blade pitch angle. Therefore, to account for the dynamics introduced by the controller, the Simulink implementation of the ROSCO controller [55] was incorporated into the model.

## 5. Numerical case studies

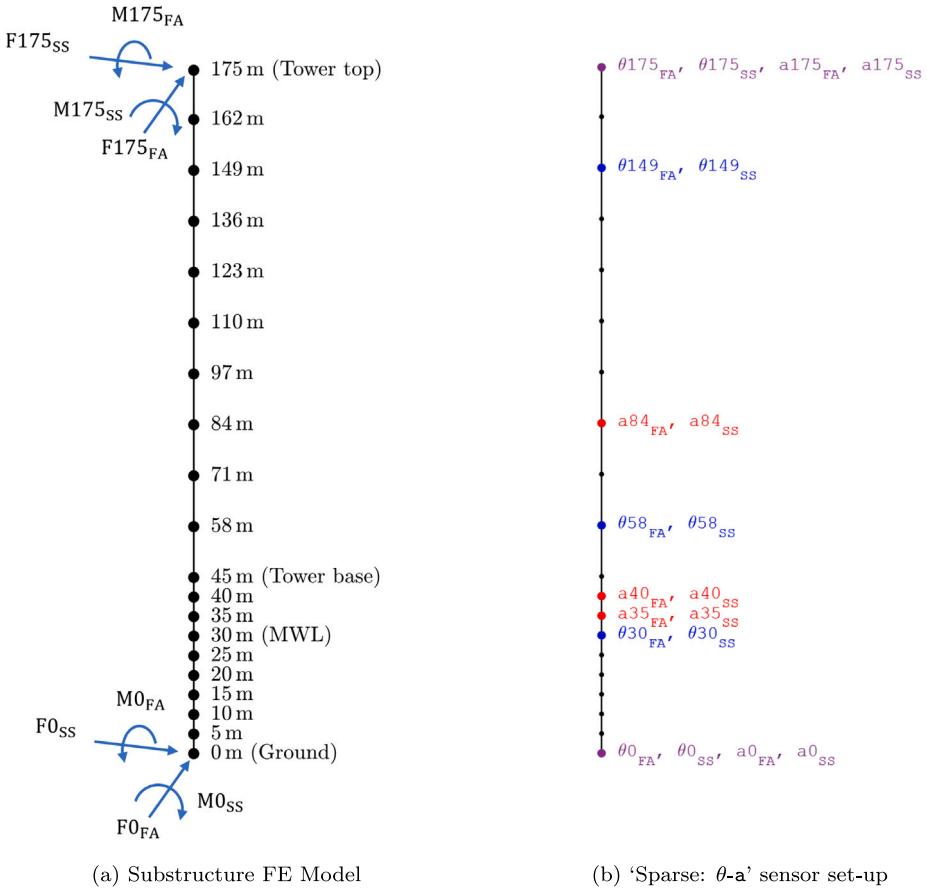
In this section, numerical case studies are presented to assess the accuracy of the proposed substructuring approach on a 15 MW offshore wind turbine with a monopile foundation. First, a description is provided of the linear finite element model of the tower substructure used for the reverse problem. Results are then presented for two case studies; one with wind loading only and a second with both wind and wave loading.

### 5.1. Linear finite element model

To provide separation between the forward and reverse problems, a different model to that presented in Section 4 was used for input-state estimation. For this model, 1D Euler–Bernoulli beam elements were used to assemble FE mass and stiffness matrices for the tower substructure, which was discretised into 19 element using the mesh shown in Fig. 3(a). The geometric and material properties are provided in Table 1. Note that these values are the same as those used for the tower and monopile substructure in the Simulink model to generate data.

Since this substructure is obtained by applying a cut just below the nacelle and another at the ground level, the user is not required to make assumptions about the RNA or soil–structure interaction. To introduce structural damping into the model, a proportional damping approach was employed [56], with a 0.8% damping ratio for the first mode. The damping matrix was then computed in proportion to the mass and stiffness matrices. It was also assumed that the response in the fore–aft (FA) and side–side (SS) directions were dominant, so that the axial and torsional degrees of freedom could be neglected.

Once the mass **M**, stiffness **K**, and damping **C** matrices were assembled, the approach outlined in Section 2.1 was applied to obtain a reduced order model. In the case studies considered here, a reduced order model with 4 retained normal modes was used, corresponding to the first 2 bending modes in the FA and SS directions. This resulted in a reduction basis comprising 16 modes; 4 rigid body modes, 4 retained normal modes and 8 attachment modes. The rigid body modes and retained free-interface normal modes are graphically presented in Fig. 4. Each of the 8 attachment modes correspond to one of the interface forces to be estimated, which include shear forces and bending moments at both the ground level and tower top, as shown in Fig. 3(a).




(a) Substructure FE Model (b) 'Sparse: $\theta$-a' sensor set-up

**Fig. 3.** (a) Linear finite element model with interface forces & moments used for input-state estimation, (b) 'Sparse: $\theta$-a' sensor set-up.

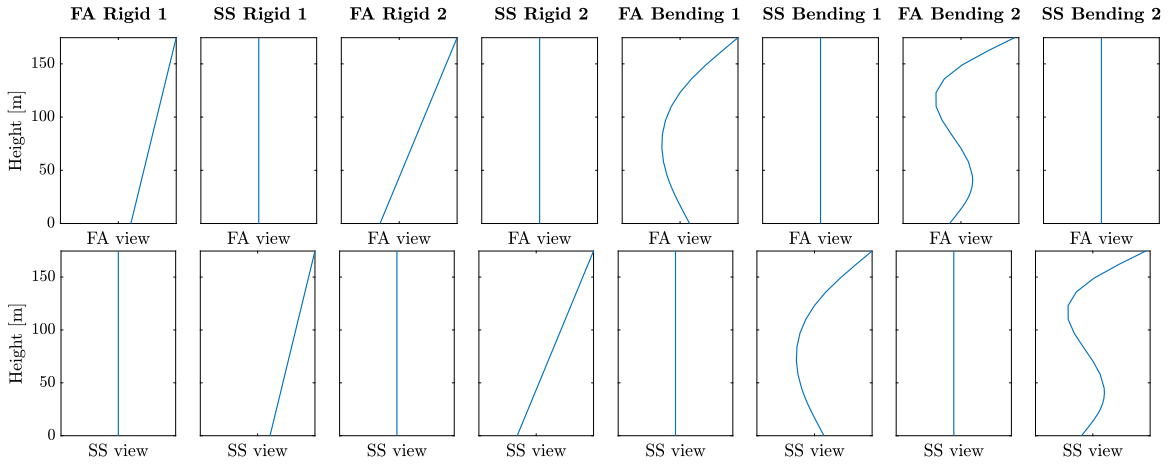



**Fig. 4.** Graphical representation of the rigid body modes and retained normal modes for the tower substructure. Each column represents an individual mode, with the top and bottom rows displaying the projections in the fore–aft (FA) and side–side (SS) directions, respectively.

### 5.1.1. Sensor set-up

Three different sensor setups were considered for the numerical case studies and are summarised in Table 2. In this paper, unless stated otherwise, all inputs, states and measurements are denoted using the notation $\alpha\beta_\gamma$, where $\alpha$ refers to the signal type (F: shear force, M: bending moment, a: acceleration, x: displacement, $\theta$: rotation), $\beta$ refers to the signal's location in metres above ground level and $\gamma$ denotes the signal's direction (FA: fore–aft, SS: side–side).

**Table 2**
Summary of sensor configurations considered in the numerical case studies. Note that at each sensor location, the response in both the fore–aft and side–side directions are used in the measurement set.

| Sensor set-up label | Number/Type of sensors | Sensor Locations |
|---|---|---|
| Dense: x-a | 40 displacement sensors at:<br>40 accelerometers at: | Every node (see Fig. 3(a))<br>Every node (see Fig. 3(a)) |
| Dense: $\theta$-a | 40 inclinometers at:<br>40 accelerometers at: | Every node (see Fig. 3(a))<br>Every node (see Fig. 3(a)) |
| Sparse: $\theta$-a | 10 inclinometers at:<br>10 accelerometers at: | 0 m, 30 m, 58 m, 149 m, 175 m (see Fig. 3(b))<br>0 m, 35 m, 40 m, 84 m, 175 m (see Fig. 3(b)) |

Two of the considered set-ups assume that sensors are installed at every node in the FE model. These sensor configurations are labelled 'Dense: x-a' and 'Dense: $\theta$-a' and consist of either accelerometers and displacements or accelerometers and inclinometers, respectively. Displacement sensors are mostly here used as a reference; in reality, displacement information is rare and costly to extract. To the contrary, acceleration measurements, usually delivered by means of MEMs technologies at low costs are the common option. In addition, tilt-meter (inclinometer) measurements can also be made available via such low cost technologies, delivering a quantity that is linked to deformation, and not its derivatives (velocity or acceleration). The latter is highly useful for improving filter estimates in virtual sensing schemes.

Although it is unlikely to be provided with such a heavily instrumented structure in practice, these dense sensor configurations were initially considered so that a comparison could be made between displacement and inclinometer measurements, without the influence of sensor placement affecting the results.

A more realistic sensor configuration is further considered and is referred to as the 'Sparse: $\theta$-a' set-up. For this sensor set-up, only inclinometers and accelerometers are considered, since these types of sensors are more likely to be installed in practice. Moreover, it was desired to keep the number of sensors to a minimum. Following the guidance in [57], to achieve stable input estimation, the number of displacement sensors (or inclinometers) has to be greater than or equal to the number of inputs. In the case studies considered here, the maximum number of inputs to be estimated is 10 (for Case Study 2). As a result, the sparse set-up includes 10 inclinometers.

To improve the invertibility conditions for input estimation, accelerometers were also included, with their count limited to 10. Furthermore, due to the higher installation cost associated with submerged sensors, an additional restriction was imposed, which only allowed one submerged location in the sensor configuration.

With the above restrictions defined, the 10 accelerometers and 10 inclinometers were distributed as shown in Fig. 3(b) and Table 2. Note that at each sensor location, the response in both the FA and SS directions are included in the measurement set. To replicate practical noise inherent in real sensors, zero mean white noise with standard deviation equal to 1% of the true signal was added to each measurement.

### 5.2. Comparison metrics

In order to evaluate the performance of the proposed approach, the following metrics were used:

1. The Time Response Assurance Criterion ($TRAC$) for two signals in the time domain, $\sigma_1(t)$ and $\sigma_2(t)$:

$$TRAC = \frac{\left[\sigma_1(t)^T \sigma_2(t)\right]^2}{\left[\sigma_1(t)^T \sigma_1(t)\right]\left[\sigma_2^T(t)\sigma_2(t)\right]} \tag{27}$$

2. The Frequency Response Assurance Criterion ($FRAC$) for two signals in the frequency domain, $\sigma_1(f)$ and $\sigma_2(f)$:

$$FRAC = \frac{\left[\sigma_1(f)^T \sigma_2(f)\right]^2}{\left[\sigma_1(f)^T \sigma_1(f)\right]\left[\sigma_2^T(f)\sigma_2(f)\right]} \tag{28}$$

3. The Mean Absolute Error between two signals in the time domain, $\sigma_1(t)$ and $\sigma_2(t)$:

$$MAE = \frac{1}{n_t}\sum_{i}^{n_t} |\sigma_1(t_i) - \sigma_2(t_i)|, \tag{29}$$

where $n_t$ is the number of data points and $i$ indicates the time step.

The $TRAC$ value allows for a comparison of two signals in the time domain and can take a value between 0 and 1, where a value of 1 indicates that two signals are perfectly in phase and value of 0 suggests that two signals are out of phase. Similar to the $TRAC$, the $FRAC$ takes a value between 0 and 1 and provides insight into the similarity of two signals in the frequency domain. In this work, all three metrics are used in combination to evaluate the performance of the estimated results. However, various scenarios exist where these metrics alone can lead to misleading conclusions. Therefore, these criteria were always complimented by a visual inspection of the results presented.

#### 5.2.1. Calculating accumulated fatigue damage

In addition to the metrics described above, the estimated response was also used to calculate the accumulate fatigue damage. In this work, the damage was estimated based on the stress response and the widely used rainflow counting method [58], which determines the number of cycles at corresponding mean stress and stress range levels. Miner's rule was then subsequently used to compute the accumulate fatigue damage, $D$:

$$D = \sum_{i=1}^{k} \frac{n_i}{N_i}, \tag{30}$$

where $n_i$ is the number of cycles in the stress time history at the $i$th stress range and $N_i$ is the number of cycles to failure for the $i$th stress range.

The number of cycles to failure, $N_i$, for a given stress range can be determined using Basquin's law,

$$\log(N_i) = \log(\overline{a}) - m\log(\Delta\sigma_0), \tag{31}$$

where $\Delta\sigma_0$ denotes the stress range for a mean stress of zero, $m$ is the Wöhler exponent, and $\overline{a}$ is a constant that is determined from the intercept of Eq. (31) on the $\log(N)$ axis. In this work, fatigue failure was assumed to occur due to circumferential butt welds failing in the axial direction. The fatigue design standards for offshore steel structures [59] suggest using $m = 5$, and $\log(\overline{a}) = 13.617$ for this type of weld.

For situations when the stress range does not centre around a mean of zero, Goodman's rule is often applied to find the equivalent zero mean stress range, $\Delta\sigma_{0,eq}$:

$$\Delta\sigma_{0,eq} = \Delta\sigma_\sigma \Big/ \left[1 - \frac{\sigma_{nz}}{\sigma_{UTS}}\right]. \tag{32}$$

Here, $\Delta\sigma_\sigma$ is the stress range for the non zero mean stress $\sigma_{nz}$, and $\sigma_{UTS}$ is the ultimate strength of the material under consideration. In this work, $\sigma_{UTS} = 500\,\text{MPa}$, since the substructure is assumed to be composed of S355 steel. As the stresses along the tower are inferred, the corresponding fatigue damage calculation will be shown along the tower and not just for the base and will be compared to the 'actual' fatigue life calculated form the simulations.

### 5.3. Case study 1: wind loading only

To generate synthetic data for this wind-only case study, a 700 s forward simulation was first carried out using the Simulink model. In this simulation, Turbsim was used to generate a turbulent wind field based on a mean wind speed of 8 m/s at hub height and 10% turbulence intensity. The aerodynamic forces calculated based on this wind field predominantly apply loading in the FA direction. However, due to combined effects of the 3D turbulent wind and the aeroelastic interaction, as well as the dynamics introduced by the controller, the substructure also experiences excitation in the SS direction. Considering these loading conditions, the substructure's input vector $\mathbf{p}$ comprises the 8 interface forces and bending moments depicted in Fig. 3(a):

$$\mathbf{p}_1 = \left[\mathtt{FO}_{\mathtt{FA}} \quad \mathtt{MO}_{\mathtt{FA}} \quad \mathtt{FO}_{\mathtt{SS}} \quad \mathtt{MO}_{\mathtt{SS}} \quad \mathtt{F175}_{\mathtt{FA}} \quad \mathtt{M175}_{\mathtt{FA}} \quad \mathtt{F175}_{\mathtt{SS}} \quad \mathtt{M175}_{\mathtt{SS}}\right]^T \tag{33}$$

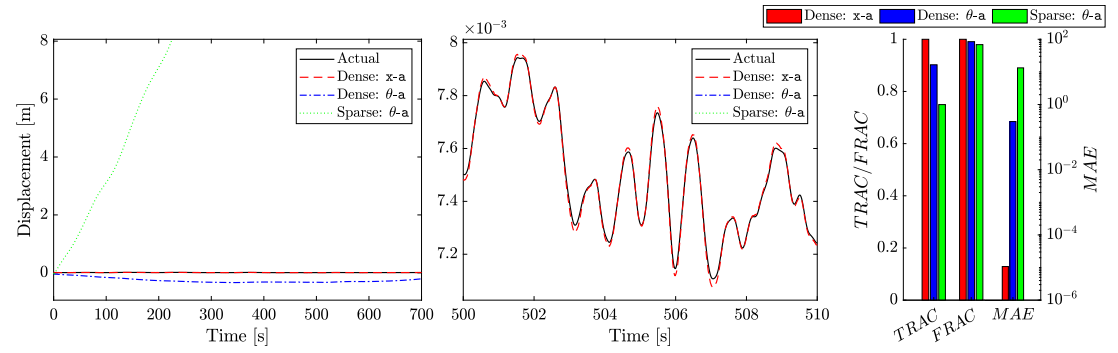
It should be noted that the interface forces and moments at the tower top are due to the combined effects of the external wind pressure, the vibration of the blades, the rotation of the rotor and the actions of the controller.

#### 5.3.1. Case study 1 results — state estimation

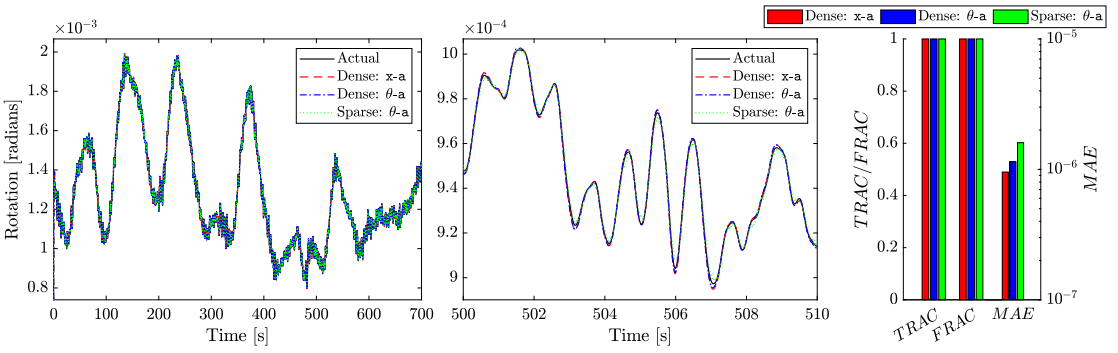
The performance of the proposed approach for state estimation is assessed by comparing the actual response with the AKF estimated response at four representative DOFs. Due to an interest in gaining further insight into the foundation behaviour, the ground level displacement and rotation were included in the set to be analysed. However, as this node is collocated with sensors, the displacements and rotations were also compared at a submerged location where no sensors were installed (15 m). In summary, the selected DOFs for comparison include the displacements and rotations at the ground level (0 m) and halfway to the mean water level (15 m): $\{\mathtt{x0}_{\mathtt{FA}}, \theta\mathtt{0}_{\mathtt{FA}}, \mathtt{x15}_{\mathtt{FA}}, \theta\mathtt{15}_{\mathtt{FA}}\}$.

Fig. 5 compares the AKF results with the actual results for all three sensor set-ups. Focusing first on the results where displacement sensors are used to supplement the accelerometers (labelled 'Dense: x-a'), the figures show that high accuracy has been achieved in the time domain response for the displacements and rotations at both locations. This conclusion is reinforced by the $TRAC$ and $FRAC$ scores being close to 1, indicating that the response has been estimated well in both the time domain and frequency domain.

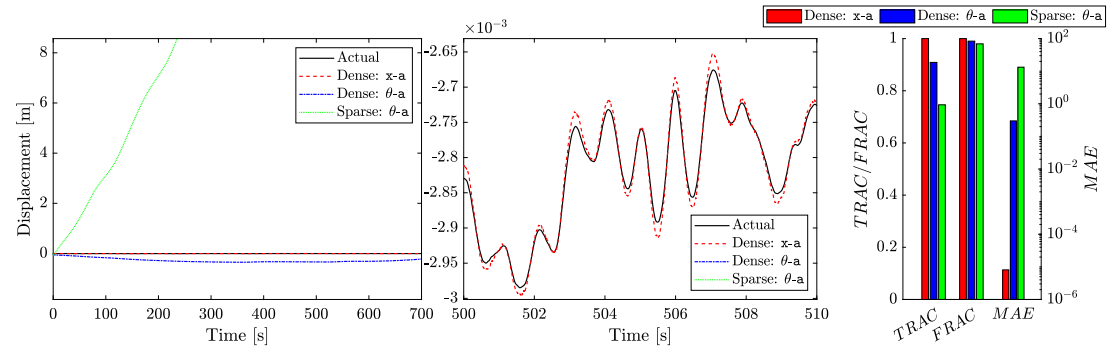
Now considering the results from the set-up with all inclinometers and accelerometers (labelled 'Dense: $\theta$-a'), it can be seen that the rotations were identified with similar accuracy to the case with displacement sensors. However, the time domain response for the displacements show that the AKF significantly diverges from the true response, so much so that the response cannot be seen in the magnified figures. This is an artefact of the unconstrained substructure, since with only measurements from rotations and accelerations, the AKF has no way to penalise the error in the absolute position of the wind turbine, which leads to the observed drift in the figures. If displacement sensor are not used, the displacements of this unconstrained structure would be unobservable. This drift also contributes to a drop in the $TRAC$ values and an increase in the $MAE$s which can be seen in Figs. 5(a) & 5(c). However, these figures also show that the corresponding $FRAC$ values remain high and similar to the values obtained for the sensor set-up with displacements, suggesting that the response was still estimated well in the frequency domain. One possible explanation for this could be the inability to estimate the substructure's rigid body motion. To strengthen this argument, Fig. 6 compares the estimated



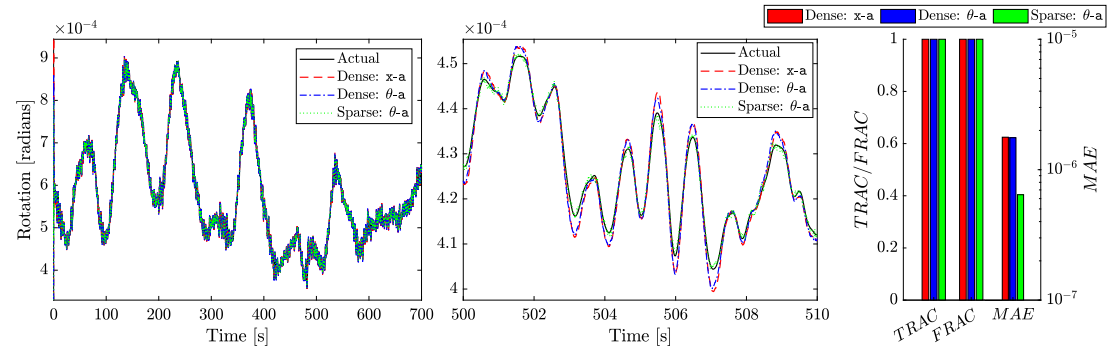

(a) Fore-Aft **displacement** at **15 m** ($\mathtt{x15_{FA}}$)




(b) Fore-Aft **rotation** at **15 m** ($\theta\mathtt{15_{FA}}$)




(c) Fore-Aft **displacement** at **0 m** ($\mathtt{x0_{FA}}$)




(d) Fore-Aft **rotation** at **0 m** ($\theta\mathtt{0_{FA}}$)

**Fig. 5.** Comparison between actual response and AKF estimated response at selected DOFs for the three sensor set-ups (Case Study 1). *From left to right:* entire response duration, magnified excerpt, $TRAC/FRAC/MAE$ values.

**Table 3**
$TRAC$, $FRAC$, and $MAE$ for selected degrees of freedom in the fore–aft and side–side directions.

| Sensor | Metric | State identifier | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Set-up | | $x0_{FA}$ | $x0_{SS}$ | $\theta 0_{FA}$ | $\theta 0_{SS}$ | $x15_{FA}$ | $x15_{SS}$ | $\theta 15_{FA}$ | $\theta 15_{SS}$ |
| Dense: x-a | $TRAC$ | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | $FRAC$ | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | $MAE$ | 8e−6 | 1e−6 | 2e−6 | 2e−7 | 1e−5 | 1e−6 | 1e−6 | 2e−7 |
| Dense: $\theta$-a | $TRAC$ | 0.909 | 0.385 | 1.000 | 1.000 | 0.902 | 0.380 | 1.000 | 1.000 |
| | $FRAC$ | 0.990 | 0.751 | 1.000 | 1.000 | 0.991 | 0.749 | 1.000 | 1.000 |
| | $MAE$ | 0.300 | 5.346 | 2e−6 | 2e−7 | 0.300 | 5.346 | 1e−6 | 3e−7 |
| Sparse: $\theta$-a | $TRAC$ | 0.746 | 0.430 | 1.000 | 1.000 | 0.750 | 0.425 | 1.000 | 1.000 |
| | $FRAC$ | 0.980 | 0.785 | 1.000 | 1.000 | 0.979 | 0.783 | 1.000 | 1.000 |
| | $MAE$ | 13.300 | 27.064 | 6e−7 | 2e−7 | 13.300 | 27.064 | 2e−6 | 5e−7 |

response in modal coordinates for the sensor set-up 'Dense: $\theta$-a'. Figs. 6(a)–6(d) show the estimated and actual rigid body modes, whereas Figs. 6(e)–6(p) show the remaining modes (normal and attachment). These figures show that the normal and attachment modes have been estimated with high accuracy and it is only the rigid body modes that exhibit the same drift that is observed in the displacement response from Figs. 5(a) and 5(c). Therefore, it can be concluded that despite the inability to estimate the absolute displacement when inclinometers are used, the contribution of the response from the internal and attachment modes is still captured with high accuracy, comparable to that achieved with displacement sensors.

Fig. 5 further indicates that using the sparse sensor set-up yields similar performance and observations as the 'Dense: $\theta$-a' sensor set-up. It is worth noting that the displacement DOFs exhibit smaller $TRAC$ values and larger $MAE$s, primarily due to a drift with larger magnitude. However, this does not affect the estimated performance in the frequency domain. Interestingly, based on Fig. 5(d) (particularly observing the $MAE$s), it could be argued that the sparse sensor set-up slightly outperforms the dense sensor set-ups. The minor differences observed between the sparse and dense configurations are due to the different observations available for each set-up. However, these results show that with well-placed sensors ensuring observability of the desired state, a reduced number of sensors can still achieve comparable accuracy to the dense configurations, reinforcing the robustness of the method.

Although Fig. 5 presents only the DOFs in the fore–aft direction, state estimation in the side–side direction showed similar trends, as reflected in the performance metrics in Table 3. Furthermore, the damage distributions introduced in Section 5.3.2 provide additional insight into state estimation performance for all nodes of the substructure, in both the fore–aft and side–side directions.

#### 5.3.2. Case study 1 results — fatigue damage

Once the displacements and rotations have been estimated, they can be used to calculate the damage for the corresponding node in the FE model, using the approach outlined in Section 5.2.1. This was carried out for every node in the FE model to obtain a distribution of the damage along the height of the substructure. According to Miner's rule (Eq. (30)), failure is assumed to occur when the accumulated damage value reaches 1. Therefore, the damage values presented here represent the contribution over a 10 min period towards this reference threshold. However, this threshold is an approximation and does not necessarily indicate actual failure, as it does not account for load sequence effects or other nonlinear damage accumulation mechanisms [60]. Nonetheless, this damage model provides a convenient framework for comparing the predicted and actual fatigue estimates.

Fig. 7 shows the damage distributions for the substructure in both the fore–aft and side–side directions. The fore–aft damage was computed from the stress time history at the outer diameter of the tower's cross-section, aligned with the dominant wind direction at each node. Similarly, the side–side damage was calculated in the perpendicular direction to the dominant wind. As a result, the fore–aft and side–side damage distributions are only influenced by their corresponding bending moments. Since the estimated nodal displacements were required to compute the stresses, the damage distributions shown in Fig. 7 also provide a compact representation for comparing the global performance of the proposed approach for state estimation. This figure shows that, despite the previously discussed drift due to the lack of displacement measurements, the AKF has accurately captured the fatigue distributions along the tower for all three sensor configurations. This is due to the fact that the rigid body motion does not contribute to the accumulated fatigue damage, so only the normal modes and attachment modes need to be estimated well to get a good estimate of the accumulated damage. Therefore, if the end goal is to estimate accumulated fatigue damage, then displacement sensors are not necessary. This is an interesting observation that demonstrates that the occurring lack of observability of the rigid body displacements does not affect the model's fatigue estimates.

Minor differences can be observed between the different sensor set-ups in Fig. 7. For example, between 120 and 150 metres above ground level, the Sparse: $\theta$-a set-up outperforms the dense configurations in the fore–aft direction but shows slightly reduced accuracy in the side–side direction. As discussed in Section 5.3.1, these differences mainly arise because different observations result in varying levels of performance. However, since the normal and attachments modes are observable for all three configurations, the different sensor setups perform comparably for fatigue estimation. This is a case where the use of the additional sensors does not result in significant improvement of the estimated quantities. As expected, even if those additional sensors appear to be redundant in the cases studied, their use does not practically lead to deteriorating the estimated quantities. The additional information they offer could nonetheless be beneficial when dealing with other cases where there would be greater modelling errors and higher noise levels.

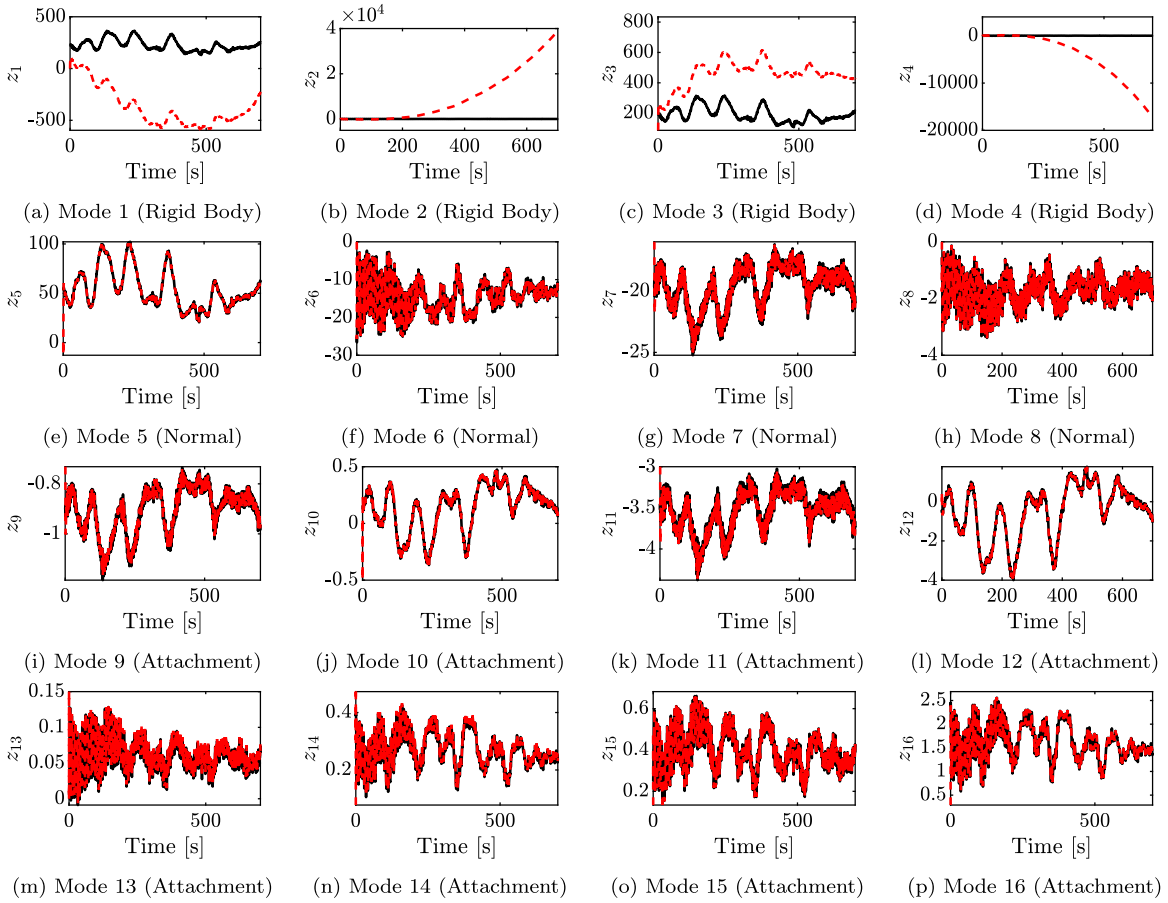



**Fig. 6.** Estimated modal displacements when using 'Dense: $\theta$-a' sensor set-up (red-dashed lines) vs. actual response (solid-black line). Sub-figures (a)–(d): *rigid body modes*, sub-figures (e)–(h): *normal modes*, sub-figures (i)–(p): *attachment modes*.

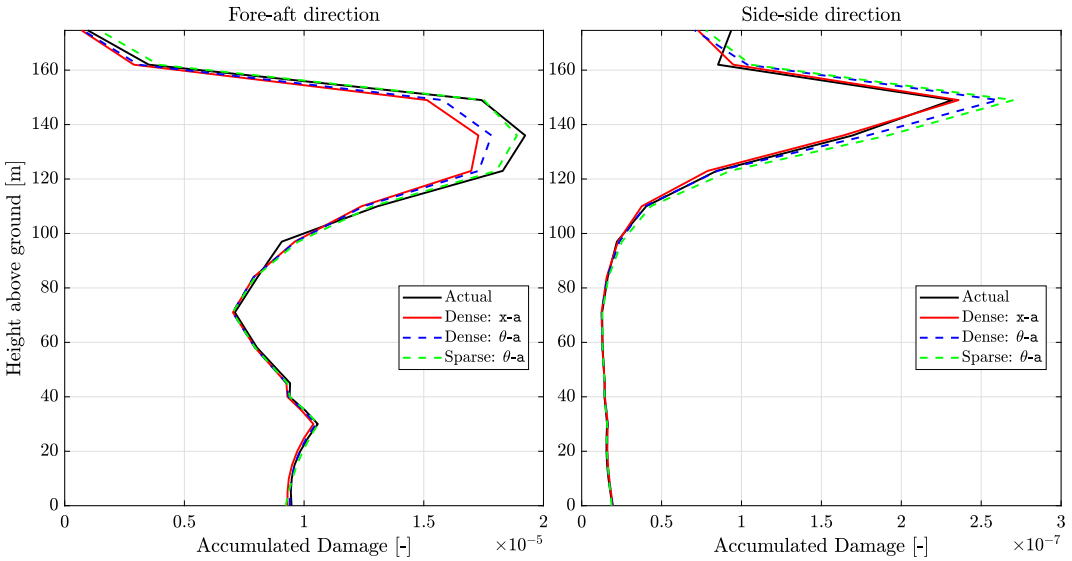



**Fig. 7.** Accumulated damage at each node in the FE model (Left figure: fore–aft damage distribution. Right figure: side–side damage distribution).

#### 5.3.3. Case study 1 results — input estimation

Figs. 8 and 9 compare the estimated inputs at the ground level interface and the tower top interface respectively. Overall, these figures show that the AKF has been able to accurately estimate the main trends in the time domain for all sensor setups, which is further indicated by $TRAC$ values close to 1 for almost all interface inputs. Moreover, the high $FRAC$ values indicate that the response in the frequency domain has also been accurately estimated. Comparing results between sensor sets 'Dense: x-a' and 'Dense: $\theta$-a', it can be seen that the two sensor setups perform competitively. This suggests that for the task of input estimation, it is reasonable to substitute displacement sensors with inclinometers without sacrificing performance. Now considering the results from the 'Sparse: $\theta$-a' setup, the figures show that all inputs have also been estimated with high performance. However, compared with the dense sensor setups, it can be seen that the $MAE$s increase for the ground level inputs (Fig. 8). This is mainly due to a time lag in the estimated inputs (most noticeable in Fig. 8(a)). This time lag is influenced by various factors, but in this case, one significant cause is the number and placement of accelerometers. This affects the terms in the direct feedthrough matrix $\mathbf{J}$ and consequently impacts the input–output coupling. This potentially motivates the use of optimal sensor placement strategies to design sensor set-ups for which this time lag is minimised. However, such an investigation is outside the scope of this paper.

An additional observation that can be made from Figs. 8 & 9, is a DC offset in the estimated response for certain inputs and sensor setups. This is most noticeable for the fore–aft shear force at the tower top interface (see Fig. 9(a)). One explanation for this offset relates to p-delta effects [61]. These effects arise due to the offset between the RNA's centre of mass and the top of the tower, which introduces a non-linear contribution to the lateral response from gravity forces. This non-linearity is captured by the Simulink model used to generate data. However, the linear FE model used for identification in the AKF does not account for these nonlinear p-delta effects. Consequently, the AKF has to compensate for this in the estimated inputs, resulting in the observed DC offset. To verify that this was the dominant cause of the offset, a less realistic forward simulation was run using the Simulink model to obtain measurement data without gravity forces included. Fig. 10(a) presents an excerpt of the estimated fore–aft shear force at the tower top for the dataset without gravity. A comparison between Figs. 10(a) and 9(a), reveals a significant reduction in the offset to a negligible level. Similar observations were made for the remaining 7 inputs, so their time domain responses are not displayed here. However, Fig. 10(b) presents a comparison of the $MAE$s across all inputs between datasets which either include or exclude gravity forces. This figure shows a reduction in the $MAE$ for all estimated inputs when gravity forces are omitted from the simulated data. Therefore, it is reasonable to conclude that the p-delta effects were a significant cause of this offset. It is interesting to note however, that this DC offset in the inputs does not impact the performance for state estimation and fatigue estimates.

### 5.4. Case study 2: wind and wave loading

In the previous case study, still water conditions were assumed, so that no forces resulted from hydrodynamic wave loading. Building upon these results, this case study used simulated data that considered both aerodynamic and hydrodynamic loading. To generate this synthetic data, the same turbulent wind field from Case Study 1 was used, with the wind field aligned in the fore–aft direction. For the hydrodynamics, unidirectional wave loading was applied, which was calculated based on a JONSWAP spectrum defined by a significant wave height of 4.5 m and a peak wave period of 11 s. To explore the impact of misaligned wind and waves, 2 simulations were run: one with a wave yaw angle of 0° and another with a wave yaw angle of 30°. Here, a yaw angle of 0° indicates that the waves are aligned in the fore–aft direction and a yaw angle of 90° indicates that the waves are aligned in the side–side direction.

As a result of the wave loading, forces and moments act on each submerged node in the FE model. Attempting to estimate each of these forces and moments individually would result in an excessive number of inputs to be estimated for the given sensor setups. However, this paper's focus is not on estimating wave inputs but on interface forces and substructure states. Therefore, two approaches were implemented to account for the wave loads, while maintaining high accuracy in the quantities of interest.

In the first approach, the input vector was identical to that of Case Study 1, which included the 8 interface inputs:

$$\mathbf{p}_1 = \begin{bmatrix} \mathrm{FO}_{\mathrm{FA}} & \mathrm{MO}_{\mathrm{FA}} & \mathrm{FO}_{\mathrm{SS}} & \mathrm{MO}_{\mathrm{SS}} & \mathrm{F175}_{\mathrm{FA}} & \mathrm{M175}_{\mathrm{FA}} & \mathrm{F175}_{\mathrm{SS}} & \mathrm{M175}_{\mathrm{SS}} \end{bmatrix}^T. \tag{34}$$

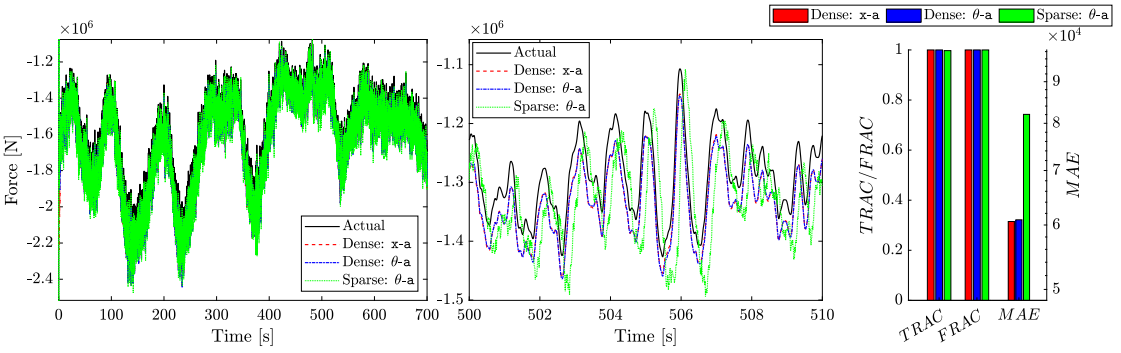
In other words, with this set of inputs, the AKF does not attempt to estimate the wave excitation, which is non-zero for Case Study 2. This approach was included primarily to investigate the necessity of estimating unmeasured wave inputs.

In the second approach, the force applied on each submerged node followed a triangular distribution. Here, the wave force at the ground level was constrained to be 0 and linearly increased to a value of Fw at the mean water level. This approximation is visually represented in Fig. 11.

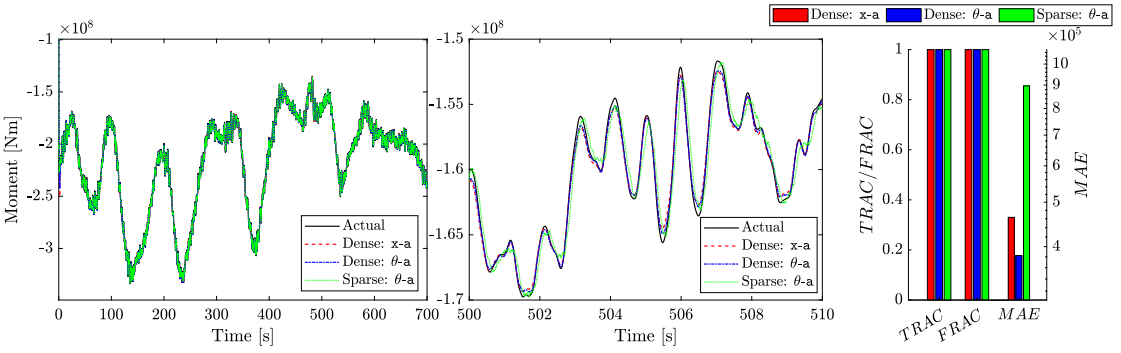
It should be noted that in the simulations the spatial distribution of the wave forces along the tower changes over time as it is frequency dependent. As such the triangular approximation here corresponds to a simplification of the actual wave distribution that generated the data. To account for potential misalignment or spreading of waves, two triangular distributions were estimated: one in the fore–aft direction and one in the side–side direction. As a result, the input vector for the second approach, $\mathbf{p}_2$, included two additional inputs:

$$\mathbf{p}_2 = \begin{bmatrix} \mathrm{FO}_{\mathrm{FA}} & \mathrm{MO}_{\mathrm{FA}} & \mathrm{FO}_{\mathrm{SS}} & \mathrm{MO}_{\mathrm{SS}} & \mathrm{F175}_{\mathrm{FA}} & \mathrm{M175}_{\mathrm{FA}} & \mathrm{F175}_{\mathrm{SS}} & \mathrm{M175}_{\mathrm{SS}} & \mathrm{Fw}_{\mathrm{FA}} & \mathrm{Fw}_{\mathrm{SS}} \end{bmatrix}^T, \tag{35}$$

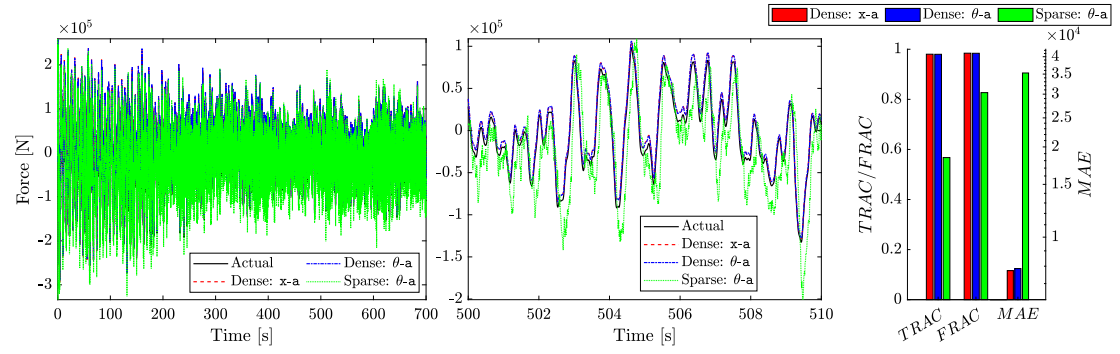
where $\mathrm{Fw}_{\mathrm{FA}}$ and $\mathrm{Fw}_{\mathrm{SS}}$ are the triangular approximation's forces at the MWL in the fore–aft and side–side directions, respectively. It should be noted that the values of the distributions at the top are assumed to be independent for identification purposes, even though in the simulation these were the projections of a wave excitation applied at a yaw angle. In other words, the identification model is not making use of this additional knowledge.



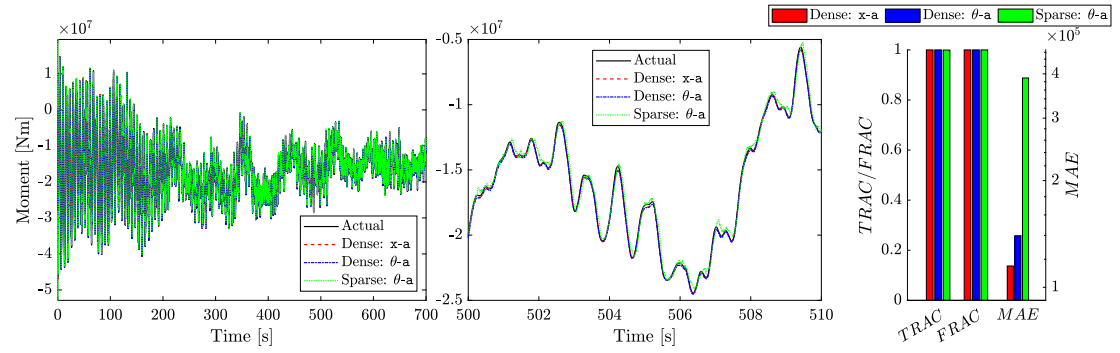

(a) Fore-Aft shear force ($\mathtt{FO}_{\mathtt{FA}}$)




(b) Fore-Aft bending moment ($\mathtt{MO}_{\mathtt{FA}}$)

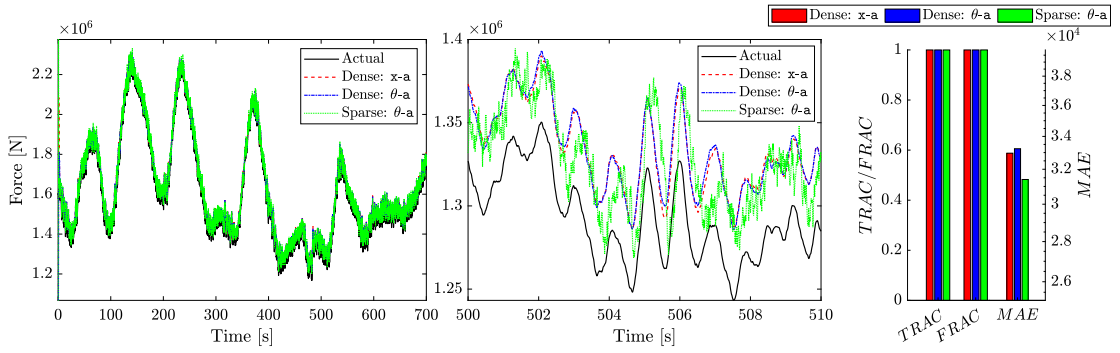



(c) Side-Side shear force ($\mathtt{FO}_{\mathtt{SS}}$)

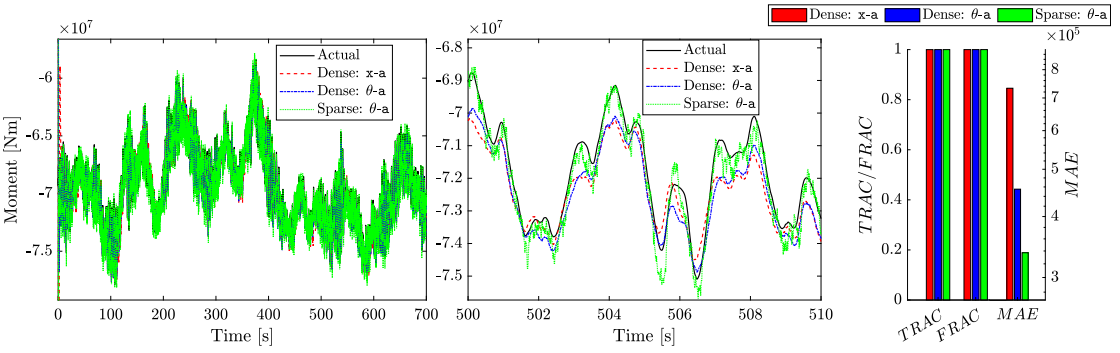



(d) Side-Side bending moment ($\mathtt{MO}_{\mathtt{SS}}$)

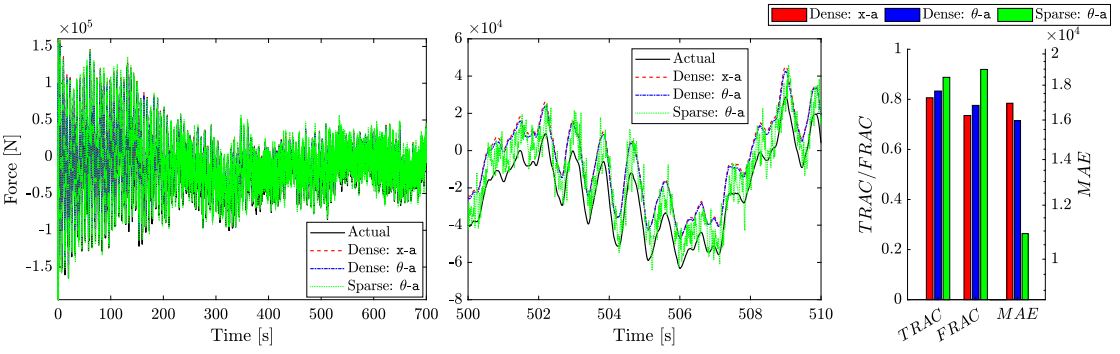
**Fig. 8.** Shear forces and bending moments at the **ground level interface** for the three sensor set-ups (Case Study 1). *From left to right:* entire response duration, magnified excerpt, $TRAC/FRAC/MAE$ values.



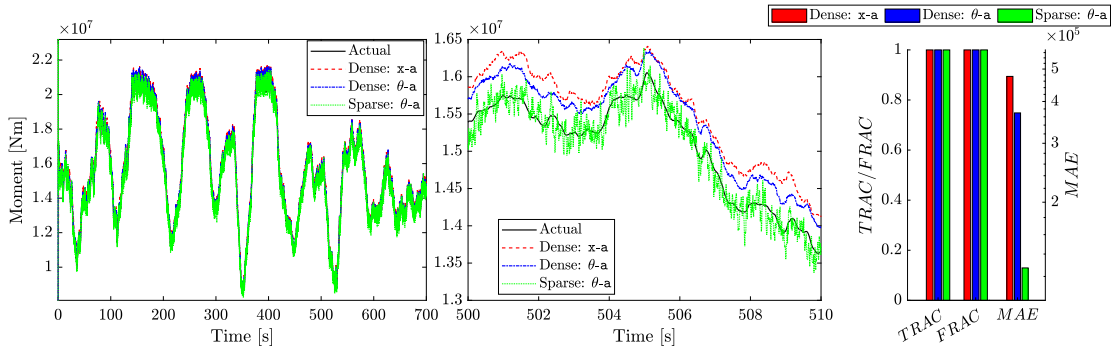

(a) Fore-Aft shear force ($\texttt{F175}_{\texttt{FA}}$)




(b) Fore-Aft bending moment ($\texttt{M175}_{\texttt{FA}}$)

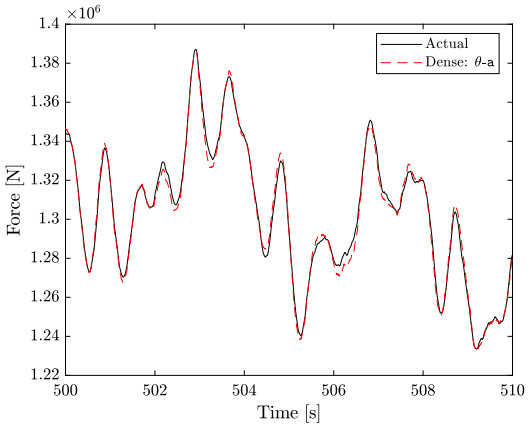



(c) Side-Side shear force ($\texttt{F175}_{\texttt{SS}}$)

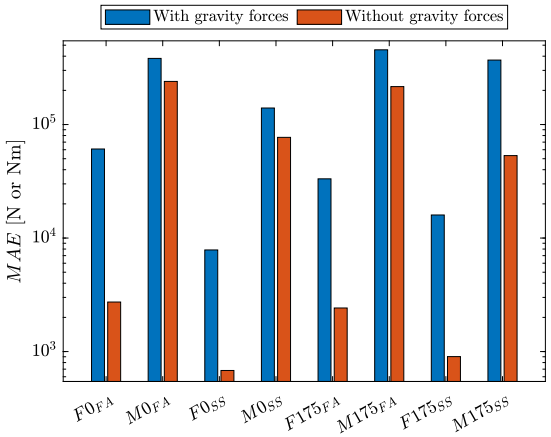



(d) Side-Side bending moment ($\texttt{M175}_{\texttt{SS}}$)

**Fig. 9.** Shear forces and bending moments at the **tower top interface** for the three sensor set-ups (Case Study 1). *From left to right:* entire response duration, magnified excerpt, $TRAC/FRAC/MAE$ values.




(a) Fore-aft shear force at the tower top interface: $\texttt{F175}_{\texttt{FA}}$. Comparison between actual response and AKF estimated response when using the 'Dense: $\theta$-a' sensor setup.




(b) $MAE$s in estimated inputs when the measured data do not include gravity forces, compared with the original results where gravity forces are included. Both sets of results are obtained using the 'Dense: $\theta$-a' sensor setup.

**Fig. 10.** Results for input estimation with synthetic data that excludes gravity forces (Case Study 1).

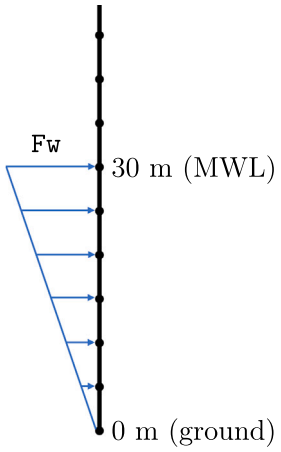



**Fig. 11.** Triangular wave loading approximation.

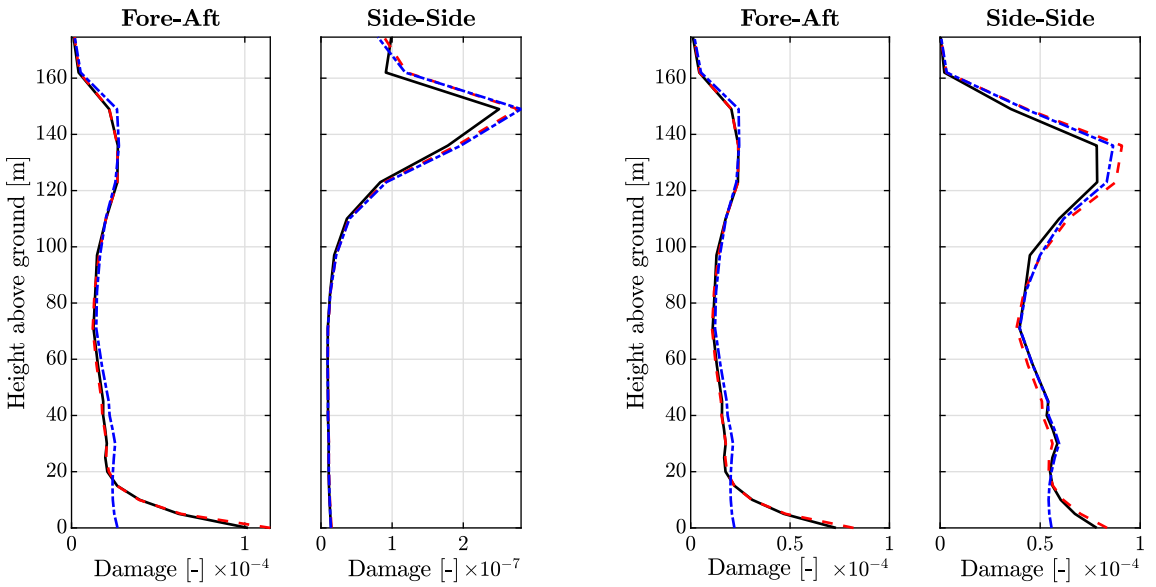
### 5.4.1. Case study 2 results

Regarding the influence of the three sensor set-ups, similar observations and conclusions to those discussed in Case Study 1 were also found here. Therefore, results are only presented for the 'Sparse: $\theta$-a' sensor set-up. First a discussion is made on the damage distributions, before presenting the results from input estimation. A comparison is drawn between the estimated results obtained from using the two input vectors $\mathbf{p}_1$ and $\mathbf{p}_2$. Since the damage distributions provide insight into the performance of state estimation for the entire substructure, the response at individual DOFs is not presented here.

Fig. 12 shows the damage distributions caused by the two excitations studied in Case study 2, corresponding to wave yaw angles of 0° and 30°.

These figures demonstrate that, overall, the fatigue damage has been accurately estimated. However, a drop in performance was observed at nodes closer to the ground level interface when the wave inputs were not estimated and were instead assumed to be zero (i.e. when using $\mathbf{p}_1$), whereas the triangular approximation achieved a level of accuracy comparable to that of Case Study 1 (see Fig. 7 vs. Fig. 12). It is interesting to note that when the waves are aligned in the fore–aft direction, the estimated damage distribution in the side–side direction is similar for both input vectors $\mathbf{p}_1$ and $\mathbf{p}_2$ (see Fig. 12(a)). This was expected due to the symmetry of the substructure, resulting in a reduced order model that is decoupled in the fore–aft and side–side directions. Therefore, when waves are aligned in the fore–aft direction, there is minimal impact on the estimated response in the side–side direction, rendering $\texttt{Fw}_{\texttt{SS}}$ unnecessary. However, when a yaw angle is present, a drop in accuracy was observed when wave inputs were excluded (see Fig. 12(b)), and the inclusion of $\texttt{Fw}_{\texttt{SS}}$ in the estimated quantities was required to restore previous levels of accuracy.

Fig. 13 presents the results for input-estimation at the ground level interface. Here it can be seen that when wave inputs are not included, the inputs are underestimated. Examining the power spectral densities reveal that although the peaks corresponding to




(a) Wave yaw angle: 0° (b) Wave yaw angle: 30°

**Fig. 12.** Case Study 2 damage distributions. (solid black line: actual damage, blue dash-dotted line: damage when no wave inputs were considered, red dashed line: damage when triangular approximation was used).

the first and second tower modes are estimated at the correct frequencies, the magnitude has been underestimated around the first mode, contributing to the errors seen in the time domain. One possible explanation for this is that the interface forces and moments have to compensate for the additional loading, which is present due to the waves but not estimated. Therefore, it can be concluded that for accurate estimation of the interface inputs, it is important to account for the presence of wave loading in the input vector. Moreover, it can be seen that using a simple triangular approximation for this purpose yields highly accurate estimated inputs.

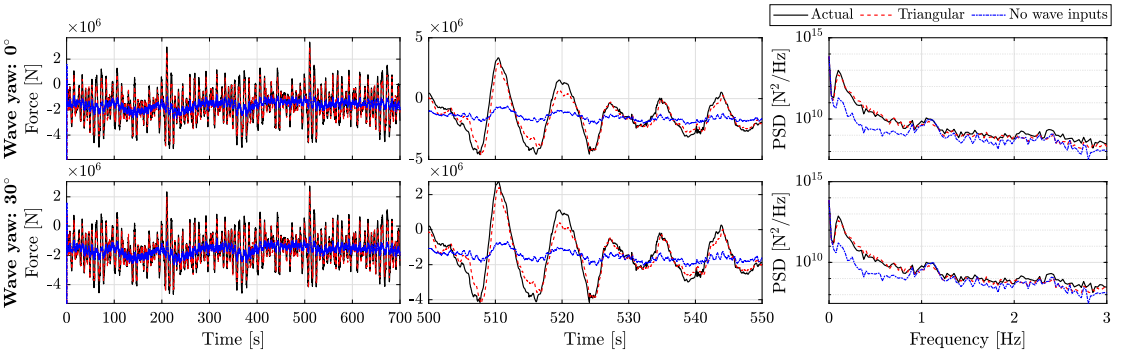
It is worth mentioning that the presence of waves had less impact on the accuracy of the estimated inputs at the tower top interface, with results comparable to those of Case Study 1. As a result, they are not included here.

## 6. Conclusions

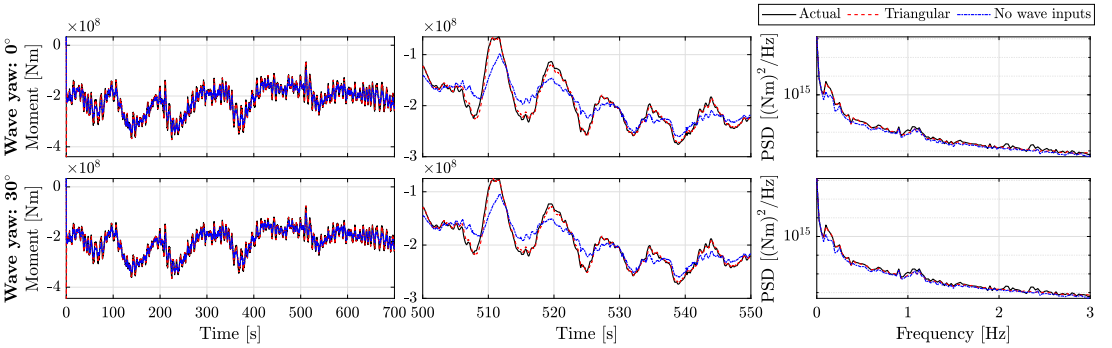
In this work, a sub-structuring approach has been applied for input-state estimation of a wind turbine sub-structure. By applying a cut at the top of the tower as well as at the ground level, the resulting sub-structure only requires modelling details which relate to the tower and monopile. This is advantageous as it significantly reduces the model uncertainty since the rotor-nacelle-assembly and soil–structure interaction do not need to be modelled.

The accuracy of the proposed approach was evaluated through numerical case studies conducted on a 15 MW offshore wind turbine. A higher fidelity wind turbine simulator was used to generate data, while a reduced order linear finite element model was used for identification purposes. These case studies demonstrated the capability of the suggested framework to achieve high performance for state and input estimation, as well as predicted fatigue damage. Regarding the interface forces, an offset was observed in the estimated interface forces, which was attributed to the linear finite element model not accounting for the non-linear p-delta effects due to gravity forces. This was confirmed by simulating a less realistic dataset, where forces due to gravity were not considered; for this case, the offset was negligible. Nonetheless, despite this offset, the overall trend in the estimated inputs was still found to be captured well, with high accuracy in the frequency domain. Similar results for the inputs were obtained irrespective of whether displacement or inclinometer measurements were employed. Moreover, the sparse sensor set-up was shown to produce accurate estimates for the interface inputs. However, due to having fewer accelerometers, the response was estimated with a time lag, and was attributed to the reduced contribution from terms in the direct feedthrough matrix $\mathbf{J}$.

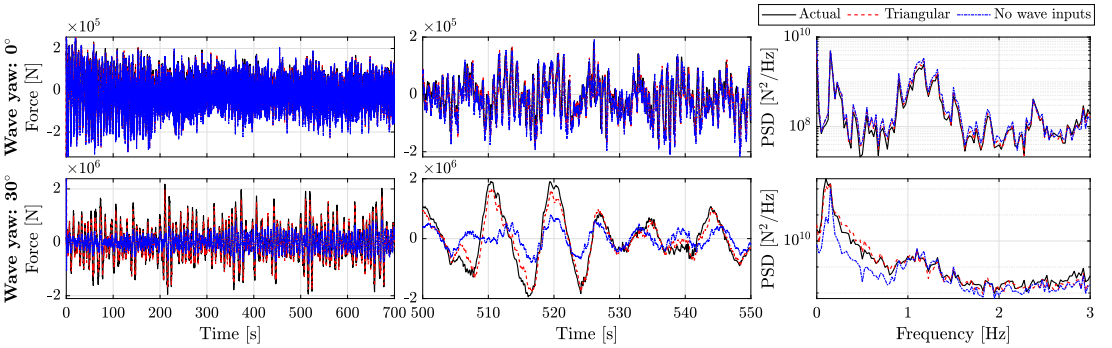
The states were also found to be estimated very accurately and were not affected by the aforementioned p-delta effects. However, when the displacement sensors were swapped for inclinometers in the measurement set, a drift was observed in the estimated displacements. This was attributed to the fact that the considered sub-structure is unconstrained, so that inclinometers and accelerometers cannot provide any information on the absolute position of the wind turbine. As a consequence of observability of this model and the specific sensor setup, the augmented Kalman filter is not able to estimate the rigid body response of the wind turbine. This conclusion was reinforced by examining the estimated modal displacements, where it was shown that all modes apart from the rigid body modes were accurately estimated. Note that if one displacement sensor is available to supplement the inclinometers and accelerometers, then the rigid body motion can also be estimated. This conclusion could potentially advise future



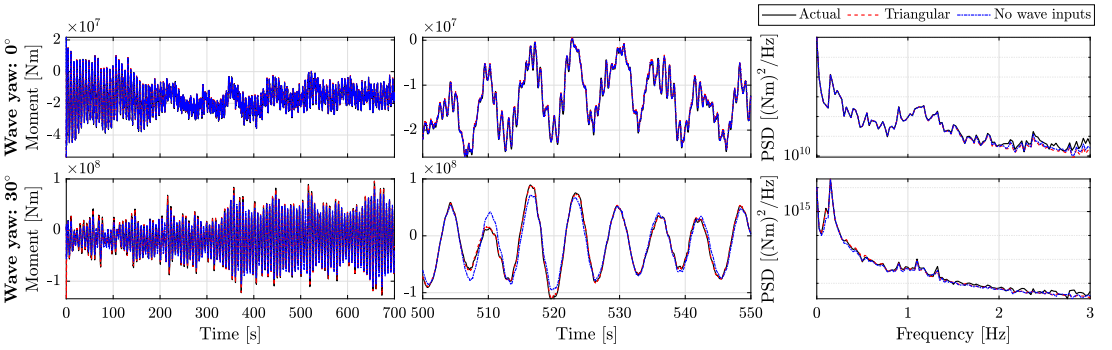

(a) Fore-aft shear force ($\mathtt{FO}_{\mathtt{FA}}$). Top: results for wave yaw = 0°. Bottom: results for wave yaw = 30°




(b) Fore-aft bending moment ($\mathtt{MO}_{\mathtt{FA}}$). Top: results for wave yaw = 0°. Bottom: results for wave yaw = 30°.




(c) Side-side shear force ($\mathtt{FO}_{\mathtt{SS}}$). Top: results for wave yaw = 0°. Bottom: results for wave yaw = 30°.




(d) Side bending moment ($\mathtt{MO}_{\mathtt{SS}}$). Top: results for wave yaw = 0°. Bottom: results for wave yaw = 30°.

**Fig. 13.** Ground level interface input estimation for Case Study 2 (wind and wave loading). *From left to right:* entire response duration, 10 s magnified excerpt, power spectral density.

measurement campaigns to consider installing at least one GPS sensor in each direction, for example in the nacelle, to provide this information. However, since the rigid body motion does not contribute to the deformation of the sub-structure, the predicted accumulated damage was found to be estimated with equal accuracy regardless of the sensor type.

In a second case study, the influence of both wind and wave loading was investigated. Here it was shown that the performance declines if inputs are not considered to account for the wave loading. Consequently, such excitations should be included in the inputs to be estimated. Furthermore, it was then demonstrated that a triangular distribution can be used to approximate the wave loading. This straightforward approximation only required two additional inputs and was shown to yield results with high accuracy in both the estimated damage distributions and the interface inputs.

Due to the significantly reduced model uncertainty, the proposed approach is considered to be a more robust method for estimating the states and inputs for the tower substructure, compared with trying to model the entire system. In addition, the method is computationally efficient and can be implemented in real-time, making it suitable for online applications.

While a key advantage of the framework lies in its minimal assumptions, meaning that very few would need to be removed for real-world applications, it is important to highlight its main limitations. In this study, the substructure model properties were assumed to be known. However, in real-world applications, there may exist uncertainties related to the definition of certain parameters, such as water level, hydrodynamic added mass coefficient, and modelling of the transition piece. If such parameters are thought to introduce significant modelling error, then the framework could be developed to include parameter estimation. In addition, accelerometers alone were found to be insufficient to accurately estimate inputs, resulting in a decrease in performance for state and fatigue estimation. Consequently, the number of inclinometers required to supplement the accelerometers used in this work may be greater than that which is typically available in practice. While collocation at the interface degrees of freedom was assumed in this work, this is not a strict requirement for successful input-state estimation; however, such an arrangement may still improve performance. An optimal sensor placement strategy was not considered in this work, and further studies are needed to explore this aspect.

While including a foundation model could allow for observability of the rigid displacements, the suggested approach allows for removing assumptions about the soil–structure-interaction model without consequences for the estimated fatigue life of the structure. Future work will look into methods to estimate soil stiffness properties based on the estimated ground level displacements and forces, while applying the substructuring approach to field data for real-world monitoring of offshore wind turbines.

## CRediT authorship contribution statement

**Harry A. Simpson:** Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Eleni N. Chatzi:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Manolis N. Chatzis:** Writing – review & editing, Supervision, Methodology, Conceptualization.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Harry Simpson reports financial support was provided by RWE Renewables. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

The authors gratefully acknowledge the financial support from RWE Renewables during this work.

## Data availability

Data will be made available on request.

## References

[1] L. Joyce, F. Zhao, Global wind energy report 2022, 2022.
[2] D. Cevasco, S. Koukoura, A. Kolios, Reliability, availability, maintainability data review for the identification of trends in offshore wind energy applications, Renew. Sustain. Energy Rev. 136 (2021) 110414.
[3] S. Vettori, E. Di Lorenzo, B. Peeters, E. Chatzi, A virtual sensing approach to operational modal analysis of wind turbine blades, in: Proceedings of ISMA2020 International Conference on Noise and Vibration Engineering, Leuven, Belgium, 2020, pp. 3579–3590.
[4] E. Akintunde, S.E. Azam, D.G. Linzell, Probabilistic bridge fatigue evaluation at virtual sensing locations using kernel density estimation, Int. J. Fatigue 176 (2023) 107885.
[5] B. Moynihan, E.M. Tronci, M.C. Hughes, B. Moaveni, E. Hines, Virtual sensing via Gaussian process for bending moment response prediction of an offshore wind turbine using SCADA data, Renew. Energy 227 (2024) 120466.
[6] A. Iliopoulos, R. Shirzadeh, W. Weijtjens, P. Guillaume, D. Van Hemelrijck, C. Devriendt, A modal decomposition and expansion approach for prediction of dynamic responses on a monopile offshore wind turbine using a limited number of vibration sensors, Mech. Syst. Signal Process. 68 (2016) 84–104.
[7] M. Song, S. Christensen, B. Moaveni, A. Brandt, E. Hines, Joint parameter-input estimation for virtual sensing on an offshore platform using output-only measurements, Mech. Syst. Signal Process. 170 (2022) 108814.

[8] M. Henkel, Validation of Virtual Sensing for the Fatigue Assessment of Subsoil and Submerged Components of Offshore Wind Turbines (Ph.D. thesis), Vrije Universiteit Brussel, 2022.
[9] M.-S. Nabiyan, F. Khoshnoudian, B. Moaveni, H. Ebrahimian, Mechanics-based model updating for identification and virtual sensing of an offshore wind turbine using sparse measurements, Struct. Control. Heal. Monit. 28 (2) (2021) e2647.
[10] R.E. Kalman, A New Approach to Linear Filtering and Prediction Problems, J. Basic Eng. 82 (1) (1960) 35–45, http://dx.doi.org/10.1115/1.3662552, arXiv:https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/82/1/35/5518977/35_1.pdf.
[11] H. Ebrahimian, R. Astroza, J.P. Conte, C. Papadimitriou, Bayesian optimal estimation for output-only nonlinear system and damage identification of civil structures, Struct. Control. Heal. Monit. 25 (4) (2018) e2128.
[12] K. Maes, M. Chatzis, G. Lombaert, Observability of nonlinear systems with unmeasured inputs, Mech. Syst. Signal Process. 130 (2019) 378–394.
[13] X. Shi, M. Chatzis, An efficient algorithm to test the observability of rational nonlinear systems with unmeasured inputs, Mech. Syst. Signal Process. 165 (2022) 108345.
[14] X. Shi, M.N. Chatzis, Lie symmetries of nonlinear systems with unknown inputs, Mech. Syst. Signal Process. 188 (2023) 110027.
[15] E. Lourens, E. Reynders, G. De Roeck, G. Degrande, G. Lombaert, An augmented Kalman filter for force identification in structural dynamics, Mech. Syst. Signal Process. 27 (2012) 446–460.
[16] S.E. Azam, E. Chatzi, C. Papadimitriou, A dual Kalman filter approach for state estimation via output-only acceleration measurements, Mech. Syst. Signal Process. 60 (2015) 866–886.
[17] S. Gillijns, B. De Moor, Unbiased minimum-variance input and state estimation for linear discrete-time systems, Automatica 43 (1) (2007) 111–116.
[18] S. Gillijns, B. De Moor, Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough, Automatica 43 (5) (2007) 934–937.
[19] S.P. Tchemodanova, M. Sanayei, B. Moaveni, K. Tatsis, E. Chatzi, Strain predictions at unmeasured locations of a substructure using sparse response-only vibration measurements, J. Civ. Struct. Heal. Monit. 11 (4) (2021) 1113–1136.
[20] K. Tatsis, E. Lourens, A comparison of two Kalman-type filters for robust extrapolation of offshore wind turbine support structure response, in: Life-Cycle of Engineering Systems:, CRC Press, 2016, pp. 209–216.
[21] P.C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, SIAM Rev. 34 (4) (1992) 561–580.
[22] K. Tatsis, V.K. Dertimanis, C. Papadimitriou, E. Lourens, E. Chatzi, A general substructure-based framework for input-state estimation using limited output measurements, Mech. Syst. Signal Process. 150 (2021) 107223.
[23] V.K. Dertimanis, E. Chatzi, S.E. Azam, C. Papadimitriou, Input-state-parameter estimation of structural systems from limited output information, Mech. Syst. Signal Process. 126 (2019) 711–746.
[24] S.E. Azam, E. Chatzi, C. Papadimitriou, A. Smyth, Experimental validation of the Kalman-type filters for online and real-time state and input estimation, J. Vib. Control 23 (15) (2017) 2494–2519.
[25] S. Vettori, E. Di Lorenzo, B. Peeters, M. Luczak, E. Chatzi, An adaptive-noise augmented Kalman filter approach for input-state estimation in structural dynamics, Mech. Syst. Signal Process. 184 (2023) 109654.
[26] K. Tatsis, E. Chatzi, E.-M. Lourens, Reliability prediction of fatigue damage accumulation on wind turbine support structures, in: Proceedings of the 2nd International Conference on Uncertainty Quantification in Computational Sciences and Engineering, National Technical University of Athens (NTUA), 2017, pp. 76–89.
[27] E. Branlard, D. Giardina, C.S. Brown, 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 Sci. 5 (3) (2020) 1155–1167.
[28] K. Maes, A. Iliopoulos, W. Weijtjens, C. Devriendt, G. Lombaert, Dynamic strain estimation for fatigue assessment of an offshore monopile wind turbine using filtering and modal expansion algorithms, Mech. Syst. Signal Process. 76 (2016) 592–611.
[29] K. Maes, G. De Roeck, G. Lombaert, A. Iliopoulos, D. Van Hemelrijck, C. Devriendt, P. Guillaume, Continuous strain prediction for fatigue assessment of an offshore wind turbine using Kalman filtering techniques, in: 2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS) Proceedings, IEEE, 2015, pp. 44–49.
[30] A. Mehrjoo, E.M. Tronci, B. Moynihan, B. Moaveni, F. Rüdinger, R. McAdam, E. Hines, Recursive Bayesian estimation of wind load on a monopile-supported offshore wind turbine using output-only measurements, Mech. Syst. Signal Process. 224 (2025) 112183.
[31] N. Noppe, A. Iliopoulos, W. Weijtjens, C. Devriendt, Full load estimation of an offshore wind turbine based on SCADA and accelerometer data, in: Journal of Physics: Conference Series, Vol. 753, IOP Publishing, 2016, 072025.
[32] N. Noppe, K. Tatsis, E. Chatzi, C. Devrient, W. Weijtjens, Fatigue stress estimation of offshore wind turbine using a Kalman filter in combination with accelerometers, in: Proceedings of International Conference on Noise and Vibration Engineering (ISMA 2018), International Conference on Uncertainty in Structural Dynamics (USD 2018), KU Leuven, Department of Mechanical Engineering, 2018, pp. 4693–6701.
[33] B.W. Byrne, G.T. Houlsby, H.J. Burd, K.G. Gavin, D.J. Igoe, R.J. Jardine, C.M. Martin, R.A. McAdam, D.M. Potts, D.M. Taborda, et al., PISA design model for monopiles for offshore wind turbines: application to a stiff glacial clay till, Géotechnique 70 (11) (2020) 1030–1047.
[34] API, Recommended Practice for Planning, Designing & Constructing Fixed Offshore Platforms, Tech. Rep. API-RP.2A-WSD, American Petroleum Institute, 2010.
[35] B.W. Byrne, R.A. McAdam, W.J. Beuckelaers, H.J. Burd, K. Gavin, G.T. Houlsby, D.J.P. Igoe, R. Jardine, C.M. Martin, et al., Cyclic laterally loaded medium scale field pile testing for the pisa project, in: 4th International Symposium on Frontiers in Offshore Geotechnics (Postponed), 2020, pp. 1323–1332.
[36] H.J. Burd, C.N. Abadie, B.W. Byrne, G.T. Houlsby, C.M. Martin, R.A. McAdam, R.J. Jardine, A.M. Pedro, D.M. Potts, D.M. Taborda, et al., Application of the PISA design model to monopiles embedded in layered soils, Géotechnique 70 (11) (2020) 1067–1082.
[37] F. Nadim, Accounting for uncertainty and variability in geotechnical characterization of offshore sites, in: Geotechnical Safety and Risk V, IOS Press, 2015, pp. 23–35.
[38] M. Henkel, W. Weijtjens, C. Devriendt, Fatigue stress estimation for submerged and sub-soil welds of offshore wind turbines on monopiles using modal expansion, Energies 14 (22) (2021) 7576.
[39] D. De Klerk, D.J. Rixen, S. Voormeeren, General framework for dynamic substructuring: history, review and classification of techniques, AIAA J. 46 (5) (2008) 1169–1181.
[40] B. Besselink, U. Tabak, A. Lutowska, N. Van de Wouw, H. Nijmeijer, D.J. Rixen, M. Hochstenbach, W. Schilders, A comparison of model reduction techniques from structural dynamics, numerical mathematics and systems and control, J. Sound Vib. 332 (19) (2013) 4403–4422.
[41] W.C. Hurty, Dynamic analysis of structural systems using component modes, AIAA J. 3 (4) (1965) 678–685.
[42] S. Rubin, Improved component-mode representation for structural dynamic analysis, AIAA J. 13 (8) (1975) 995–1006.
[43] R.H. MacNeal, A hybrid method of component mode synthesis, Comput. Struct. 1 (4) (1971) 581–601.
[44] R.R. Craig Jr., C.-J. Chang, A Review of Substructure Coupling Methods for Dynamic Analysis, vol. 2, NASA Langley Res. Center Advan. in Eng. Sci., 1976.
[45] S. Vettori, E.D. Lorenzo, R. Cumbo, U. Musella, T. Tamarozzi, B. Peeters, E. Chatzi, Kalman-based virtual sensing for improvement of service response replication in environmental tests, in: Model Validation and Uncertainty Quantification, Volume 3: Proceedings of the 38th IMAC, a Conference and Exposition on Structural Dynamics 2020, Springer, 2020, pp. 93–106.

[46] S. Vettori, E. Di Lorenzo, B. Peeters, E. Chatzi, Assessment of alternative covariance functions for joint input-state estimation via Gaussian process latent force models in structural dynamics, Mech. Syst. Signal Process. 213 (2024) 111303.
[47] J. Meng, R.A. McAdam, M.N. Chatzis, A simulink model for the dynamic analysis of floating wind turbines, in: "Dynamics of Civil Structures, Volume 2": Conference Proceedings of the Society for Experimental Mechanics Series, Springer, 2024, pp. 229–239.
[48] H. Simpson, K.E. Tatsis, I. Abdallah, E. Chatzi, M.N. Chatzis, Estimating the foundation parameters of offshore wind turbines through Bayesian model updating, in: Journal of Physics: Conference Series, Vol. 2647 - 11, IOP Publishing, 2024, 112008.
[49] E. Gaertner, J. Rinker, L. Sethuraman, F. Zahle, B. Anderson, G.E. Barter, N.J. Abbas, F. Meng, P. Bortolotti, W. Skrzypinski, et al., IEA Wind TCP Task 37: Definition of the IEA 15-Megawatt Offshore Reference Wind Turbine, Tech. Rep., National Renewable Energy Lab.(NREL), Golden, CO (United States, 2020.
[50] R.R. Craig Jr., M.C. Bampton, Coupling of substructures for dynamic analyses., AIAA J. 6 (7) (1968) 1313–1319.
[51] M.O. Hansen, Aerodynamics of wind turbines. Earthscan, James James 8 (9) (2008) 14.
[52] B. Jonkman, L. Kilcher, TurbSim User's Guide: Version 1.06. 00, National Renewable Energy Laboratory, Golden, CO, USA, 2012.
[53] J. Morison, J.W. Johnson, S.A. Schaaf, The force exerted by surface waves on piles, J. Pet. Technol. 2 (05) (1950) 149–154.
[54] K. Hasselmann, T.P. Barnett, E. Bouws, H. Carlson, D.E. Cartwright, K. Enke, J. Ewing, A. Gienapp, D. Hasselmann, P. Kruseman, et al., Measurements of wind-wave growth and swell decay during the joint north sea wave project (JONSWAP), Ergaenzungsheft Zur Dtsch. Hydrogr. Z. Reihe A (1973).
[55] NREL, Rosco. Version 2.3.0, 2021, URL https://github.com/NREL/rosco.
[56] S. Adhikari, Damping Models for Structural Vibration (Ph.D. thesis), University of Cambridge, 2001.
[57] K. Maes, E. Lourens, K. Van Nimmen, E. Reynders, G. De Roeck, G. Lombaert, Design of sensor networks for instantaneous inversion of modally reduced order models in structural dynamics, Mech. Syst. Signal Process. 52 (2015) 628–644.
[58] M. Matsuishi, T. Endo, Fatigue of Metals Subjected to Varying Stress, Vol. 68, Japan Society of Mechanical Engineers, Fukuoka, Japan, 1968, pp. 37–40, 2.
[59] D.N.V. AS, et al., Fatigue design of offshore steel structures, Rev. 3 (2011).
[60] L.R. Curtis, B.M. Ayyub, A meta-analysis of Miner's rule, in: 2024 Annual Reliability and Maintainability Symposium, RAMS, IEEE, 2024, pp. 1–6.
[61] R. Gaiotti, B.S. Smith, P-delta analysis of building structures, J. Struct. Eng. 115 (4) (1989) 755–770.