



Design and implementation of a location-agnostic and control-oriented load surrogate model for wind turbines

Adrien Guilloiré, Andreas Vad, Filippo Campagnolo, and Carlo L. Bottasso

Wind Energy Institute, Technical University of Munich, 85748 Garching bei München, Germany

Correspondence: Carlo L. Bottasso (carlo.bottasso@tum.de)

Abstract.

Accurate estimation of wind turbine fatigue loads is essential for wind farm design, optimization, and control applications. Load surrogate models have recently emerged as a promising approach to reduce the computational cost associated with high-fidelity load calculations. However, while previous studies have primarily focused on demonstrating predictive accuracy, considerably less attention has been devoted to understanding how such surrogates should be designed and deployed in practice.

This study systematically investigates the design and implementation of location-agnostic wind turbine fatigue load surrogate models. A comprehensive database of aeroelastic simulations covering free-stream, partially waked, and fully waked operating conditions is used to evaluate alternative inflow representations, control descriptors, number of turbulent realizations, and training dataset sizes. The practical deployment of the resulting surrogate formulation is further assessed using multiple inflow-estimation approaches, including observer-based inflow reconstruction and steady-state wake models.

A key finding of this work is that there exists a finite level of physical complexity sufficient to achieve converged performance for a general wind turbine load surrogate model. Among the investigated options, an inflow representation consisting of four sector-averaged wind speed quantities and one rotor-averaged turbulence parameter (denoted SAWS 4A + RATI) emerges as the simplest formulation capable of reproducing converged predictive accuracy. Likewise, steady-state control setpoints are shown to provide a sufficient description of control actions, approximately 10–12 turbulent realizations are required for robust convergence with respect to turbulence variability, and surrogate performance converges with a relatively modest number of training cases.

The results further demonstrate that the same design recommendations remain valid across all investigated implementation environments. Although significant performance degradation occurs when training and deployment rely on different inflow-description spaces, high predictive accuracy is consistently recovered when the surrogate is trained and deployed using the same inflow representation. These findings provide practical design guidelines for the development of fatigue-load surrogate models and demonstrate that accurate, robust, and computationally efficient load prediction can be achieved using a small set of physically meaningful reduced-order descriptors.



1 Introduction

25 The ability to efficiently predict structural loads is essential in many wind energy applications, including wind farm design, operational optimization, wind farm control, remaining useful life assessment, lifetime extension studies, and standards-based design and certification activities. In all these applications, large numbers of operating conditions, turbine configurations, or control strategies must often be evaluated while accounting for their impact on fatigue loading. Since structural loads directly influence maintenance costs, component reliability, certification compliance, and overall plant economics, their accurate prediction is a key requirement for informed decision-making throughout the lifecycle of a wind farm.

30 A major challenge lies in the computational cost associated with load prediction. High-fidelity aeroelastic simulation tools are capable of accurately estimating turbine loads under a wide range of operating and inflow conditions, and remain the reference approach for standards-compliant design and certification analyses. However, the large number of simulations required in such studies, as well as in applications involving large design spaces, repeated optimization, uncertainty quantification, or real-time decision-making, can result in substantial computational costs. Computationally efficient load surrogate models therefore represent a promising complementary approach, enabling rapid load estimation while retaining sufficient accuracy for engineering applications.

Recent years have witnessed increasing research efforts devoted to the development of load surrogate models for wind turbine and wind farm applications (Dimitrov et al., 2018; Bossanyi, 2022; Shaler et al., 2022; He et al., 2022; Liew et al., 2023b; Pettas and Cheng, 2024; Liew et al., 2024; Witter et al., 2025). Although encouraging results have been already demonstrated, many existing models remain tailored to specific sites, inflow conditions, or control strategies, thereby limiting their applicability outside the conditions under which they were trained. In particular, the combined effects of wake-induced inflow heterogeneity and control actions remain difficult to represent in a generic and computationally efficient manner.

45 An approach toward generic load surrogate modeling was introduced in Guilloré et al. (2024). Rather than relying on site-specific inflow descriptors, the method characterizes the local rotor inflow using sector-averaged wind speed and turbulence intensity quantities extracted across the rotor disk. These variables are then related to 10-minute damage equivalent loads (DELs) through an artificial neural network. Because the inputs represent only the local inflow experienced by the turbine, the resulting model is independent of the upstream wind farm configuration and can therefore be applied across different layouts and operating conditions. The model demonstrated promising predictive capabilities and successfully generalized to an unseen wind farm.

While these results established the feasibility of generic load surrogate modeling, several fundamental design questions remain unanswered in the extensive literature on load surrogate modeling. More specifically, it remains unclear which modeling choices are most important for constructing a load surrogate that is simultaneously accurate, robust, computationally efficient, and suitable for wind farm control applications. Addressing these questions is the primary motivation of the present study.

55 The central hypothesis of this work is that the development of an effective load surrogate model must be approached as a design problem. Rather than focusing solely on the choice of machine-learning architecture, the predictive capability and practical usefulness of the resulting surrogate are expected to depend strongly on the selection of input variables, the generation



of the training dataset, and the intended application environment. Consequently, this work investigates three fundamental questions.

60 First, what is an appropriate representation of the turbine inflow for generic load prediction? The chosen inputs must be sufficiently rich to capture the dominant mechanisms governing fatigue loading, while remaining simple enough to be estimated by lower-order engineering flow models (Jensen, 1983; Bastankhah and Porté-Agel, 2016; Göcmen et al., 2016) and by turbine-level wind sensing techniques (Bottasso et al., 2018; Schreiber et al., 2020; Bertelè et al., 2024; Onnen et al., 2026). This trade-off is particularly important for wind farm control applications, where computationally expensive inflow representations would
65 undermine the practical benefits of the surrogate model. In this context, the present work evaluates different reduced-order descriptions of the rotor inflow and investigates the degree of spatial resolution required to accurately predict fatigue loads.

Second, how much training data is required to develop a reliable load surrogate model? The generation of high-fidelity aeroelastic training datasets constitutes a significant computational investment. It is therefore important to quantify the number of operating conditions, simulation cases, and turbulent realizations necessary to achieve statistical convergence and satisfactory
70 predictive performance. This paper investigates the influence of both the number of turbulent seeds and the overall size of the training database on the resulting model accuracy.

Third, how robust is the resulting load surrogate model when deployed outside its training conditions? A surrogate intended for wind farm control should generalize to previously unseen wind farm layouts and operating conditions while maintaining acceptable predictive accuracy. Furthermore, practical applications often rely on imperfect estimates of the model inputs ob-
75 tained either from steady-state wake models in offline optimization frameworks or from wind observers in real-time control applications. The present work therefore evaluates the robustness of the proposed surrogate model against unseen wind farm configurations and investigates its compatibility with both engineering wake models and observer-based inflow estimates.

A key objective throughout this study is the development of a load surrogate model that is both location-agnostic and control-oriented. Here, location-agnostic refers to a formulation that depends only on the local turbine inflow rather than
80 on farm-specific quantities such as turbine spacing, layout geometry, or wake overlap metrics. Control-oriented refers to a representation that remains applicable under a broad range of turbine- and wind-farm-level control actions. This enables direct integration within actual and future load-aware wind farm control frameworks, including applications involving yaw-based wake steering, power derating, wake mixing, and other coordinated control strategies.

The main contributions of this paper can be summarized as follows:

- 85 1. A systematic evaluation of load surrogate model design choices, including inflow representation, control representation, training database size, and turbulent sampling requirements.
2. A quantitative assessment of the trade-off between model simplicity and predictive accuracy, leading to the identification of reduced-order inflow and control representations that are sufficiently informative for fatigue-load prediction.
3. A demonstration of the applicability of the proposed surrogate-model formulation using different sources of inflow
90 information, including engineering wake models and wind-observer estimates.



4. Design guidelines for the development of computationally efficient, generic, and control-oriented load surrogate models.

A recurring theme throughout this work is the identification of reduced-order inflow and control representations that achieve an appropriate balance between predictive accuracy and implementation complexity. The objective is not to maximize predictive performance at any cost, but rather to determine the minimum amount of information required to achieve satisfactory load-prediction accuracy. Consequently, when multiple surrogate formulations yield comparable performance, preference is given to the simplest representation in terms of input dimensionality, physical interpretability, and practical estimability. While more elaborate formulations may provide marginal improvements in predictive accuracy, they often require additional information that may be difficult to obtain from engineering wake models, wind observers, or other practical sources. For this reason, surrogate design and surrogate implementation are considered jointly throughout the present study, with the aim of identifying formulations that are both accurate and readily deployable in realistic wind energy applications.

The scope of the present study is intentionally limited to atmospheric and wake conditions representative of conventional utility-scale wind farm operation. The training and validation datasets consider standard atmospheric boundary layer conditions and do not include effects such as wind veer, low-level jets, or other complex atmospheric phenomena. Consequently, the conclusions regarding the optimal spatial resolution of the rotor inflow representation may not directly generalize to such conditions.

Furthermore, wake fields are generated using models that do not explicitly represent curl-wake effects. The conclusions regarding the number of inflow sectors required to characterize rotor-scale flow heterogeneity should therefore be interpreted within this modeling framework. Future investigations may revisit these findings using wake models that incorporate explicit curl-wake dynamics. Finally, the present work focuses primarily on far-wake operating conditions. Near-wake effects are not explicitly assessed because they are not represented by the wake modeling framework used to generate the datasets. Future studies involving enhanced wake models and large eddy simulations could provide further insight into whether a finer spatial representation of the rotor inflow becomes necessary in the presence of strong near-wake flow structures.

The remainder of the paper is organized as follows. Section 2 introduces the proposed location-agnostic and control-oriented load surrogate model and its underlying formulation. Section 3 describes the datasets used throughout the study. Section 4 evaluates the impact of key design choices on surrogate model accuracy and provides recommendations for model development. Section 5 illustrates how the proposed framework can be implemented using different sources of inflow information. Finally, Section 6 summarizes the main conclusions and discusses opportunities for future research.

2 Formulation of a location-agnostic and control-oriented load surrogate model

This work approaches load surrogate modeling as a design problem. Rather than focusing solely on the choice of machine-learning algorithm, the predictive performance, generality, and practical applicability of a load surrogate model are expected to depend strongly on a series of modeling decisions. These include the selection of reduced-order inflow descriptors, the representation of turbine control actions, the generation of the training database, and the intended deployment environment. Figure 1 summarizes the overall framework adopted in this study.

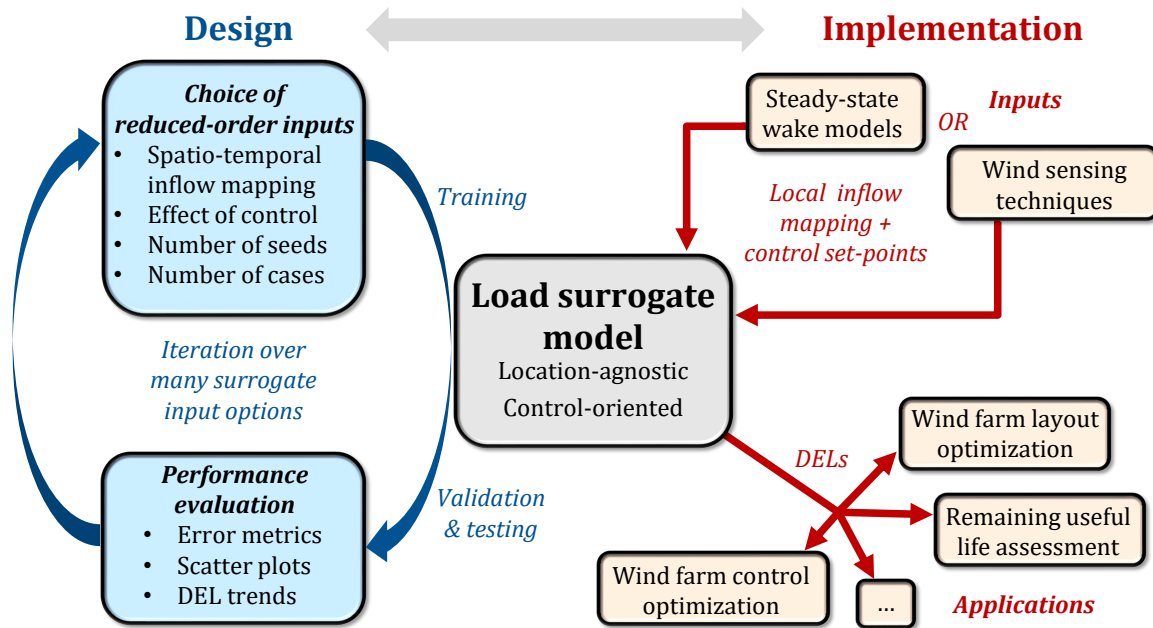


Figure 1. General framework of the proposed load surrogate model. The development process combines surrogate design choices, such as the selection of reduced-order inflow and control representations, with iterative performance evaluation to construct and refine a surrogate capable of predicting fatigue loads across diverse operating conditions, thereby supporting a wide range of applications.

The left-hand side of Fig. 1 illustrates the surrogate-model design process. Different candidate representations of the local inflow, turbine operating state, and training strategy are combined to construct a family of load surrogate models. The performance of the different configurations is subsequently evaluated in order to identify solutions that achieve a suitable compromise between predictive accuracy, computational efficiency, and implementation complexity.

The right-hand side of Fig. 1 highlights potential deployment environments for the resulting surrogate model. A key objective of this work is to formulate the surrogate using quantities that can be obtained from a variety of information sources, including high-fidelity simulations, engineering wake models, and wind-observer estimates. By relying on reduced-order representations of the local inflow and turbine operating state, the proposed framework aims to remain independent of specific wind farm layouts while retaining applicability to a broad range of wind energy applications, including wind farm design, operational optimization, load-aware control, and lifetime assessment.

To maintain the focus of the present manuscript on the formulation, training, and validation of the load surrogate model, the application examples illustrated in the bottom right of Fig. 1 are not investigated further herein. The objective of this work is primarily to establish and evaluate a location-agnostic and control-oriented surrogate formulation rather than to assess its performance within specific decision-making frameworks. Nevertheless, the proposed surrogate, or closely related



variants thereof, has already been employed in several application-oriented studies, including load-aware wind farm control optimization (Braunbehrens et al., 2024; Anand et al., 2025; Lokken et al., 2026; van der Hoek et al., 2026), wind farm layout optimization with explicit fatigue-load considerations (Kainz et al., 2026; Bay et al., 2026), turbine lifetime assessment (Vad et al., 2026b), and more general holistic wind farm modelling frameworks (Vad et al., 2026a). These studies illustrate the broad applicability of the present surrogate approach, while a detailed investigation of such applications remains beyond the scope of the current paper.

The formulation of the proposed load surrogate model is detailed in the following sections. First, the mapping between reduced-order inflow and control parameters and turbine fatigue loads is introduced. Subsequently, different representations of the inflow and control states are defined and compared, before describing the machine-learning framework used to approximate the resulting surrogate functions.

2.1 Damage equivalent loads

The objective of the proposed surrogate model is to predict fatigue loads from reduced-order descriptions of the local inflow conditions and turbine operating state. Fatigue loads are quantified using DELs, which are widely employed in wind-turbine design and load assessment because they condense an arbitrary load history into equivalent constant-amplitude loads producing the same fatigue damage.

For a given load time series, the DEL is computed using conventional rainflow cycle counting according to

$$\text{DEL} = \left(\frac{\sum_j n_j L_j^m}{N_{\text{eq}}} \right)^{1/m}, \quad (1)$$

where n_j is the number of load cycles of amplitude L_j , N_{eq} is the equivalent number of cycles, and m is the Wöhler exponent of the corresponding S-N curve (Miner, 1945). Following common wind-energy practice, $N_{\text{eq}} = 600$ is used to represent a reference frequency of 1 Hz over a 10 min period (International Electrotechnical Commission, 2019; Sutherland, 1999). The values $m = 10$ and $m = 4$ are used for blade and steel structural components, respectively.

The formulation presented in this work is independent of the specific structural quantity considered. In principle, any load channel for which a DEL can be computed may be used as a surrogate output. The specific load channels selected for the present study are described in Sect. 3.3.2.

For load quantities characterized by two orthogonal components, a projected DEL may additionally be defined by projecting the instantaneous load vector onto multiple directions and retaining the direction yielding the largest fatigue damage. This approach is adopted in the present work to identify the most critical loading direction at a given structural cross section.

2.2 Surrogate formulation

The objective of the present work is to construct computationally efficient surrogate models capable of predicting fatigue loads from reduced-order descriptions of the local inflow conditions and turbine operating state. For a given structural load channel l , the corresponding DEL_l is assumed to depend primarily on the local inflow experienced by the rotor and on the applied turbine control actions.



170 The local flow conditions at the rotor disk are represented by a vector of reduced-order *inflow* parameters, denoted by $\mathbf{i}$, while the turbine operating state is represented by a vector of *control* parameters, denoted by $\mathbf{c}$. The load surrogate model for load channel l is therefore formulated as

$$\text{DEL}_l = \mathcal{L}_l(\mathbf{i}, \mathbf{c}), \quad (2)$$

where $\mathcal{L}_l$ denotes the unknown mapping between inflow conditions, control actions, and fatigue loads.

175 While the approximation of $\mathcal{L}_l$ is an important component of the surrogate-model development process, the primary objective of this work is to determine which information should be provided to the surrogate model in the first place. In particular, suitable definitions of the inflow and control descriptors $\mathbf{i}$ and $\mathbf{c}$ are sought that are rich enough to accurately predict fatigue loads, yet sufficiently simple to be estimated by low-order engineering models and wind observers. Different inflow and control representations are therefore systematically investigated and compared throughout the paper.

180 Candidate inflow representations for $\mathbf{i}$ are introduced in Sect. 2.3, while the control representations in $\mathbf{c}$ are presented in Sect. 2.4. The mapping functions $\mathcal{L}_l$ are approximated using artificial neural networks (ANNs), as described in Sect. 2.5.

2.3 Mapping of local inflow into reduced-order quantities

This subsection describes the mapping of the local rotor inflow onto a reduced-order representation. The reduced-order inflow quantities investigated in this work are first defined in Sect. 2.3.1. Subsequently, Sect. 2.3.2 discusses how these quantities
185 can be estimated in practical simulation and operational environments. Finally, Sect. 2.3.3 establishes the theoretical equivalence between alternative inflow representation spaces for load prediction and discusses its implications for surrogate model deployment.

2.3.1 Definition of reduced-order inflow quantities

The selection of appropriate inflow descriptors constitutes one of the central design choices in load surrogate modeling. In
190 principle, the local inflow experienced by a wind turbine rotor may be represented using a wide range of reduced-order techniques, including modal decompositions, basis-function expansions, or machine-learning-based latent representations (Liew et al., 2023b, 2024; Ramaswamy et al., 2026). However, such approaches often introduce quantities that are difficult to interpret physically and challenging to estimate in practical applications.

The present work therefore adopts a geometric description of the rotor inflow based on spatially averaged quantities over the
195 rotor disk and its sectors, an approach first developed by Bottasso et al. (2018) for application in wind observers. This choice is motivated by three main considerations. First, the resulting variables retain a direct physical interpretation and can be related to well-known inflow characteristics such as wind speed, turbulence intensity, and wake-induced asymmetries. Second, the same quantities can be obtained from a variety of information sources, including high-fidelity simulations, engineering wake models, and wind observers. Third, such quantities can be readily incorporated into control-oriented applications, where computational
200 simplicity and observability are essential requirements.



The objective of this section is therefore to define a family of reduced-order inflow descriptors with increasing levels of spatial and temporal resolution, ranging from global rotor-integrated quantities to sector-based representations capable of capturing partial-wake conditions.

In all the following, $u(\psi, r, t)$ denotes the instantaneous streamwise wind velocity component at azimuthal position ψ , radial position r , and time instant t over the rotor disk. The variables r and t are defined over the intervals $0 \leq r \leq R$ and $0 \leq t \leq T$, where R denotes the rotor radius and T the duration of the considered time window.

The simplest reduced-order description of the inflow is obtained by condensing the rotor-resolved velocity field into a small number of scalar quantities. Two alternative formulations are considered in this work. The first computes local statistics prior to spatial averaging and is therefore referred to as the *averaged* formulation. The second first performs the spatial averaging of the velocity field and subsequently computes the associated statistics, and is therefore referred to as the *effective* formulation. Accordingly, the rotor-averaged wind speed (RAWS), rotor-averaged turbulence intensity (RATI), rotor-effective wind speed (REWS), and rotor-effective turbulence intensity (RETI) are defined as

$$\text{RAWS} = \frac{1}{\pi R^2} \int_0^{2\pi} \int_0^R \left(\frac{1}{T} \int_0^T u(\psi, r, t) dt \right) r dr d\psi, \quad (3a)$$

$$\text{RATI} = \frac{1}{\pi R^2} \int_0^{2\pi} \int_0^R \left(\frac{\sqrt{\frac{1}{T} \int_0^T \left(u(\psi, r, t) - \frac{1}{T} \int_0^T u(\psi, r, t) dt \right)^2 dt}}{\frac{1}{T} \int_0^T u(\psi, r, t) dt} \right) r dr d\psi \cdot 100, \quad (3b)$$

$$\text{REWS} = \frac{1}{T} \int_0^T \left(\frac{1}{\pi R^2} \int_0^{2\pi} \int_0^R u(\psi, r, t) r dr d\psi \right) dt, \quad (3c)$$

$$\text{RETI} = \frac{\sqrt{\frac{1}{T} \int_0^T \left(\frac{1}{\pi R^2} \int_0^{2\pi} \int_0^R u(\psi, r, t) r dr d\psi - \text{REWS} \right)^2 dt}}{\text{REWS}} \cdot 100. \quad (3d)$$

The quantities RAWS and REWS both provide a single scalar measure of the mean inflow velocity across the rotor disk, while RATI and RETI characterize the corresponding turbulence intensity. The distinction between the two formulations lies in the order in which spatial averaging and statistical operations are performed.

Because the averaging operator is linear, RAWS and REWS become mathematically equivalent when evaluated over identical spatial and temporal domains, up to numerical discretization errors. In contrast, turbulence intensity involves the computation of a standard deviation, which is a nonlinear operator. Consequently, the operations of spatial averaging and turbulence-intensity estimation do not commute, leading in general to different values of RATI and RETI.

From a physical perspective, the averaged quantities describe the local inflow conditions before aggregation over the rotor disk, whereas the effective quantities characterize the fluctuations of an equivalent rotor-averaged velocity signal. Both formulations are therefore retained and evaluated in this work as potential inputs to the load surrogate model.



230 By convention, RATI and RETI are expressed in percentage by multiplying the corresponding ratios by 100.

While rotor-integrated quantities provide a compact representation of the inflow, they cannot explicitly capture spatial heterogeneity across the rotor disk. This limitation becomes particularly important under waked operating conditions, where significant velocity and turbulence gradients may exist across different regions of the rotor. To account for such effects while retaining a reduced-order description, the present work extends the rotor-integrated quantities to sector-based quantities.

235 To map the local inflow experienced by a turbine with greater level of details than single values over the rotor, this work introduces the use of sector-averaged wind speeds (SAWS), sector-averaged turbulence intensities (SATI), sector-effective wind speeds (SEWS) and sector-effective turbulence intensities (SETI) defined as the following scalar quantities

$$\text{SAWS}_s(\psi_s^-, \psi_s^+) = \frac{1}{(\psi_s^+ - \psi_s^-)(R^2/2)} \int_{\psi_s^-}^{\psi_s^+} \int_0^R \left(\frac{1}{T} \int_0^T u(\psi, r, t) dt \right) r dr d\psi, \quad (4a)$$

$$240 \quad \text{SATI}_s(\psi_s^-, \psi_s^+) = \frac{1}{(\psi_s^+ - \psi_s^-)(R^2/2)} \int_{\psi_s^-}^{\psi_s^+} \int_0^R \left(\frac{\sqrt{\frac{1}{T} \int_0^T \left(u(\psi, r, t) - \frac{1}{T} \int_0^T u(\psi, r, t) dt \right)^2 dt}}{\frac{1}{T} \int_0^T u(\psi, r, t) dt} \right) r dr d\psi \cdot 100, \quad (4b)$$

$$\text{SEWS}_s(\psi_s^-, \psi_s^+) = \frac{1}{T} \int_0^T \left(\frac{1}{(\psi_s^+ - \psi_s^-)(R^2/2)} \int_{\psi_s^-}^{\psi_s^+} \int_0^R u(\psi, r, t) r dr d\psi \right) dt, \quad (4c)$$

$$\text{SETI}_s(\psi_s^-, \psi_s^+) = \frac{\sqrt{\frac{1}{T} \int_0^T \left(\frac{1}{(\psi_s^+ - \psi_s^-)(R^2/2)} \int_{\psi_s^-}^{\psi_s^+} \int_0^R u(\psi, r, t) r dr d\psi - \text{SEWS}_s(\psi_s^-, \psi_s^+) \right)^2 dt}}{\text{SEWS}_s(\psi_s^-, \psi_s^+)} \cdot 100. \quad (4d)$$

245 The sector-based quantities defined in Eqs. (4a)–(4d) depend on the choice of the azimuthal boundaries ψ_s^- and ψ_s^+ that define each sector. Different choices of sector number and placement lead to different reduced-order representations of the rotor inflow and therefore constitute distinct load-surrogate designs.

In the present work, nine sector configurations are considered to investigate the influence of spatial inflow resolution on surrogate-model performance. Each configuration is identified by two parameters: (i) the number of azimuthal sectors used to
 250 discretize the rotor disk, and (ii) a placement indicator specifying whether a sector is centered on the vertical axis (configuration A) or whether a sector boundary coincides with the vertical axis (configuration B). The considered configurations are illustrated in Fig. 2.

The rotor-integrated quantities introduced previously can be interpreted as particular cases of the sector-based formulation. When the entire rotor disk is represented by a single sector spanning the full azimuthal range $0 \leq \psi < 2\pi$, SAWS_1 and SATI_1
 255 become equivalent to RAWs and RATI, respectively. Similarly, SEWS_1 and SETI_1 become equivalent to REWS and RETI. The sector-based formulation therefore provides a generalized framework encompassing both rotor-integrated and spatially resolved inflow representations.

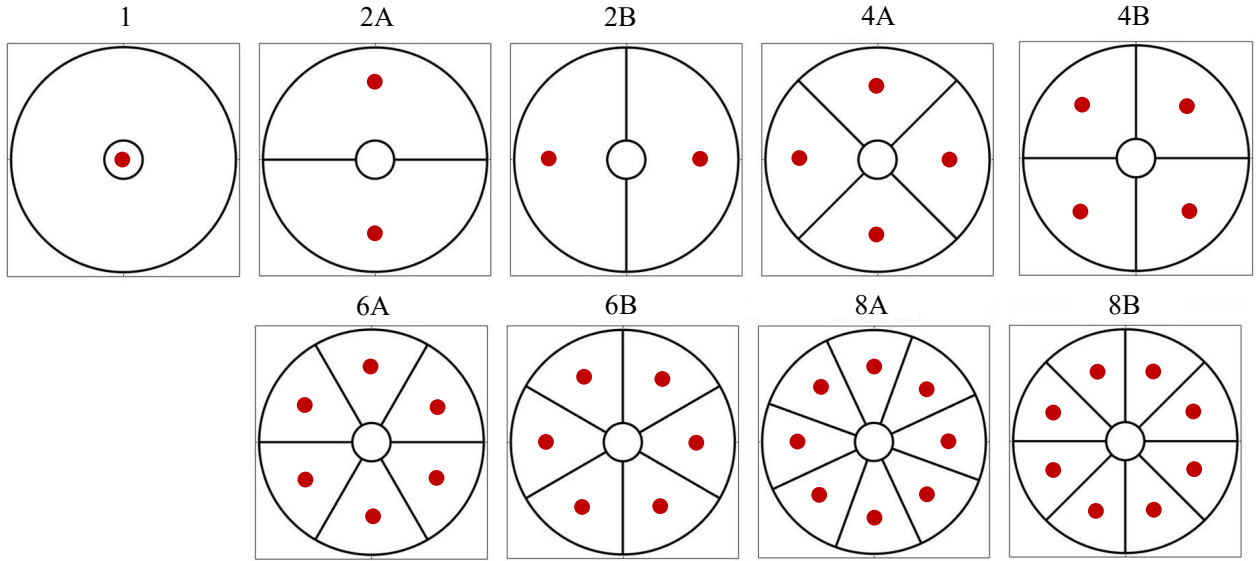


Figure 2. Schematic illustration of the sector configurations considered in this study. The notation indicates the number of azimuthal sectors and their placement relative to the vertical axis.

Among the configurations illustrated in Fig. 2, the four-sector configuration with sectors centered on the vertical axis (denoted 4A) was introduced in Guilloré et al. (2024) as a reduced-order inflow representation for load surrogate modeling. The present study extends that work by comparing the 4A configuration against eight alternative sector arrangements in order to assess the influence of spatial inflow discretization on surrogate-model performance.

With the azimuthal coordinate defined such that $\psi = 0$ corresponds to the upward vertical direction and increases clockwise when looking downstream, the four sector-averaged wind speeds of the 4A configuration are obtained from Eq. (4a) as

$$\text{SAWS}_{4A,1,\text{"up"}} = \text{SAWS}_s \left(\psi_s^- = -\frac{\pi}{4}, \psi_s^+ = \frac{\pi}{4} \right), \quad (5a)$$

$$\text{SAWS}_{4A,2,\text{"right"}} = \text{SAWS}_s \left(\psi_s^- = \frac{\pi}{4}, \psi_s^+ = \frac{3\pi}{4} \right), \quad (5b)$$

$$\text{SAWS}_{4A,3,\text{"down"}} = \text{SAWS}_s \left(\psi_s^- = \frac{3\pi}{4}, \psi_s^+ = \frac{5\pi}{4} \right), \quad (5c)$$

$$\text{SAWS}_{4A,4,\text{"left"}} = \text{SAWS}_s \left(\psi_s^- = \frac{5\pi}{4}, \psi_s^+ = -\frac{\pi}{4} \right). \quad (5d)$$



The corresponding sector-averaged turbulence intensities, denoted $\text{SATI}_{4A,i}$, are obtained from Eq. (4b) using the same azimuthal boundaries. Likewise, the sector-effective quantities $\text{SEWS}_{4A,i}$ and $\text{SETI}_{4A,i}$ follow directly from Eqs. (4c) and (4d), respectively. The notation introduced above is subsequently generalized to all sector configurations considered in this study.

275 An alternative reduced-order representation of the inflow was introduced by Schreiber et al. (2020), in which the spatial variability of the rotor inflow is condensed into two shear parameters: the rotor-averaged horizontal shear (RAHS) and rotor-averaged vertical shear (RAVS). In contrast to the sector-based quantities defined previously, which preserve local information in several independent sectors, the RAHS and RAVS descriptors seek to characterize the dominant large-scale inflow gradients using only two scalar quantities. Following Schreiber et al. (2020), these quantities are defined as

$$\text{RAHS} = \frac{3}{2} \frac{(\text{SAWS}_{4A,4,\text{"left"}} - \text{SAWS}_{4A,2,\text{"right"}})}{(\text{SAWS}_{4A,4,\text{"left"}} + \text{SAWS}_{4A,2,\text{"right"}})}, \quad (6a)$$

280

$$\text{RAVS} = \ln \left(\frac{\text{SAWS}_{4A,1,\text{"up"}}}{\text{SAWS}_{4A,3,\text{"down"}}} \right) \left(\ln \left(\frac{z_H + 2/3R}{z_H - 2/3R} \right) \right)^{-1}. \quad (6b)$$

where z_H denotes the hub height and R the rotor radius. The vertical shear parameter, RAVS, is derived by assuming a classical power-law representation of the vertical velocity profile, while the horizontal shear parameter, RAHS, is obtained from a linear approximation of the lateral velocity gradient across the rotor disk.

285 Physically, RAHS provides a compact measure of wake-induced asymmetries, enabling the characterization of partial-wake situations in which one side of the rotor experiences substantially different wind speeds than the other. The quantities RAHS and RAVS may therefore be interpreted as a further reduction of the information contained in the four-sector representation.

Equations (6a) and (6b) assume that the sector-averaged wind speeds of the 4A configuration sample the inflow field at effective locations corresponding to $\pm 2/3R$ along the horizontal and vertical directions, respectively, following the local wind-
 290 speed reconstruction framework proposed by Bottasso et al. (2018).

It should be noted that the proposed inflow descriptors are not intended to reconstruct the complete inflow field experienced by the rotor. Rather, their purpose is to identify the minimum amount of information required to predict fatigue loads with sufficient accuracy. The considered representations therefore span a range of increasing spatial complexity, from rotor-integrated quantities to sector-based descriptions and shear parameters. Together, they provide a framework for investigating one of the
 295 central questions of this study: how much inflow information is truly necessary for accurate and practically deployable load surrogate models?

However, defining reduced-order inflow descriptors is only one aspect of the problem. In practical applications, the quantities contained in i are generally not known exactly and must instead be estimated from available sources of information. The feasibility and accuracy with which a given inflow representation can be obtained may ultimately be just as important as its
 300 ability to predict turbine loads. Consequently, the suitability of an inflow descriptor should not be judged solely by its predictive performance, but also by the practicality of its estimation in the intended deployment environment.

The following section therefore discusses how the inflow quantities introduced above can be obtained from different sources, including simulation-based environments, engineering wake models, and measurement- or observer-based approaches.



2.3.2 Estimation of reduced-order inflow quantities

305 The inflow descriptors introduced in the previous section were defined using the complete rotor-resolved velocity field $u(\psi, r, t)$. In practical applications, however, such information is generally not available, and the quantities contained in $\hat{\mathbf{i}}$ must instead be estimated from available sources of information. To distinguish these estimates from the corresponding reference quantities obtained from the underlying rotor-resolved flow field $u(\psi, r, t)$, an estimated inflow vector is denoted by $\hat{\mathbf{i}}$.

The possible sources of inflow information can be broadly classified into two categories: simulation-based sources and
310 physical measurements.

Simulation-based sources of local inflow mapping include:

- Dynamic flow simulations, such as large eddy simulation (LES) models (Calaf et al., 2010; Churchfield et al., 2012) or dynamic wake meandering (DWM) models (Larsen, 2007). These approaches provide time-resolved estimates of the flow field and can therefore be used to derive both averaged and effective inflow quantities. LES can be performed
315 using dedicated high-fidelity wind-farm simulation frameworks such as SOWFA (Churchfield and Lee, 2012) or AMR-Wind (Kuhn et al., 2025). Validated DWM software implementations include FAST.Farm (Jonkman et al., 2018) and HAWC2Farm (Liew et al., 2023a). The HAWC2Farm DWM implementation has since been incorporated into the broader Dynamiks framework (Pedersen et al., 2023a; Steiner et al., 2026).
- Steady-state engineering wake models, such as the Jensen/Park model (Jensen, 1983), the Bastankhah–Porté-Agel Gaussian wake model (Bastankhah and Porté-Agel, 2016), and various other analytical formulations (Göçmen et al., 2016).
320 These models provide computationally efficient estimates of wake-induced velocity deficits and turbulence levels within a farm and are particularly attractive for large-scale optimization studies. Software frameworks such as PyWake (Pedersen et al., 2023b) and FLORIS (NREL, 2025) provide implementations of such engineering wake models.

Physical sources of local inflow mapping include:

- 325 – Direct flow measurements, obtained for example from scanning lidars (Bodini et al., 2017; Letizia et al., 2023), particle image velocimetry (Li et al., 2020), uncrewed aerial systems (Wetz and Wildmann, 2023), or other flow-sensing technologies. These approaches provide spatially and temporally resolved measurements of the local flow and can therefore provide direct observations of the inflow conditions.
- Indirect measurements based on state-estimation techniques, often referred to as wind observers (Bottasso et al., 2018; Schreiber et al., 2020; Bertelè et al., 2024; Onnen et al., 2026). In this case, the inflow quantities are reconstructed from
330 standard turbine measurements and dynamic models rather than measured directly.

The present work focuses on three representative sources of inflow information. First, dynamic wake simulations are used to generate the reference inflow quantities employed for surrogate-model development, as detailed in Sect. 4. Second, alternative estimates obtained from engineering wake models and wind observers are used to demonstrated the implementation of the
335 proposed surrogate formulation in different environments in Sect. 5.



2.3.3 Consistency between inflow representation and surrogate deployment

It is important to note that the inflow quantities represented by the exact i and the estimated $\hat{i}$ are not necessarily identical. Different sources of inflow information represent the flow field with different levels of fidelity and over different temporal and spatial scales. For example, dynamic wake models explicitly represent wake meandering and turbulent fluctuations, whereas
 340 steady-state engineering wake models describe only mean-flow quantities. Consequently, two sources of inflow information may provide conceptually equivalent descriptors while representing fundamentally different flow dynamics, thereby yielding different values of the reduced-order quantities. This distinction plays an important role in the actual deployment of load surrogate models, as demonstrated in Sect. 5.

From the perspective of load surrogate modeling, the quality of the estimated inflow quantities may be just as important as
 345 the choice of inflow representation itself. An inflow descriptor that exhibits excellent predictive performance when extracted from high-fidelity simulations may become less informative when estimated from lower-order models or measurement-based approaches. Conversely, a simpler descriptor that can be estimated more reliably may ultimately yield superior practical performance.

More generally, consider two alternative inflow representations, denoted by i_1 and i_2 . To make the discussion more concrete
 350 and easier to follow, say for example that i_1 is obtained from the wake meandering flow model used to perform the aero-servo-elastic simulations that yield the DELs; i_2 is instead computed on the lower fidelity engineering model used later for a desired application. For a given load channel l , the surrogate formulation of Eq. (2) may be expressed as

$$\text{DEL}_l = \mathcal{L}_{l,1}(i_1, c), \quad (7)$$

where $\mathcal{L}_{l,1}$ denotes the mapping associated with the inflow representation i_1 . Referring back to the previous example, this
 355 means that the surrogate has been trained on the higher fidelity inflow descriptors.

Although i_1 and i_2 may differ numerically and may originate from substantially different models or measurement systems, as in our example, suppose that they are related through a mapping

$$i_2 = \mathcal{G}(i_1). \quad (8)$$

If this mapping is one-to-one over the considered operating domain, an inverse mapping exists such that

$$360 \quad i_1 = \mathcal{G}^{-1}(i_2), \quad (9)$$

and substituting this relation into the original load mapping yields

$$\text{DEL}_l = \mathcal{L}_{l,1}(\mathcal{G}^{-1}(i_2), c). \quad (10)$$

Defining

$$\mathcal{L}_{l,2}(i_2, c) = \mathcal{L}_{l,1}(\mathcal{G}^{-1}(i_2), c), \quad (11)$$



365 the load surrogate can equivalently be written as

$$\text{DEL}_l = \mathcal{L}_{l,2}(\mathbf{i}_2, \mathbf{c}). \quad (12)$$

In practice this means that, although $\mathbf{i}_2$ is different from $\mathbf{i}_1$, as long as there is a one-to-one relationship between the two, one can feed the surrogate with $\mathbf{i}_2$ and still retrieve the correct DELs. However, crucially, for our example this means that one has to train the surrogate (i.e., learn the $\mathcal{L}_{l,2}(\cdot, \cdot)$ mapping) based on the $\mathbf{i}_2$ inflows of the lower fidelity model. On the contrary, 370 if one were to train the surrogate on the $\mathbf{i}_1$ inflows of the higher fidelity model, then, when it comes to using the model fed by $\mathbf{i}_2$, one would have

$$\text{DEL}_l \neq \mathcal{L}_{l,1}(\mathbf{i}_2, \mathbf{c}), \quad (13)$$

because the learnt mapping is

$$\text{DEL}_l = \mathcal{L}_{l,1}(\mathbf{i}_1, \mathbf{c}) \quad (14)$$

375 but

$$\mathbf{i}_1 \neq \mathbf{i}_2. \quad (15)$$

Therefore, the correct DELs would not be retrieved.

To summarize this long discussion, for our example it is advisable to train the model so that it learns the mapping between the DELs (obviously generated on the higher fidelity meandering model) and the inflows generated by the lower fidelity 380 engineering model. In fact, although it might be tempting to train based on the higher fidelity inflows, at usage time, when the model can only be fed with the lower fidelity inflows, any discrepancy between the two will invariably degrade the quality and correctness of the retrieved DELs.

This important observation provides the mathematical motivation for training and deploying surrogate models using a consistent source of inflow information. When a surrogate is trained using inflow quantities derived from one model or measurement 385 framework and later evaluated using inflow quantities derived from a different framework, the learned mapping implicitly changes, potentially degrading predictive performance. We have used this approach in previous work (Anand et al., 2025; Vad et al., 2026a) and, although we achieved good results, these required careful tuning of the models to ensure the closest possible agreement between the two different representations of the inflow. Such agreement, however, may not always be achievable and can be laborious to obtain. We therefore recommend training and deploying the surrogate within a consistent inflow-description 390 space, allowing it to learn the required transformation directly and thereby avoiding an additional source of model mismatch.

These theoretical considerations provide the foundation for the implementation strategy adopted in this work. Their practical implications are discussed and quantitatively assessed in Sect. 5.

2.4 Mapping of control actions into reduced-order quantities

The formulation of the control descriptor $\mathbf{c}$ follows the same overarching philosophy as that of the inflow descriptor $\mathbf{i}$. In 395 both cases, the objective is to construct a reduced-order representation of a high-dimensional and time-varying process while



retaining the information most relevant to load prediction. Whereas $\mathbf{z}$ provides a compact description of the local inflow conditions encountered by the rotor, $\mathbf{c}$ aims to characterize the turbine operational state and the control actions through which these inflow conditions are converted into aerodynamic loading and structural response.

Several control-related quantities can be expected to influence damage equivalent loads. Yaw misalignment directly modifies the orientation of the rotor relative to the incoming flow and thereby alters the distribution of aerodynamic loading across the rotor disk. Similarly, power-demand signals determine the turbine operating objective and therefore influence the corresponding aerodynamic and structural loading state. Additional information regarding the turbine operating condition is contained in quantities such as rotor speed, blade pitch angle, rotor torque, and aerodynamic power, which collectively describe how the turbine responds to a given inflow condition.

An open question, however, is the level of detail required to represent the turbine operating state for load-surrogate applications. One possibility is that the 10 min mean values of the relevant control and operating quantities provide sufficient information. Alternatively, damage equivalent loads may also depend on the temporal variability of these quantities, since fatigue accumulation originates from load cycles rather than mean loading alone. To account for this possibility, both the 10 min mean values and the corresponding 10 min standard deviations of rotor speed, blade pitch angle, rotor torque, and aerodynamic power are considered as candidate components of $\mathbf{c}$. From a physical perspective, these standard deviations may contain information regarding operating fluctuations and controller activity that is not captured by the corresponding mean values.

To investigate the amount of control-related information required for accurate load prediction, multiple alternative formulations of $\mathbf{c}$ are evaluated. Starting from a minimal representation that only includes yaw misalignment, progressively richer descriptions are obtained through the addition of power-demand signals, mean operating-state quantities, and finally their associated temporal standard deviations. The predictive performance of these alternative formulations is subsequently compared to identify the level of control-related information required for the development of a control-oriented and location-agnostic load surrogate.

2.5 Approximation of the load-mapping function

Various supervised machine-learning techniques can be considered to approximate the unknown load-mapping function $\mathcal{L}_l$, including polynomial response surfaces, polynomial chaos expansions, Kriging and Gaussian-process regression, interpolation-based approaches, tree-based ensemble methods, support vector regression, and artificial neural networks (ANNs) (Dimitrov et al., 2018; Shaler et al., 2022; He et al., 2022; van der Hoek et al., 2026). In the context of wind-energy applications, however, feed-forward ANNs have repeatedly demonstrated accurate approximation of complex nonlinear relationships involving inflow conditions, turbine operating states, and structural loads (Shaler et al., 2022; Vad et al., 2026a). The present study therefore adopts ANN-based surrogate models throughout.

The suitability of ANNs for this application is further supported by the Universal Approximation Theorem, which states that neural networks containing at least one hidden layer and a sufficient number of neurons can approximate arbitrary continuous functions to any desired accuracy under suitable conditions (Hornik et al., 1989). While the theorem does not prescribe a



specific architecture or training procedure, it motivates the use of ANNs as a general-purpose framework for nonlinear function
430 approximation.

A separate surrogate is trained for each load channel considered in this work. Consequently, the adopted framework follows a multiple-input single-output (MISO) architecture, in which the inflow vector i and control vector c jointly form the surrogate inputs, while a single fatigue-load quantity constitutes the output. This choice allows each load channel to be approximated independently and avoids imposing a common functional structure across fundamentally different loading mechanisms.

435 The adopted models are fully connected feed-forward ANNs comprising an input layer, one or more hidden layers using hyperbolic tangent activation functions, and a single linear output neuron. Network weights are optimized using the Adam algorithm (Kingma and Ba, 2015) with the mean-squared error as loss function.

Prior to training, all input and output variables are standardized using statistics computed from the training subset. The available data are randomly partitioned into training and validation sets, with 80 % of the samples used for training and the
440 remaining 20 % reserved for validation.

To quantify surrogate-model performance throughout this study, the normalized root mean square error (NRMSE) is employed, defined as

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2}}{y_{\max} - y_{\min}}, \quad (16)$$

where y_k and $\hat{y}_k$ denote the true and predicted values associated with sample k , N is the total number of samples in the dataset,
445 and $y_{\max}$ and $y_{\min}$ denote the maximum and minimum values of the target variable in the training dataset. The NRMSE is evaluated on the training, validation, and testing datasets and is reported in percentage throughout this work. The normalization range ($y_{\max} - y_{\min}$) is fixed using the training dataset for each load channel.

To assess the statistical robustness of the reported performance metrics, a non-parametric bootstrap procedure is additionally employed. For each evaluated dataset, 1000 bootstrap samples of equal size are generated by random resampling with replace-
450 ment. The NRMSE is computed for each bootstrap realization, yielding an empirical distribution of performance estimates. Reported confidence intervals correspond to the 2.5th and 97.5th percentiles of this distribution, and therefore represent 95 % confidence intervals.

Particular attention is given to hyperparameter optimization (HPO). Because the dimensionality of the surrogate inputs varies substantially across the different formulations investigated in this work, using a single ANN architecture for all cases
455 could introduce an unintended bias in the comparison. Consequently, an independent HPO is performed for every surrogate formulation considered.

The HPO procedure consists of an exhaustive discrete grid search over two architectural hyperparameters: the number of hidden layers and the number of neurons per hidden layer. Hidden-layer counts of 1, 2, 3, and 4 are considered, while neuron counts of 8, 16, 32, and 64 are evaluated, resulting in a total of 16 candidate architectures for each surrogate. Every candidate
460 model is trained for a maximum of 400 epochs using a batch size of 16. Early stopping is employed using the validation loss as monitoring metric, with a patience of 50 epochs and restoration of the best-performing weights.



Model selection is performed using the validation-set NRMSE. The ANN architecture yielding the lowest validation NRMSE is retained as the final surrogate for the corresponding load channel and descriptor formulation. Illustrative examples of the resulting grid-search outcomes are provided in Appendix F.

465 In this work, the surrogate models are implemented using TensorFlow (Abadi et al., 2016). Fixed random seeds and deterministic training settings are employed throughout to ensure fully reproducible model-training and hyperparameter-optimization procedures.

This architecture-specific optimization procedure is important for the interpretation of the results presented in the following sections. Since every candidate definition of the surrogate inputs is paired with its own optimized ANN architecture, the
470 observed performance differences can be attributed primarily to the information content of the selected descriptors rather than to an arbitrary choice of neural-network complexity.

3 Datasets used for the analysis

The surrogate framework proposed in Sect. 2 is intentionally formulated independently of the source of the training data. In principle, the methodology can be trained and applied using any dataset providing the required surrogate inputs and outputs,
475 regardless of whether these originate from numerical simulations, wind-tunnel experiments, or full-scale wind-farm measurements. The proposed formulation therefore does not rely on any specific modelling framework and is intended to be applicable across a broad range of wind-energy applications.

The present work focuses on the design, implementation and validation of the surrogate methodology using simulation data. Numerical simulations provide direct access to local inflow conditions throughout the wind farm together with detailed
480 aero-servo-elastic response measurements, making them particularly well suited for the systematic investigation of surrogate design choices considered in this study. At the same time, an important objective throughout the development is to maintain a formulation that remains sufficiently generic to be applied to alternative data sources in future work.

For this reason, the surrogate methodology presented in Sect. 2 is separated from the specific datasets employed herein. The remainder of this section describes the generation of the training, validation and testing datasets used in the present study,
485 before they are subsequently employed for surrogate design, implementation and evaluation in Sect. 4 and Sect. 5.

3.1 Training and validation datasets

The data-driven surrogate models introduced in Sect. 2.5 require a training dataset covering a broad range of inflow conditions, wake interactions, and wind-farm control actions. To generate this dataset, aero-servo-elastic wind-farm simulations are performed using FAST.Farm (Jonkman et al., 2018). Three IEA 3.4 MW reference wind turbines (Bortolotti et al., 2019) are
490 arranged in a row and subjected to varying atmospheric conditions, turbine spacings, wind directions, yaw misalignment angles, and power-demand set-points, as illustrated on Fig. 3.

The use of three turbines allows the generated database to include free-stream conditions as well as single- and double-wake interactions. Although larger wind farms may exhibit more complex wake patterns and, in some cases, influence from more

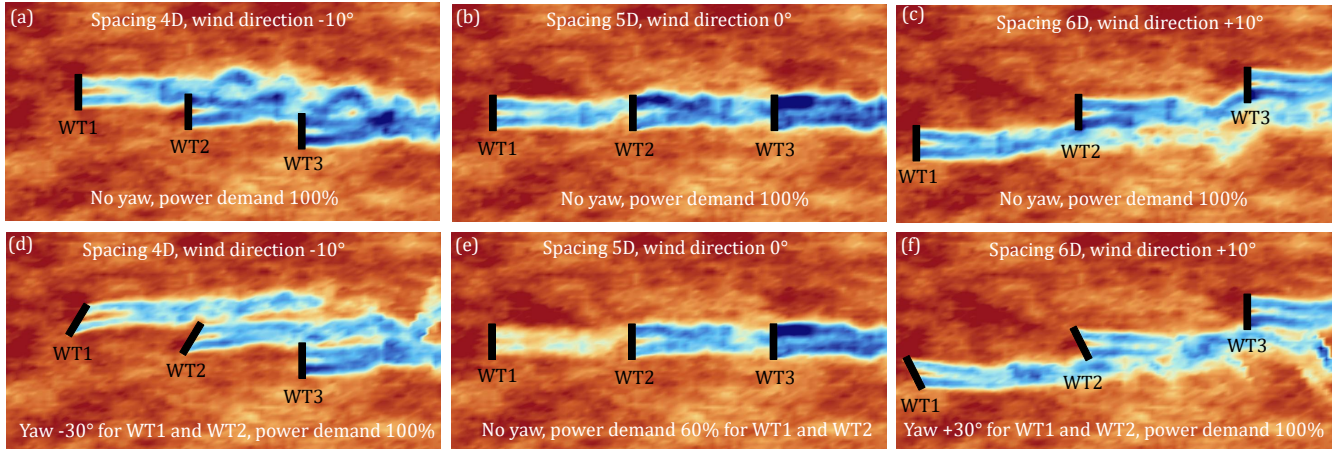


Figure 3. Illustrative examples of six operating and layout configurations considered in the three-turbine FAST.Farm simulations used to generate the training and validation datasets. Turbine spacing, wind direction, yaw misalignment angle, and power-demand set-points vary across the cases, as indicated in each subpanel. The background shows an instantaneous snapshot of the hub-height streamwise velocity field computed with FAST.Farm (Jonkman et al., 2018) and visualized using ParaView (Ayachit, 2015).

than two upstream turbines, the present configuration already generates a broad range of local inflow conditions, wake deficits, turbulence levels, and wake-overlap scenarios at the turbine level. Since the surrogate models are formulated using exclusively local inflow and control descriptors, rather than information about the global wind-farm layout, increasing the number of turbines in the training farm would not necessarily introduce fundamentally new input-output relationships for the surrogate to learn. The present configuration therefore provides a computationally efficient framework for generating representative local operating conditions and associated structural loads. The ability of the resulting surrogates to generalize beyond this training configuration is later assessed using an independent five-turbine wind farm presented in Sect. 3.2.

The aero-servo-elastic time-domain response of each turbine is simulated using the AeroDyn, ElastoDyn and ServoDyn modules of the OpenFAST framework (NREL, 2026). Wake interactions within the wind farm are modelled using FAST.Farm (Jonkman et al., 2018), which is based on the DWM framework originally proposed by Larsen (2007). The DWM approach provides a computationally efficient representation of wake transport, recovery and meandering while retaining the dominant unsteady mechanisms governing turbine loading within wind farms. This computational efficiency is an essential requirement of the present work, where the generation of training, validation and testing datasets requires tens of thousands of simulation cases, making higher-fidelity coupled LES–OpenFAST approaches computationally impractical. The FAST.Farm implementation has been extensively validated against large eddy simulations for both wake dynamics and structural loading (Shaler and Jonkman, 2021), and against field measurements of fatigue loading (Kretschmer et al., 2021), providing confidence in its suitability for the generation of surrogate-training datasets. The specific FAST.Farm implementation employed in the present work corresponds to the wake-added turbulence formulation developed by Kretschmer et al. (2021). In addition to the stan-



dard DWM representation of wake deficits and meandering, this formulation accounts for wake-added turbulence, which has been shown to contribute significantly to fatigue loading in wind farms. The default DWM model parameters provided with FAST.Farm are used throughout this work.

515 Farm-scale ambient turbulent inflow fields are generated using the standard Kaimal spectral turbulence model implemented in TurbSim (Jonkman, 2009). The considered simulations focus exclusively on normal atmospheric and operating conditions representative of conventional wind-farm operation. Extreme atmospheric events, gusts, low-level jets, veer, shutdowns and fault scenarios are outside the scope of the present work.

A particular emphasis is placed on the convergence of load statistics with respect to atmospheric turbulence realizations. 520 While the use of six turbulent seeds is common practice in many single-turbine studies (International Electrotechnical Commission, 2019), substantially larger variability can be observed when dynamic wake interactions are included. To ensure robust estimation of both surrogate inputs and outputs, every operating condition is therefore simulated using 24 independent turbulent realizations. The influence of the number of turbulent seeds on surrogate performance is investigated explicitly in Sect. 4.5.

The baseline turbine controller is the Fortran-based DISCON controller distributed with the IEA 3.4 MW reference turbine 525 (Bortolotti et al., 2019). In the present work, the controller was modified in-house to include curtailment operation through a prescribed power-demand set-point. Curtailment is implemented by triggering entry into Region III operation once the desired power level is reached, with the power-demand set-point expressed as a percentage of rated power. The resulting changes in rotor speed, blade pitch angle, thrust coefficient and power curve are detailed in Appendix B. For completeness, the modified controller was benchmarked against the state-of-the-art ROSCO implementation presented by Abbas et al. (2022). Excellent 530 agreement was obtained in terms of power production, rotor speed, blade pitch angle and, most importantly, DELs, providing confidence that the implemented controller accurately represents current industrial control behaviour.

Yaw-misalignment control is represented through prescribed nacelle orientations. The yaw degree of freedom is disabled in ElastoDyn, such that the nacelle orientation remains fixed throughout each 10 min simulation. Desired yaw misalignment angles are imposed directly through the turbine initial conditions rather than through a dynamic yaw-control system.

535 Figure 3 illustrates several representative examples from the training simulation database. The complete considered parameter ranges are summarized in Table 1.

Table 1. Parameter ranges used to generate the training and validation datasets considered in this study.

Category	Parameter	Values / Range
Atmospheric	Ambient wind speed	6–12 m s ^{−1}
	Ambient turbulence intensity	6 and 12 %
	Vertical shear exponent	0–0.4
Wind-farm configuration	Wind direction	−20° to +20°
	Turbine spacing	2.5–6 D
Control	Yaw misalignment	−30° to +30°
	Power demand	5–100 %



The turbine spacing listed in Table 1 is applied simultaneously between WT1–WT2 and WT2–WT3, as illustrated in Fig. 3. In contrast, yaw misalignment angles and power-demand set-points are assigned independently to WT1 and WT2, thereby generating a wide variety of controlled wake conditions affecting downstream turbines. Power demand is expressed as a percentage of rated power.

The selected atmospheric conditions are intentionally concentrated below and around rated operation, rather than uniformly spanning the entire turbine operating envelope. This choice reflects the fact that aerodynamic wake losses are particularly pronounced in the below-rated to near-rated regime (Baas et al., 2023; Porchetta et al., 2026), while wake-control strategies have also demonstrated substantial benefits in this operating range (Simley et al., 2021; Meyers et al., 2022).

A full factorial combination of all parameters would lead to a prohibitively large number of simulations. Therefore, a subset of 10 261 operating conditions is selected to provide broad coverage of the parameter space while limiting the computational cost. Considering all three turbines separately, the resulting database contains 30 783 turbine-level samples linking local inflow descriptors, control variables, and structural-load responses.

The complete list of simulated cases, together with the corresponding input parameters and post-processed outputs, is provided through the publicly available dataset accompanying this work (Guilloré et al., 2026). The distributions of the resulting surrogate inputs and outputs are analysed in Sect. 3.3.3.

A key aspect of the proposed methodology is that no information regarding the absolute position of a turbine within the wind farm is provided to the surrogate. The training samples consist exclusively of local inflow descriptors and control quantities. After extracting all turbine-level samples, the complete dataset is randomly shuffled prior to training. This ensures that the surrogate learns relationships between local operating conditions and structural loads, rather than farm-specific spatial configurations, thereby enabling location-agnostic predictions. As described more specifically in Sect. 2.5, the full 30 783 turbine-level samples are split into 80 % used for training and 20 % reserved for validation.

3.2 Testing dataset

To assess the generalization capabilities of the proposed surrogate models, an independent testing dataset is generated using a wind-farm configuration that differs substantially from the three-turbine layouts employed during training. The objective is to evaluate whether the surrogates remain accurate when exposed to local inflow conditions arising in a larger wind farm and under combinations of wake interactions not explicitly represented during training.

The testing dataset is generated using FAST.Farm following the same modelling assumptions and simulation settings as those employed for the training and validation datasets. In contrast to the aligned three-turbine configurations used during training, the testing wind farm consists of five IEA 3.4 MW turbines arranged in an irregular layout, as shown in Fig. 4. The selected configuration intentionally creates a broad range of operating conditions, including free-stream turbines, partial wake interactions, and multiple wake overlaps.

While the wind farm layout and wind direction remain fixed to the configuration shown in Fig. 4, the atmospheric and control cases are varied according to the ones summarized in Table 2.

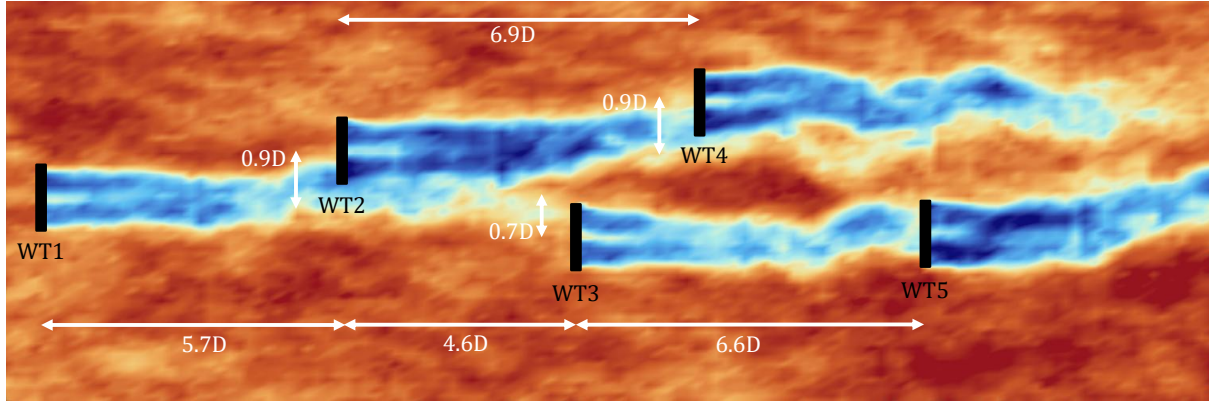


Figure 4. Visualization of the independent five-turbine wind farm used to generate the testing dataset. All distances are expressed in terms of rotor diameters (D). The background shows an instantaneous snapshot of the hub-height streamwise velocity field computed with FAST.Farm (Jonkman et al., 2018) and visualized using ParaView (Ayachit, 2015).

Table 2. Parameter ranges used to generate the testing dataset considered in this study. The wind farm configuration (wind direction and inter-turbine spacings) is limited to the one shown in Fig. 4.

Category	Parameter	Values / Range
Atmospheric	Ambient wind speed	8, 10 and 12 m s^{-1}
	Ambient turbulence intensity	6 %
	Vertical shear exponent	0.2
Control	Yaw misalignment	$-30^\circ : 10^\circ : +30^\circ$
	Power demand	40:10:100 %

570 The yaw misalignment angles and power-demand set-points listed in Table 2 are assigned independently to WT1, WT2 and WT3, thereby generating a wide variety of controlled wake conditions affecting downstream turbines. A total of 147 wind-farm operating cases are simulated, each including 24 independent turbulent realizations, consistently with the training and validation datasets. Considering all five turbines, the resulting database contains 735 turbine-level samples for testing the surrogates. The complete list of testing cases is provided through the publicly available dataset accompanying this work (Guilloré et al., 2026).

575 The distributions of the resulting surrogate inputs and outputs are analysed in Sect. 3.3.3.

No sample belonging to the testing dataset is used at any stage of the surrogate training or validation process. The testing cases differ from the training dataset in terms of wind-farm layout, number of turbines, and resulting wake interaction patterns. Consequently, the reported testing results evaluate the ability of the surrogates to predict structural loads for previously unseen wind-farm configurations. A key advantage of the proposed formulation is that these markedly different wind-farm scenarios can still be described through local inflow quantities that remain within the surrogate input space, enabling robust predictions across a broad range of layouts without requiring training on each specific farm configuration.



3.3 Extraction of surrogate inputs and outputs

Once the training, validation and testing simulations have been completed, the resulting time-series are post-processed into reduced-order quantities suitable for surrogate modelling. The objective of this post-processing step is to transform the high-dimensional inflow and load measurements into a set of compact descriptors and target quantities that can be efficiently handled by the surrogate models.

3.3.1 Reduced-order inflow and control descriptors

The reduced-order inflow quantities defined in Sect. 2.3 are extracted using virtual sensors embedded within the DWM environment. More specifically, nine virtual sensors measuring the undisturbed wind speed are placed along each blade within the AeroDyn module of OpenFAST (NREL, 2026). Their radial locations are selected such that each sensor represents an equal rotor area, while their average radial position corresponds approximately to two-thirds of the blade span, following the methodology of Bottasso et al. (2018).

In a post-processing step, the measurements from the 27 rotating sensors are transformed into a fixed frame of reference to reconstruct the local inflow field $u(\psi, r, t)$ across the rotor disk. The reduced-order inflow descriptors introduced in Sect. 2.3 are subsequently computed from this reconstructed inflow field, resulting in the complete family of rotor- and sector-based inflow quantities investigated in this work.

Figure 5 illustrates a representative partial-wake realization extracted from the training database together with the corresponding SAWS and SATI descriptors for all nine sector configurations introduced in Fig. 2. The considered example corresponds to a wake impingement occurring predominantly on the right-hand side of the rotor disk (when viewed looking downstream). Similar inflow conditions are encountered, for example, by WT2 and WT3 in the configuration shown in panel (c) of Fig. 3, as well as by WT2 and WT4 in the testing wind farm shown in Fig. 4.

Figure 5 highlights how increasing the number of sectors provides a progressively finer description of the inflow heterogeneity experienced across the rotor disk. The refinement ranges from simple rotor-averaged quantities (equivalent to SAWS 1 and SATI 1) to the most detailed eight-sector discretization considered in this work. Alternative sector counts and sector arrangements could also be envisioned without fundamentally modifying the methodology. The resulting reduced-order quantities are supplied to the surrogate as input features of increasing complexity, with the dimensionality of the inflow vector $\mathbf{i}$ increasing as the number of sectors is refined, from 2 components for rotor-averaged quantities up to 16 components for the eight-sector discretization. A key objective of this study is therefore to determine the minimum level of inflow discretization required to obtain accurate and robust load predictions, a question addressed in Sect. 4.4.

Other approaches considered for estimating the required inflow quantities, namely wind observers and steady-state engineering wake models, are discussed separately in Sect. 5.1 and Sect. 5.2, respectively.

The reduced-order control descriptors considered as candidates for vector $\mathbf{c}$ are obtained either directly from the prescribed simulation inputs (yaw misalignment angle and power-demand set-point) or from selected turbine-response signals. For the latter, the 10 min mean value and corresponding 10 min standard deviation of rotor speed, blade pitch angle, rotor torque and

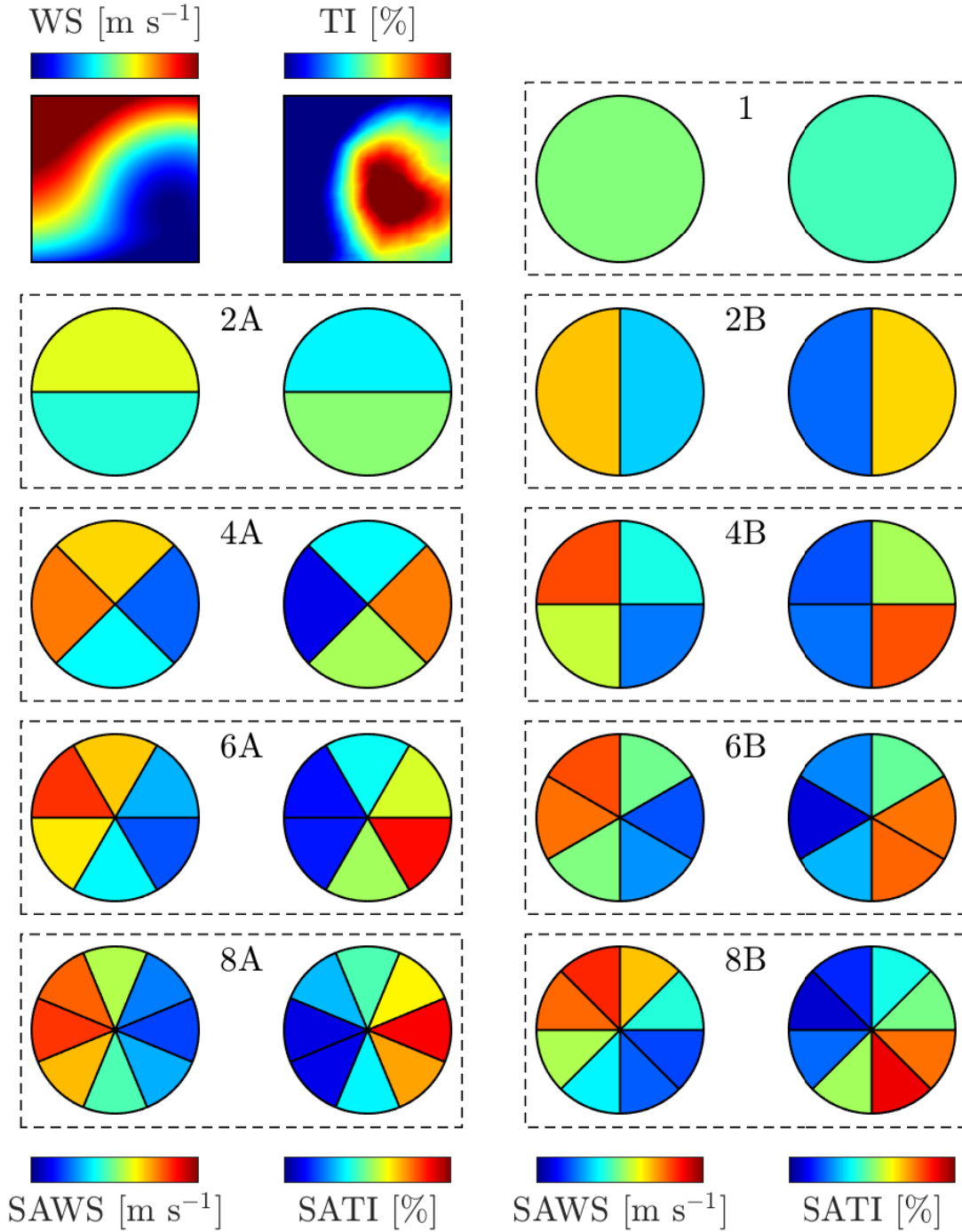


Figure 5. Application of the sector discretizations introduced in Fig. 2 to a representative partial-wake realization extracted from the generated dataset. Wake impingement occurs predominantly on the right-hand side of the rotor disk, in this example. For each sector configuration, the resulting $SAWS$ are displayed on the left and the corresponding $SATI$ on the right. The view is looking downstream, i.e. facing the rotor disk.



615 aerodynamic power are computed and used as reduced-order control parameters. The relative merits of these different inputs are investigated in Sect. 4.3.

3.3.2 Considered fatigue load channels

The proposed surrogate methodology is applicable to any fatigue load quantity available from measurements or numerical simulations. In the present work, four projected DEL channels are considered. These quantities are derived from eight underlying
 620 load signals corresponding to four structural locations, each characterized by two orthogonal load components.

For each load location, the two orthogonal load components are projected onto directions spanning the corresponding cross-sectional plane with an angular resolution of 10° . A DEL is then computed for each projected load time series using Eq. (1), and the maximum value among all projection directions is retained. This projected DEL represents the direction experiencing the highest fatigue damage and provides a scalar quantity representative of the most critical loading condition at the considered
 625 location. Furthermore, this projection procedure removes the dependence on the orientation of the local coordinate system and reduces each pair of orthogonal load components to a single scalar fatigue metric.

The four projected DEL channels used throughout this study, together with their underlying orthogonal load components, are:

- DEL Blade root projected: blade root in-plane and out-of-plane bending moments.
- 630 – DEL Shaft projected: shaft out-of-plane and yaw bending moments.
- DEL Yaw bearing projected: tower top fore-aft and side-side bending moments.
- DEL Tower base projected: tower bottom fore-aft and side-side bending moments.

In practice, the direction associated with the maximum projected DEL depends on the operating condition and therefore cannot be represented by a single angular location. For the blade root and tower base, the maximum projected DEL is typically
 635 aligned close to the in-plane and fore-aft directions, respectively, owing to the larger amplitudes of these load components. For the shaft and yaw bearing, the two orthogonal components are of comparable magnitude, causing the direction of maximum projected DEL to vary substantially with the operating condition, particularly with yaw misalignment.

3.3.3 Distributions of surrogate inputs and outputs

Figure 6 summarizes the distributions of selected surrogate inputs and outputs across the training, validation and testing
 640 datasets. The displayed quantities correspond to representative inflow and control descriptors used throughout the study, together with the four projected DEL channels considered as surrogate outputs.

Overall, the testing dataset exhibits distributions that are consistent with those of the training and validation datasets while remaining sufficiently different to provide a meaningful assessment of surrogate generalization capabilities. This reflects the distinct layout, wake-interaction patterns and control setpoints of the testing wind farm compared to the training and validation
 645 setup. In particular, several test cases are located close to the boundaries of the training distributions, especially for the SAWS

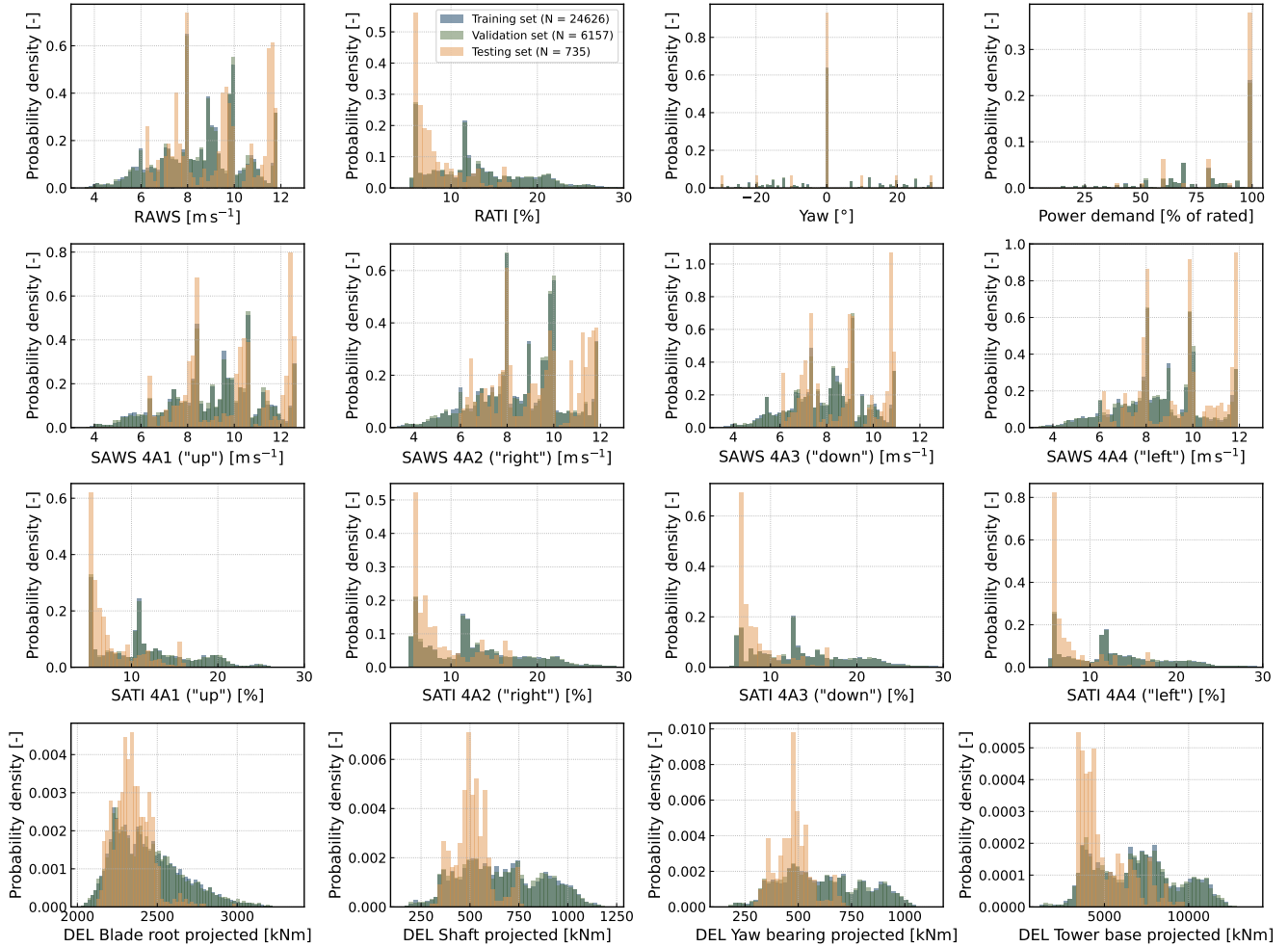


Figure 6. Distributions of selected surrogate inputs and outputs across the training, validation and testing datasets. The displayed inflow descriptors correspond to the rotor-averaged case (option 1) and the four-sector discretization (option 4A) introduced in Figs. 2 and 5, together with the yaw-misalignment and power-demand set-points. The four projected DEL channels used as surrogate outputs are shown in the bottom row.



and SATI descriptors, while remaining fully contained within the range covered by the training and validation datasets. Consequently, the proposed surrogates are evaluated under a challenging yet physically meaningful interpolation regime, probing regions of the input space close to the limits encountered during training while avoiding extrapolation beyond the supported domain.

650 The distributions of the yaw-misalignment and power-demand set-points are strongly concentrated around 0° and 100%, respectively, reflecting the predominance of baseline operation in the simulated datasets. Nevertheless, non-zero yaw-misalignment and power-curtailment conditions remain well represented. The apparent imbalance is partly due to the aggregation of samples from all turbines, whereby controlled upstream turbines and nominally operated downstream turbines similarly contribute to the dataset. This is further accentuated by the fact that yaw-misalignment and power-curtailment set-points are often varied
655 independently, such that curtailed turbines frequently contribute samples at 0° yaw misalignment and vice versa.

4 Impact of design choices on load surrogate model performance

This section quantifies the influence of key surrogate design parameters on prediction accuracy and uses the resulting observations to derive practical recommendations for the development of load surrogate models intended for wind-farm applications.

The analysis is organized as follows. Section 4.1 compares alternative spatio-temporal inflow representations and identifies
660 suitable reduced-order descriptors of the rotor-scale wind field. Section 4.2 evaluates whether effective inflow quantities provide advantages over their averaged counterparts. Section 4.3 examines different representations of turbine control set-points. Section 4.4 investigates the influence of the number and positioning of inflow sectors used to characterize rotor-scale inflow heterogeneity. Section 4.5 evaluates the impact of the number of turbulent realizations used to estimate the target fatigue loads employed during surrogate training. Finally, Section 4.6 assesses the sensitivity of surrogate performance to the size of the
665 training database. The resulting observations are used to derive practical recommendations for the development of accurate and computationally efficient load surrogate models.

4.1 Influence of the spatio-temporal inflow representation

This first analysis investigates how the choice of inflow representation influences the ability of the surrogate to generalize across the wide range of operating conditions encountered within a wind farm, including unwaked, partially waked, and fully
670 waked situations. The influence of control actions is intentionally excluded at this stage in order to isolate the effect of the inflow mapping.

To this end, a reduced subset of the dataset is considered for which the turbines operate on the maximum power curve, corresponding to zero yaw misalignment and a power demand of 100 %. As a result, the training dataset is reduced from 24 626 to 6 559 cases, the validation dataset from 6 157 to 1 640 cases, and the testing dataset from 735 to 447 cases.

675 The predictive performance obtained with different strategies for reducing the local inflow conditions into low-dimensional representations is presented in Fig. 7. The investigated representations correspond to increasing levels of complexity in the description of the rotor inflow field and are based on the quantities defined in Sect. 2.3. In the present analysis, all inflow



representations are formulated using averaged quantities. The performance implications of adopting effective quantities instead are investigated subsequently in Sect. 4.2.

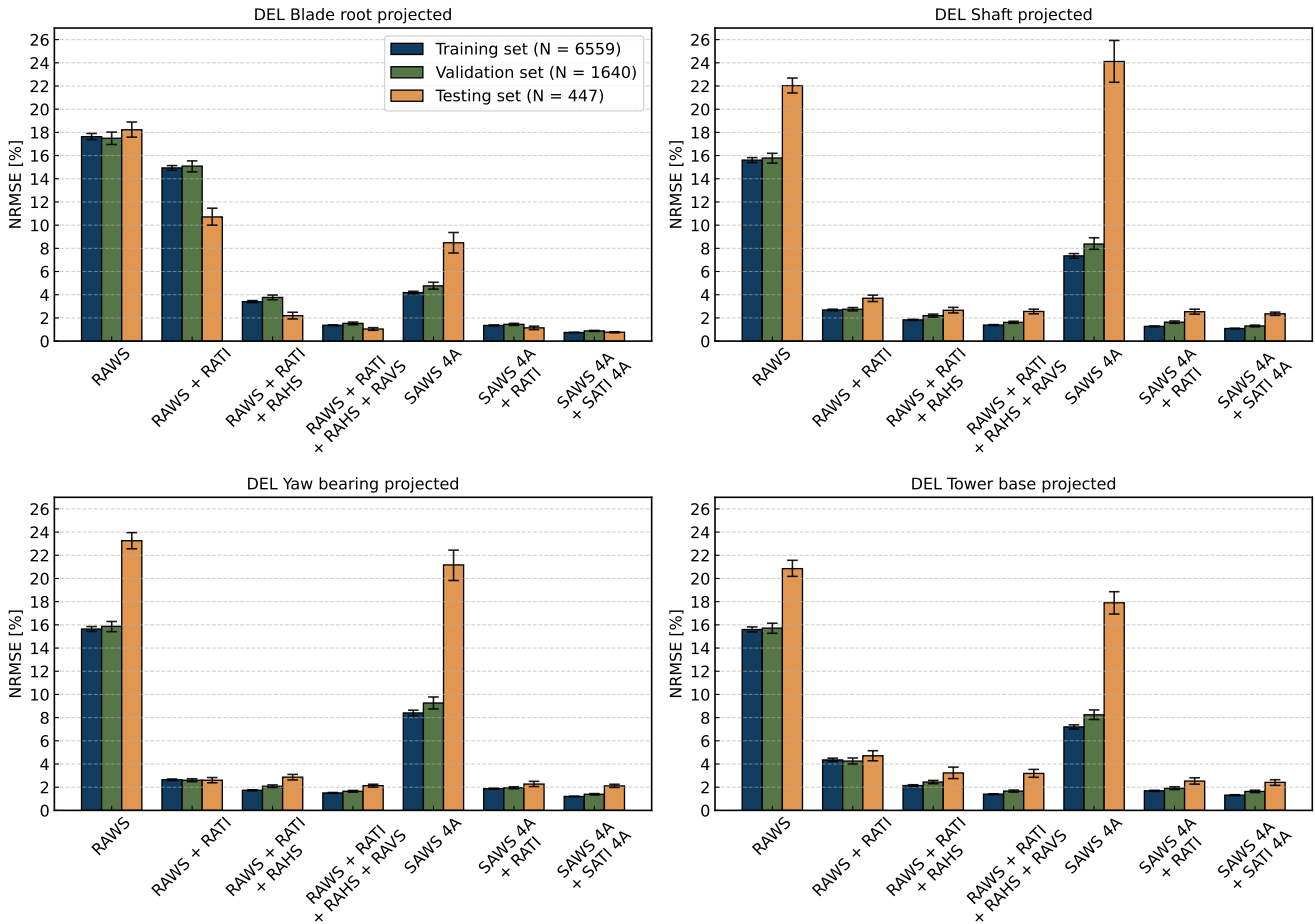


Figure 7. Performance of surrogates obtained using different strategies for mapping the spatio-temporal inflow field into reduced-order parameters based on averaged inflow quantities. The bars show the deterministic NRMSE computed on the full training, validation, and test sets, respectively. The whiskers indicate the corresponding 95% confidence intervals estimated using non-parametric bootstrapping (1000 resamples with replacement).

Overall, the results in Fig. 7 reveal a clear and consistent trend across all considered load channels: prediction errors decrease as additional information describing the rotor-scale inflow field is incorporated into the surrogate inputs. Equally important, the relative ranking of the investigated inflow representations remains remarkably consistent across the training, validation, and testing datasets. The resulting agreement in NRMSE values indicates that the identified trends are not specific to a particular



subset of operating conditions and provides strong evidence that the proposed surrogates generalize well to unseen inflow
685 situations.

The nature of the most beneficial inflow descriptors depends on the considered load channel. For blade-root DELs, the largest reduction in NRMSE is obtained when information describing rotor-scale inflow heterogeneity is introduced through RAHS, RAVS, or SAWS 4A. This behavior is physically consistent with the fact that rotating blades repeatedly encounter different inflow conditions while traversing full and partial wakes. Blade fatigue loads are therefore strongly governed by spatial variations of the inflow field across the rotor disk. In contrast, the shaft, yaw-bearing, and tower-base DELs exhibit their largest improvements once turbulence-intensity information is included through RATI or SATI 4A, indicating a stronger
690 sensitivity to temporal inflow fluctuations and the integrated turbine response.

This distinction is illustrated by the SAWS 4A representation. Although it relies exclusively on wind-speed information, the introduction of four sector-averaged wind-speed quantities substantially reduces blade-root prediction errors compared to purely rotor-averaged representations. However, the absence of turbulence-intensity information still leads to noticeably larger
695 errors for the shaft, yaw-bearing, and tower-base DELs. The results therefore indicate that surrogate models intended to predict multiple load channels require information describing both rotor-scale velocity heterogeneity and the overall turbulence level experienced by the turbine.

An interesting observation is that the RAHS/RAVS formulation proposed by Schreiber et al. (2020) achieves performance very similar to the four-sector representation previously introduced by Guilloré et al. (2024). This similarity can be understood by the fact that RAHS and RAVS are directly derived from the same four sector-averaged wind-speed quantities according to Eqs. (6a) and (6b). The comparable prediction accuracies suggest that, for the considered operating conditions and load metrics, fatigue loads are influenced more strongly by the magnitude of the rotor-scale asymmetries than by the precise wind-speed values observed within each sector. Nevertheless, the direct SAWS 4A representation yields slightly lower NRMSE
705 values while avoiding an additional level of abstraction and is therefore retained in the remainder of this work.

The results further show that augmenting SAWS 4A with a single rotor-averaged turbulence-intensity parameter provides nearly all achievable performance gains. In particular, SAWS 4A + RATI performs almost identically to SAWS 4A + SATI 4A, despite the latter incorporating a much more detailed spatial description of turbulence intensity. This suggests that, once rotor-scale wind-speed heterogeneity has been represented explicitly, additional spatial refinement of the turbulence-intensity field
710 provides little additional information for the prediction of the considered fatigue loads. From a practical perspective, this is particularly attractive because rotor-averaged turbulence intensity is considerably easier to estimate from reduced-order wake models than its spatial distribution across the rotor disk.

The associated 95 % bootstrap confidence intervals remain narrow for all considered datasets and become progressively smaller as the inflow representation becomes more informative. For the best-performing configurations, the confidence intervals
715 are barely distinguishable from the corresponding NRMSE bars, indicating limited sensitivity of the reported performance metrics to dataset resampling and providing further evidence of the statistical robustness of the observed trends.

Overall, SAWS 4A + RATI emerges as the most attractive inflow representation among the investigated alternatives. It consistently achieves near-optimal performance across all considered load channels while remaining substantially simpler than



SAWS 4A + SATI 4A. The combination therefore provides the best compromise between predictive accuracy, robustness,
 720 interpretability, and implementation complexity, and is selected as the recommended inflow representation for the proposed
 location-agnostic load surrogate.

Figure 8 provides additional insight into the effect of the inflow representation by comparing measured and predicted DELs
 for three representative surrogate configurations of increasing complexity: RAWs, RAWs + RATI, and SAWS 4A + RATI. For
 clarity, the training, validation, and testing datasets are merged into a single set containing all 8 646 operating conditions for
 725 the evaluation presented on Fig. 8.

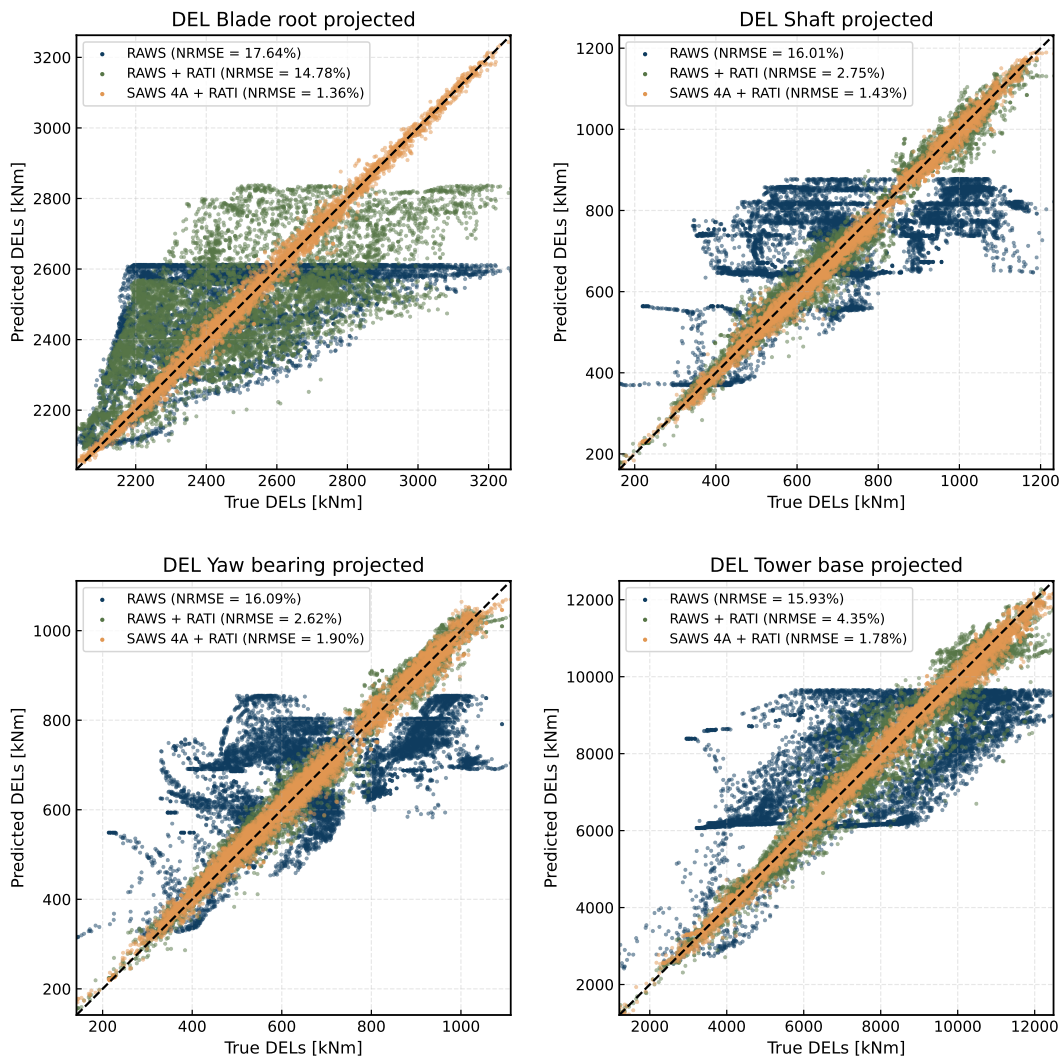


Figure 8. Scatter plots from load surrogates with increasing level of spatio-temporal descriptors for the local wind field. Here, all three (training, validation + testing) set are merged together (for the 8646 cases that do not have any special control set-points, i.e. full power).



The progression from RAWs to RAWs + RATI highlights the important role of turbulence intensity for several load channels. In particular, the shaft, yaw-bearing, and tower-base DELs already exhibit a clear clustering around the diagonal once RATI is introduced. This observation is consistent with the trends reported in Fig. 7 and confirms that a significant fraction of the variability of these integrated turbine loads can be explained through the combined knowledge of the rotor-averaged wind speed and rotor-averaged turbulence intensity. The subsequent transition from RAWs + RATI to SAWS 4A + RATI further tightens the cloud of points around the one-to-one line, demonstrating that additional information describing rotor-scale inflow heterogeneity still improves the prediction accuracy, although to a lesser extent than for blade loads.

A markedly different behavior is observed for the blade-root DELs. Even after the introduction of RATI, the RAWs + RATI representation remains unable to reproduce the full range of observed fatigue loads. The scatter plots reveal a pronounced upper limit in the predicted DELs, visible as a horizontal ceiling of prediction values. Consequently, many operating conditions associated with elevated blade loads are systematically underpredicted. This behavior indicates that a representation based solely on rotor-averaged quantities lacks sufficient information to distinguish between inflow fields producing similar mean wind speeds and turbulence intensities but substantially different levels of rotor-scale asymmetry.

The situation changes considerably once the four sector-averaged wind-speed quantities are introduced through the SAWS 4A + RATI representation. The artificial ceiling observed for RAWs and RAWs + RATI largely disappears, and the predicted blade-root DELs become much more closely aligned with the one-to-one line. The additional sector information allows the surrogate to recognize whether elevated fatigue loads originate from heterogeneous inflow conditions across the rotor disk, such as those encountered during partial-wake operation. In this sense, the scatter plots provide direct evidence that rotor-scale flow heterogeneity is a primary driver of blade fatigue loads and cannot be adequately represented through rotor-averaged inflow quantities alone.

Overall, the scatter plots confirm and illustrate further the conclusions previously drawn from the NRMSE analysis. Rotor-averaged quantities combined with turbulence intensity already provide an adequate description of the inflow conditions governing the shaft, yaw-bearing, and tower-base DELs. In contrast, accurate prediction of blade fatigue loads requires explicit information on the spatial distribution of wind speed across the rotor disk. The SAWS 4A + RATI representation therefore provides the most balanced description of both temporal and spatial inflow characteristics, explaining its superior and more consistent performance across all considered load channels.

4.2 Averaged vs. effective inflow representation

The previous subsection established the influence of the level of complexity used to describe the rotor inflow field when averaged quantities are adopted. A related question is whether the predictive performance of the surrogate can be further improved by replacing these averaged quantities with their effective counterparts. To address this question, the different inflow representations introduced in Sect. 2.3 are reformulated using effective quantities and compared against their averaged counterparts.

To provide additional context for this comparison, the values of the averaged and effective inflow quantities defined in Sect. 2.3 are reported in Appendix C using scatter plots constructed from the combined training and validation datasets. As expected, the averaged and effective wind-speed quantities (RAWs versus REWS and SAWS versus SEWS) are nearly identi-



cal, whereas noticeable differences are observed between the corresponding turbulence-intensity quantities (RATI versus RETI and SATI versus SETI). These differences arise from the distinct averaging operations involved in their definitions and indicate that averaged and effective turbulence-intensity descriptors cannot be considered interchangeable. The key question, therefore, is whether either formulation provides a systematic advantage when employed as an input to the load surrogate model. The results of this comparison are presented in Fig. 9.

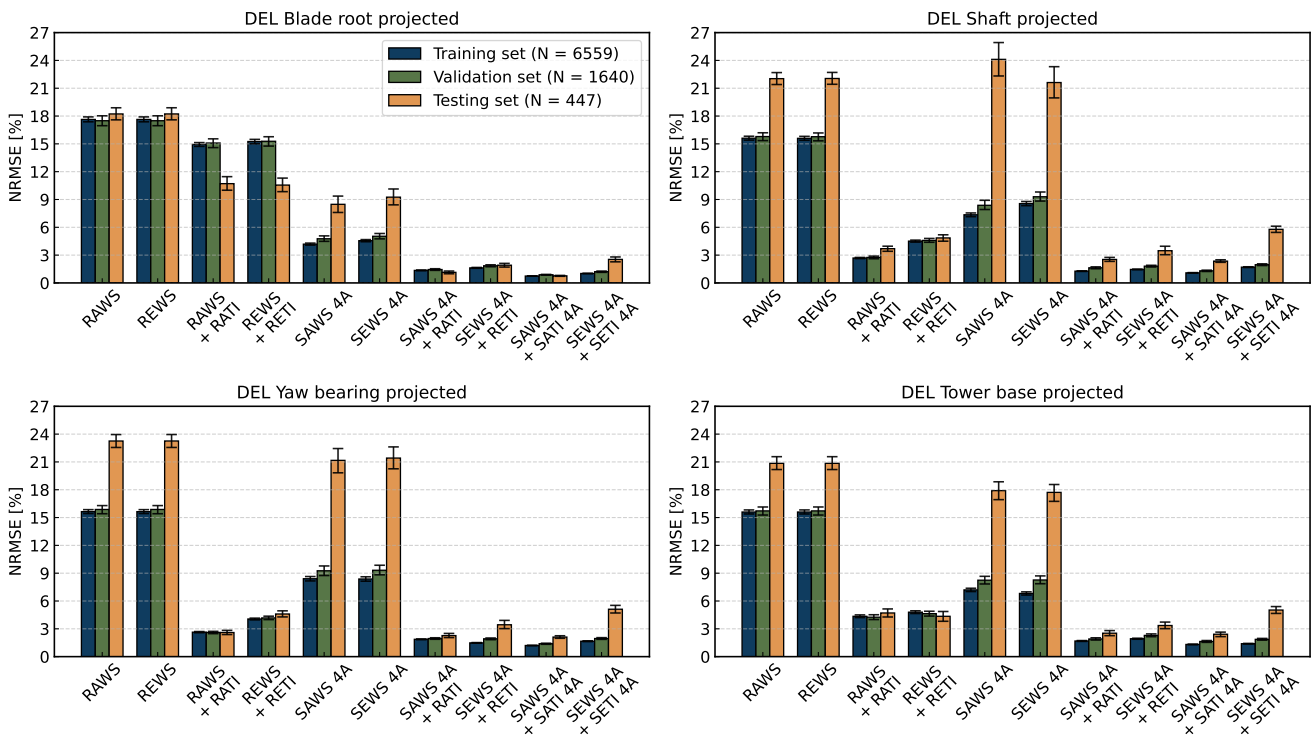


Figure 9. Performance of surrogates for different strategies of spatio-temporal inflow mapping into averaged vs. effective quantities. The bars show the deterministic NRMSE computed on the full training, validation, and test sets, respectively. The whiskers indicate the corresponding 95% confidence intervals estimated using non-parametric bootstrapping (1000 resamples with replacement).

What can be observed in Fig. 9 is that, for each considered inflow representation, the surrogate performance obtained using effective inflow quantities closely tracks those obtained using the corresponding averaged quantities. The differences in NRMSE are generally small and remain within the uncertainty levels estimated by bootstrapping. This result is not necessarily intuitive, given that averaged and effective quantities are constructed through fundamentally different operations and may exhibit noticeably different numerical values, particularly for the turbulence-intensity descriptors, as seen in Appendix C.

The similarity in predictive performance nevertheless indicates that both formulations contain largely equivalent information from the perspective of the surrogate model. In other words, although quantities such as RATI and RETI or SATI and SETI are



not numerically identical, the machine-learning model is able to learn an equally accurate mapping between these descriptors and the target DELs. This observation suggests that no single formulation can be regarded as superior for load-surrogate development. Instead, the consistency between training and application appears to be of greater importance than the specific choice between averaged and effective quantities.

The conclusions drawn in the previous subsection therefore remain essentially unchanged when effective quantities are employed. In particular, the progressive improvement observed when enriching the inflow representation from rotor-averaged to sector-based descriptors is preserved, irrespective of whether averaged or effective quantities are used.

These findings provide additional flexibility for practical implementations. For example, the combination SAWS 4A + RATI may be viewed as more consistent with steady-state engineering flow-model approaches (Crespo and Hernández, 1996; Ishihara and Qian, 2018; NREL, 2025; Pedersen et al., 2023b), whereas SEWS 4A + RETI is conceptually closer to observer-based frameworks (Bottasso et al., 2018; Schreiber et al., 2020; Bertelè et al., 2024; Onnen et al., 2026) in which rotor-averaged wind-speed time series are first reconstructed and subsequently transformed into reduced-order descriptors. The comparable performance of these alternatives suggests that either formulation can be employed successfully for load-surrogate modelling, provided that the selected definition is used consistently during both surrogate training and deployment.

For the remainder of this study, the averaged formulation (SAWS 4A + RATI) is adopted because it yields slightly lower NRMSE values and is more physically intuitive. This choice also simplifies the presentation of the subsequent analyses.

4.3 Influence of control set-points

The previous subsections addressed the definition of the inflow descriptor $\hat{i}$. The present subsection focuses on the definition of the control vector c .

The analysis is now extended to the complete dataset, including cases with yaw misalignment and curtailment. Consequently, all data distributions shown in Fig. 6 are considered here and throughout the remainder of the paper. The two upper right panels of Fig. 6 show the distributions of the yaw-misalignment and power-demand operating conditions introduced in this analysis.

The question addressed in this subsection is which control information is required to accurately capture the influence of control actions on DELs while maintaining a compact surrogate formulation. Figure 10 compares surrogate-model performance for different definitions of c . In all cases, the inflow conditions are represented by the previously selected formulation SAWS 4A + RATI, while the control vector is progressively enriched to provide increasing information about the turbine operating state and applied control actions, as defined in Sect. 2.4.

The results indicate that – as expected – neglecting the control vector c altogether leads to poor predictive performance once cases involving yaw misalignment and curtailment are included in the training, validation, and test datasets. In contrast to the analyses presented previously, the same inflow conditions may now correspond to multiple turbine operating states and therefore to substantially different DELs. Although the inflow descriptor SAWS 4A + RATI provides an accurate representation of the local flow conditions, it does not contain sufficient information to distinguish between the operating states associated with different yaw and curtailment settings. Consequently, markedly different DELs may be associated with similar inflow conditions, resulting in comparatively large prediction errors.

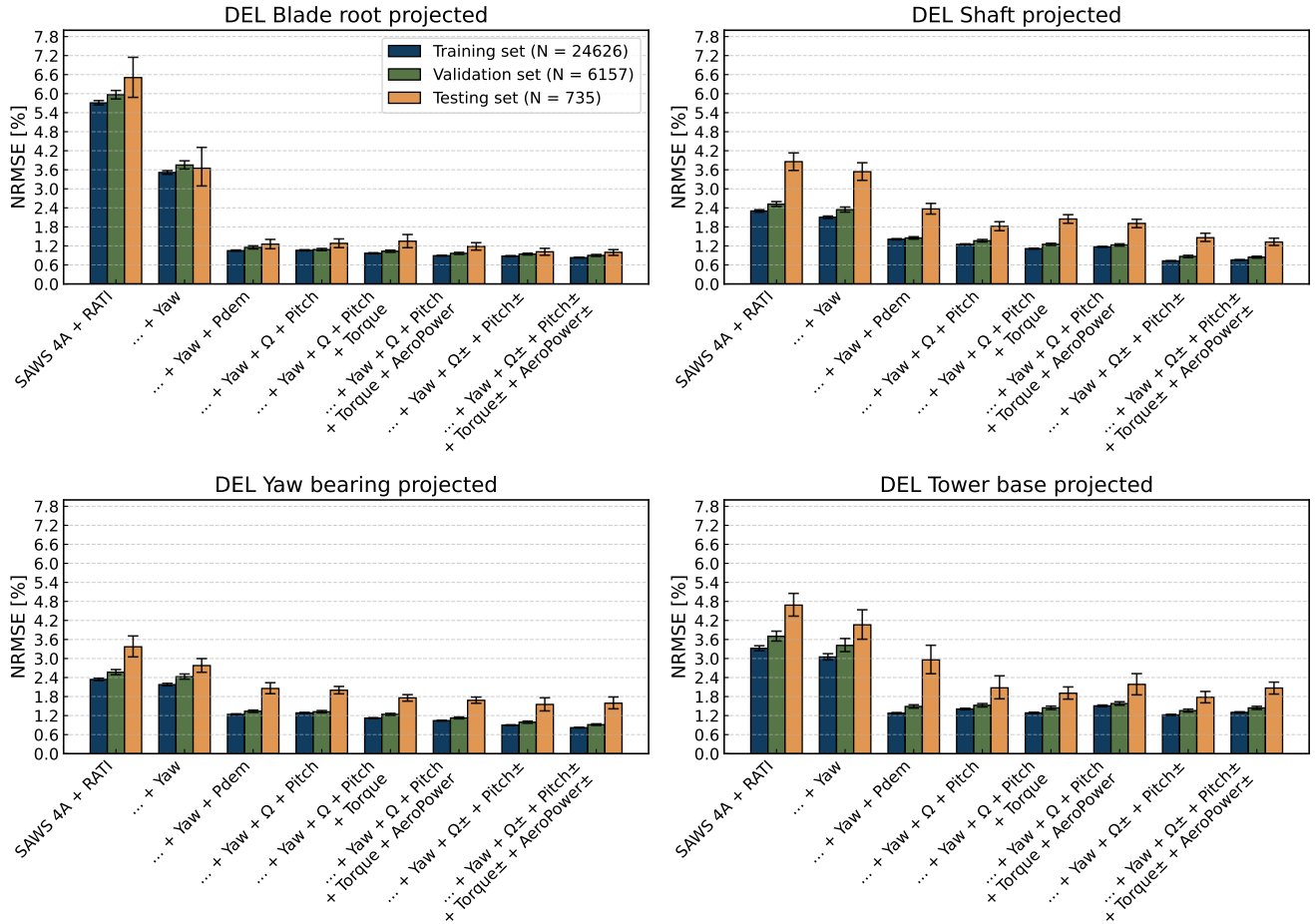


Figure 10. Performance of surrogates for different strategies of representing the control vector $\mathbf{c}$. Here all cases, including yaw misalignments and various levels of curtailment, are included in the training, validation, and test datasets. The inflow is represented using SAWS 4A + RATI in all cases. In the labels, Ω denotes the rotor speed, while the symbol $\pm$ indicates that the corresponding 10 min standard deviation is included in addition to the mean value of the quantity. The bars show the deterministic NRMSE computed on the full training, validation, and test sets, respectively. The whiskers indicate the corresponding 95% confidence intervals estimated using non-parametric bootstrapping (1000 resamples with replacement).



As in the previous analyses, the NRMSE values obtained on the training, validation, and test sets closely follow one another for all surrogate formulations considered. Furthermore, the associated bootstrap confidence intervals remain extremely narrow and centered around the deterministic NRMSE values. This consistency indicates that the observed trends are highly robust and not driven by a small subset of samples, thereby providing confidence that the conclusions drawn from Fig. 10 remain valid across a wide range of resampled datasets.

Introducing the yaw-misalignment angle improves the surrogate performance for all load channels, confirming that this control state must be represented explicitly. Incorporating the power-demand signal leads to a further improvement, as it enables the surrogate to distinguish between operating conditions associated with different levels of curtailment. Once both yaw misalignment and power demand are included, most of the ambiguity introduced by the controlled operating conditions is resolved, and the prediction errors approach those obtained with considerably richer definitions of c .

The impact of including c is particularly pronounced for the blade-root DELs, which exhibit NRMSE values of approximately 6% when no control information is provided. This behavior is physically consistent with the strong sensitivity of blade-root loading to rotor asymmetries induced by yaw misalignment and to changes in rotor speed associated with different curtailment levels. By contrast, the remaining load channels already achieve NRMSE values of approximately 3–4% without any explicit control mapping, indicating that they are predominantly driven by the characteristics of the inflow itself. Nevertheless, incorporating c remains essential, as it reduces the prediction errors for all load channels to approximately 1.5–2.5%, thereby approaching the best performance levels obtained in the present study.

Further enrichment of the control vector through the addition of rotor speed, blade pitch, rotor torque, and aerodynamic power statistics produces little further reduction in NRMSE, indicating that the surrogate performance has largely reached a plateau. Interestingly, the same conclusion remains valid when the corresponding 10 min standard deviations are included as additional inputs. Although these quantities provide information regarding the dynamic turbine response and controller activity beyond the mean operating state, they contribute only marginal improvements in predictive accuracy. This suggests that the dominant influence of the control actions is already captured by the mean operating-condition parameters. From a practical perspective, this result is particularly noteworthy because obtaining quantities such as rotor speed, blade pitch, rotor torque, aerodynamic power, and especially their temporal standard deviations would require substantially more information than is typically available from reduced-order steady-state flow models. Given the limited reduction in NRMSE achieved by these additional inputs, their inclusion appears difficult to justify for the present application.

An interesting observation can be made by comparing Fig. 10 with Fig. 7. The NRMSE range considered here extends only to approximately 7.8%, whereas Fig. 7 spans values reaching approximately 26%. This difference highlights the dominant role of the inflow vector i in the overall surrogate accuracy. Indeed, improving the inflow mapping reduces the NRMSE from values of approximately 20% to about 3%, whereas the introduction of c further reduces the errors from approximately 6% to about 2% once controlled operating conditions are included. These observations should not be interpreted as diminishing the importance of control descriptors, since accurate representations of both i and c are required to achieve low prediction errors. Rather, they indicate that the largest source of uncertainty originates from the description of the inflow, while the control mapping provides the additional information required to resolve the variability associated with different turbine operating modes.



The results indicate that a favorable balance between predictive accuracy and descriptor complexity is already achieved using a control vector composed only of yaw misalignment and power demand. Although more elaborate formulations provide marginal accuracy gains, the prediction errors have effectively reached a plateau, suggesting that the dominant influence of the control actions has already been captured. From this perspective, the additional variables provide refinement rather than
845 fundamentally new predictive capability.

This observation is particularly relevant from an implementation perspective. In practical applications, the surrogate is intended to operate in conjunction with engineering flow models, which typically provide estimates of local inflow and wake conditions and can therefore provide mean power, but do not contain turbine operational variables such as rotor speed, blade pitch, rotor torque, or their temporal standard deviations. Requiring such information would therefore increase the complex-
850 ity of the overall framework and partly undermine the objective of maintaining a compact and broadly applicable surrogate formulation.

One limitation of this formulation is that the power-demand signal is directly linked to the particular curtailment strategy used to generate the training data, which is illustrated in Appendix B. If the surrogate is expected to be applied across multiple curtailment implementations, a more generic description of the turbine operating state is desirable. In such cases, a formulation
855 based on yaw misalignment together with mean rotor speed and mean blade-pitch angle constitutes an attractive alternative, as these quantities can be interpreted as control-agnostic parameters of the turbine operating condition.

For the remainder of this work, the formulation based on yaw misalignment and power demand is retained. This option represents the simplest control description considered while still providing satisfactory predictive performance across all investigated load channels. It therefore offers the most favorable compromise between surrogate accuracy, implementation simplicity, and
860 practical applicability.

4.4 Influence of the number and position of sectors

The previous subsection established a suitable formulation for the control vector $\mathbf{c}$, based on yaw misalignment and power demand. Having now identified both the inflow descriptor type and control representation, the present analysis focuses on a practical design question: how many inflow sectors are required to achieve converged surrogate performance? While Sect. 4.1
865 demonstrated the benefits of sector-averaged wind-speed descriptors under baseline operating conditions, the present analysis revisits this question using the complete dataset, including yaw-misalignment and curtailment cases.

As described in Sect. 2.5, an independent grid-search HPO optimization is performed for every surrogate formulation considered throughout this study. This step is particularly important in the present analysis, since the dimensionality of the surrogate inputs varies substantially across the candidate formulations. Depending on the number of sectors used to define $\mathbf{i}$, the total
870 number of inputs ranges from only a few variables to more than a dozen. Illustrative examples of the resulting grid-search outcomes for the present sector-discretization study are provided in Appendix F. The optimal ANN architectures vary between formulations, reflecting the differing complexity of the associated learning tasks. Consequently, the differences reported in this section can be attributed to the physical information contained in the alternative definitions of $\mathbf{i}$ rather than to an artefact of ANN architecture selection.



875 Following the conclusions of Sect. 4.2, the turbulence-intensity representation is fixed to a single rotor-averaged value (RATI) throughout the present analysis. Since SAWS 4A + RATI was found to provide performance nearly identical to SAWS 4A + SATI 4A, only the spatial discretization of the sector-averaged wind-speed representation is varied here. This choice also improves practical applicability, as rotor-resolved turbulence-intensity distributions are not always available in all reduced-order steady-state inflow models.

880 The analysis below therefore focuses exclusively on the influence of the number and position of SAWS sectors used to describe the inflow state. Nine sector configurations are considered, ranging from the rotor-averaged representation (1) to increasingly refined discretizations (2A, 2B, 4A, 4B, 6A, 6B, 8A, and 8B), as illustrated in Fig. 2 and Fig. 5. In all cases, turbulence intensity is represented by RATI, while the control vector is fixed to the combination of yaw misalignment and power demand identified in Sect. 4.3.

885 The blade root DEL exhibits the strongest dependence on the sector discretization. Starting from the rotor-averaged formulation (RAWS + RATI), the NRMSE remains close to 9%, confirming that rotor-averaged inflow quantities are insufficient to characterize the inflow features governing blade fatigue loading. The most significant reduction in NRMSE is observed only when the inflow representation becomes capable of resolving left-right asymmetries across the rotor. While the SAWS 2A configuration provides little improvement relative to the rotor-averaged formulation, the SAWS 2B configuration yields a substantial reduction in prediction error. This behavior is physically consistent with the dominant role of partial-wake conditions, which primarily manifest themselves as lateral inflow non-uniformities across the rotor disk. Moving to four sectors results in another marked reduction in error and marks the onset of a clear performance plateau. The SAWS 4A and SAWS 4B configurations both achieve blade-root DEL NRMSE values close to 1–1.5%, while the training, validation, and test results closely follow one another throughout the entire convergence study.

890 The same overall convergence trend is observed for the shaft, yaw-bearing, and tower-base DELs. Starting from the rotor-averaged formulation, progressively richer sector-based representations lead to lower prediction errors, while little additional benefit is obtained beyond four sectors. Consequently, the SAWS 4A and SAWS 4B configurations again define the onset of a clear performance plateau, indicating that the dominant load-relevant inflow features are already captured by a relatively coarse spatial discretization of the rotor disk.

895 Although this general behavior is remarkably consistent across all load channels, some differences can nevertheless be observed. For the shaft and yaw-bearing DELs, the distinction between the 2A and 2B sector orientations remains comparatively modest, and the improvement in NRMSE occurs relatively gradually as the inflow description is enriched from RAWS to SAWS 2A, SAWS 2B, and finally the four-sector representations. The tower-base DEL exhibits a stronger sensitivity to the sector orientation, with SAWS 2B providing a noticeably larger reduction in NRMSE than SAWS 2A. This observation suggests that horizontal wake-driven inflow heterogeneity may play a more important role for tower-base loading than for the shaft and yaw-bearing channels. A possible explanation is that wake meandering, which predominantly develops in the lateral direction, introduces inflow fluctuations that are more effectively captured by a left-right partition of the rotor than by an up-down partition.

900

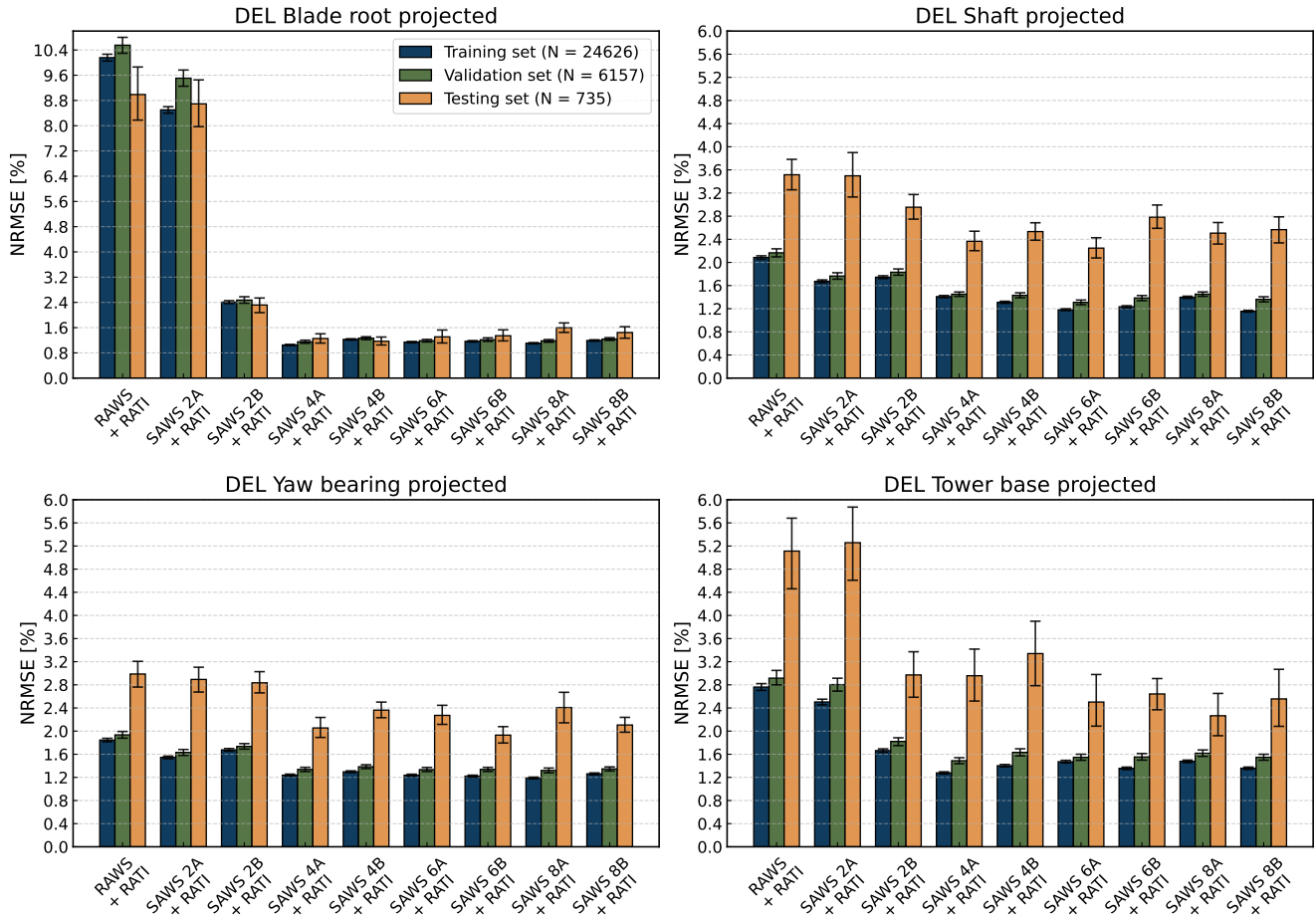


Figure 11. Performance of surrogates by increasing number and position of sector-averaged wind speeds (SAWS) and fixed single rotor-averaged turbulence intensities (RATI). The sectors are illustrated in Fig. 2 and Fig. 5. The bars show the deterministic NRMSE computed on the full training, validation, and test sets, respectively. The whiskers indicate the corresponding 95% confidence intervals estimated using non-parametric bootstrapping (1000 resamples with replacement).



Taken together, the results reveal a remarkably consistent convergence behavior across all four load channels. In all cases, moving from a rotor-averaged representation to a four-sector description yields substantial improvements in predictive performance, whereas further increases in sector resolution provide little additional benefit. In several cases, the training and validation NRMSE values continue to decrease slightly when moving from four to six or eight sectors, while the corresponding test-set errors remain unchanged or even increase. Although such behavior is common in machine-learning applications, it is particularly noteworthy in the present context because the test set consists exclusively of simulations performed in the independent five-turbine wind farm shown in Fig. 4. The results therefore suggest that increasing the number of inflow descriptors beyond four sectors may improve the fit to configurations represented in the training data while providing little benefit, or even slightly degrading performance, when the surrogate is applied to substantially different flow situations. This convergence behavior constitutes an important result of the present study, as it demonstrates that highly refined rotor-flow descriptions are not required to construct accurate load surrogates. Rather, a small number of appropriately distributed inflow descriptors already captures most of the load-relevant information contained in the rotor-resolved flow field. This conclusion clearly holds limitedly to the inflow conditions considered here. The inclusion of – for example – veer and low-level jets might lead to different conclusions.

Finally, the SAWS 4A and SAWS 4B configurations achieve remarkably similar performance across all load channels. This finding indicates that once four measurements distributed across the rotor are available, the surrogate is able to infer the relevant mapping between inflow heterogeneity and structural loading largely independently of the precise sector orientation. From a physical perspective, both four-sector layouts provide sufficient information to characterize the dominant horizontal and vertical non-uniformities associated with wake interactions, ambient shear, and yawed operation. At the same time, the similarity between the two configurations is reassuring because it demonstrates that the observed convergence is not tied to a unique or special sector arrangement. Nevertheless, SAWS 4A consistently yields slightly lower NRMSE values than SAWS 4B, particularly on the independent test set. Given these small but systematic advantages, together with its intuitive up-right-down-left interpretation, the SAWS 4A + RATI formulation is retained in the remainder of the study as the preferred compromise between predictive accuracy, descriptor compactness, and practical applicability.

While the previous figures quantified surrogate performance through aggregate NRMSE values, it does not illustrate how the different inflow representations reproduce the underlying variations of the fatigue loads across individual operating conditions. To provide additional insight, Fig. 12 compares surrogate predictions against reference DEL values for a representative subset of cases extracted from the independent five-turbine testing wind farm shown in Fig. 4.

Along the horizontal axis, the yaw-misalignment angles of WT1, WT2, and WT3 are varied simultaneously. This generates a continuous range of operating conditions involving both changing control actions and changing wake-deflection and wake-overlap patterns throughout the wind farm. The figure therefore provides a complementary perspective to the NRMSE analysis by illustrating how the surrogate predictions evolve across a broad range of inflow and control conditions.

Three representative inflow formulations are compared, selected based on the conclusions of the previous analyses: RAWs, RAWs + RATI, and SAWS 4A + RATI. In all cases, the control vector c includes yaw misalignment and power demand. It should be noted that the operating conditions shown in Fig. 12 are not artificial examples chosen to exaggerate surrogate



945 differences. Rather, they correspond to realistic variations of turbine loads arising from changes in wake steering within the wind farm. While the ambient wind speed and turbulence intensity remain fixed, the changing yaw settings modify wake deflection and wake interactions, thereby giving rise to the load variations shown in the figure.

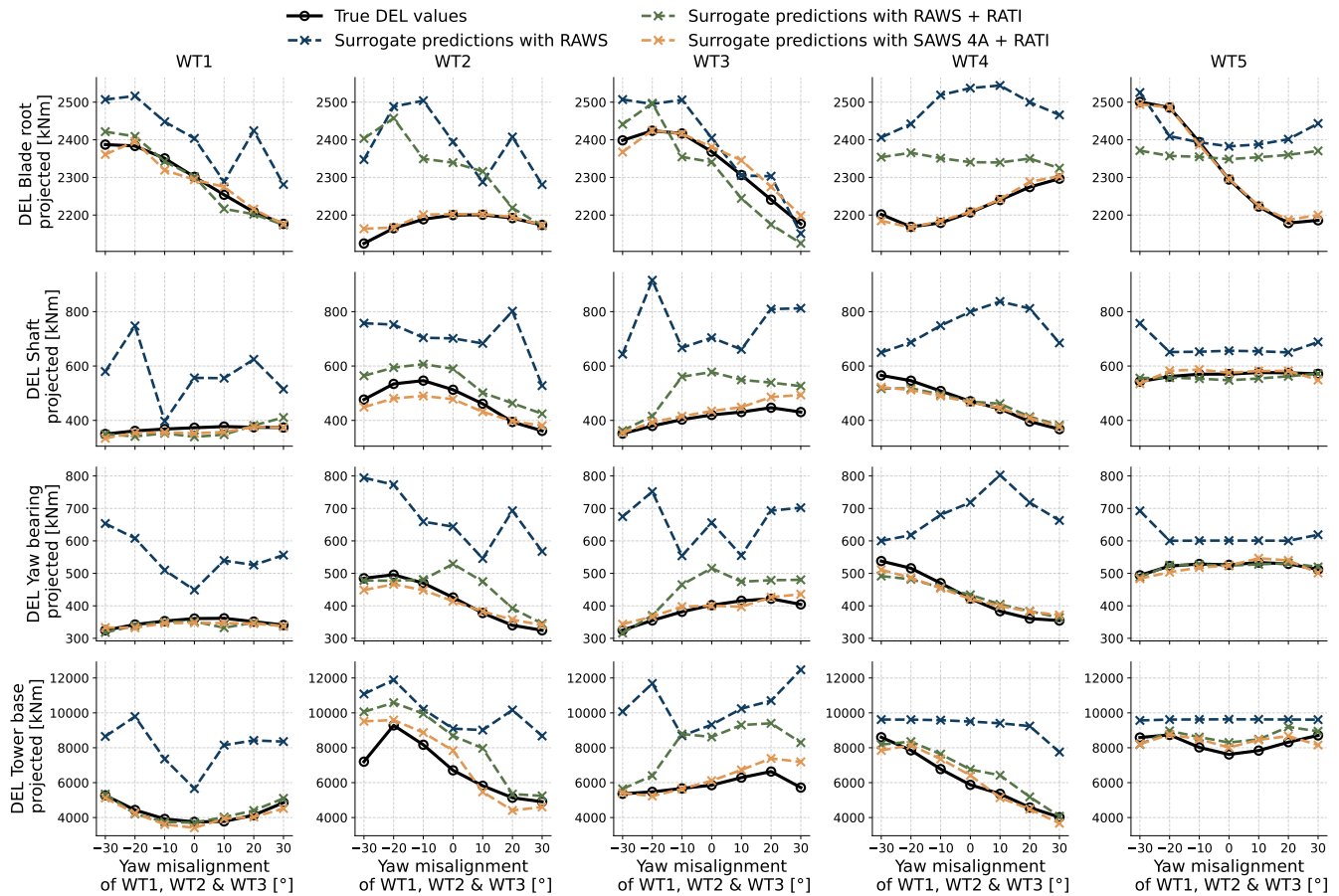


Figure 12. Illustrative comparison between surrogate predictions (dashed lines) and reference values (solid black lines) obtained on the independent testing wind farm for three representative load-surrogate formulations. Rows correspond to the four load channels, while columns correspond to the five turbines. The turbine locations are shown in Fig. 4. Along the horizontal axis, the yaw misalignment angles of WT1, WT2, and WT3 are varied simultaneously, inducing different wake-deflection and wake-overlap conditions within the wind farm.

950 For WT1, WT2, and WT3, the surrogate must simultaneously capture the influence of the local inflow conditions and the applied control actions, since these turbines are actively yawed throughout the considered operating sequence. The good agreement between surrogate predictions and reference values therefore demonstrates the ability of the proposed framework to account for both inflow-induced and control-induced variations in fatigue loading. For WT4 and WT5, which are not actuated directly, the observed load variations arise primarily from changes in the local inflow caused by wake deflection and wake



interactions generated by the upstream turbines. Accurate predictions for these downstream turbines therefore provide a direct assessment of the effectiveness of the inflow representation itself.

The conclusions previously drawn from the NRMSE analysis are strongly supported by the load trends displayed in Fig. 12. For the blade-root DELs, pronounced variations are observed as the wake trajectories evolve across the wind farm. These variations are reproduced accurately only by the SAWS 4A + RATI formulation. For WT1, the RAWS + RATI formulation already provides reasonable predictions, which is unsurprising given that this turbine operates outside of wake influence. For the downstream turbines, however, the limitations of rotor-averaged inflow quantities become evident. In particular, the RAWS + RATI surrogate often predicts nearly constant blade-root DELs despite the substantial variations observed in the reference data. Since this formulation contains no explicit information about the spatial distribution of the inflow across the rotor disk, it cannot adequately distinguish between different levels of wake impingement. The addition of the four sector-averaged wind-speed quantities resolves this limitation and enables the surrogate to capture the detailed load variations associated with wake deflection and partial-wake operation.

A different behavior is observed for the shaft, yaw-bearing, and tower-base DELs. Consistently with the previous analyses, these load channels are governed to a greater extent by the overall turbulence level experienced by the turbine. Consequently, the RAWS + RATI formulation already captures a substantial fraction of the observed trends, as RATI implicitly contains information regarding wake-added turbulence and wake-meandering effects through a single rotor-averaged quantity. Nevertheless, the most accurate agreement with the reference DELs is again obtained using SAWS 4A + RATI. Across all turbines and operating conditions, this formulation provides the most consistent reproduction of both the magnitude and evolution of fatigue loads. The results therefore further support the conclusion that a compact representation combining four sector-averaged wind speeds with a single rotor-averaged turbulence-intensity parameter captures the dominant spatial and temporal inflow characteristics governing fatigue loading.

4.5 Influence of the number of turbulent seeds

The analyses presented so far have considered training, validation, and testing datasets generated using averages over 24 turbulent realizations for each operating condition. An important practical question is therefore how many turbulent seeds are actually required to obtain a sufficiently accurate load surrogate model. Since the generation of additional turbulent realizations represents a significant computational cost, identifying the minimum number of seeds required for convergence is of direct practical relevance.

To address this question, surrogate models are trained using progressively increasing numbers of turbulent seeds, while their predictive performance is evaluated against reference quantities computed using all 24 available seeds. In other words, for a given number of seeds N_{seed} , the surrogate inputs and target DEL values are first averaged over only N_{seed} turbulent realizations during training. The resulting surrogate predictions are then compared against the corresponding 24-seed averaged quantities in Fig. 13. Within the scope of the present analysis, the 24-seed averages are therefore treated as the highest-fidelity reference quantities available to quantitatively study the convergence behavior of the formulated surrogate.



985 Because multiple combinations of turbulent realizations can be selected for a given value of N_{seed} , the analysis is repeated using ten independent random selections of seeds. The solid lines in Fig. 13 represent the mean NRMSE obtained from these ten realizations, while the shaded envelopes indicate the corresponding minimum and maximum values.

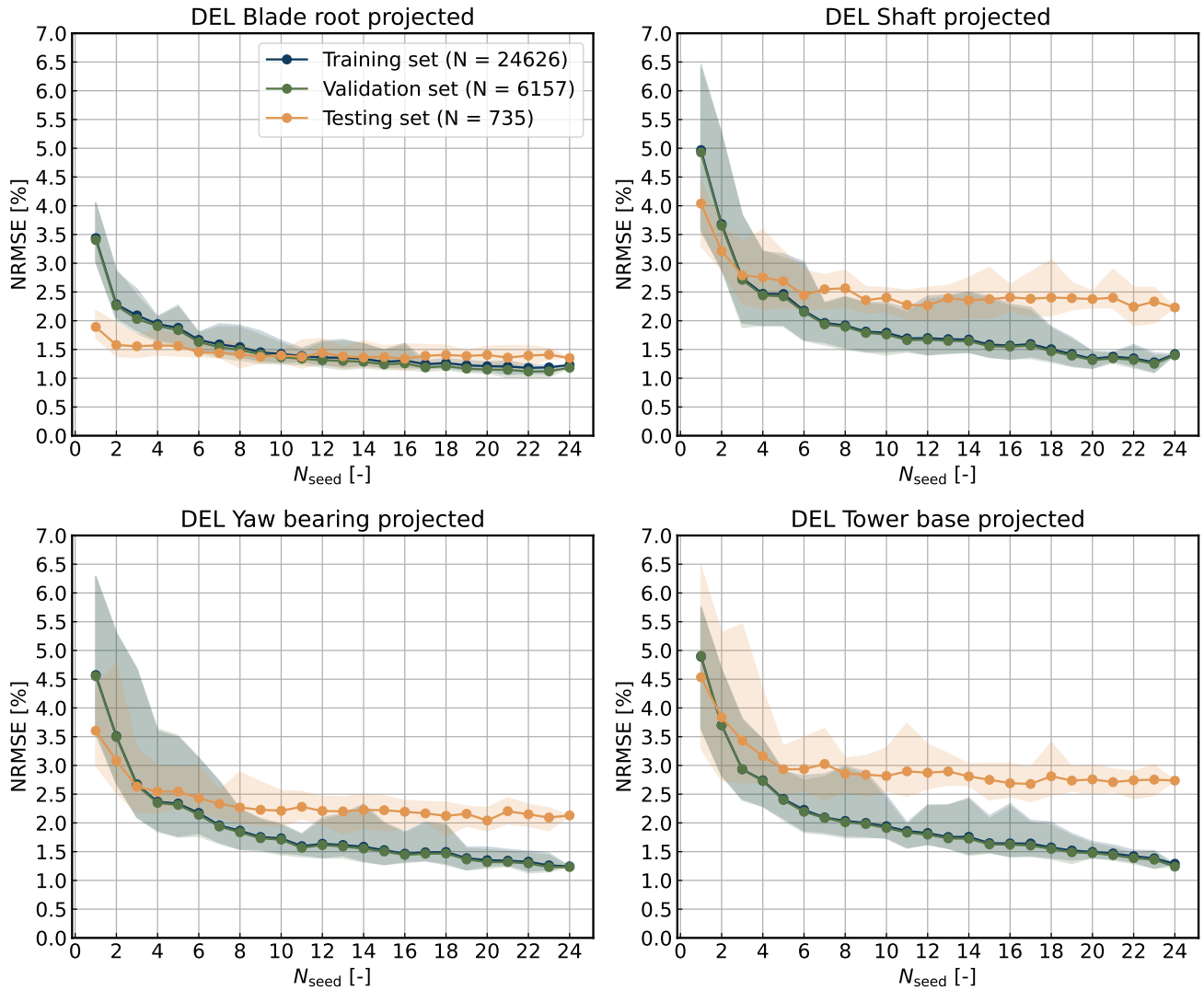


Figure 13. Performance of surrogates as a function of the number of turbulent seeds used to construct the training quantities. The inflow is represented using SAWS 4A + RATI. At each point, surrogate predictions are evaluated against reference quantities obtained from averages over all 24 turbulent realizations. The solid lines indicate the mean NRMSE obtained from ten independent random seed selections, while the shaded areas denote the corresponding minimum and maximum values.



The overall trend is remarkably consistent across all considered load channels. Surrogate performance converges rapidly as the number of turbulent seeds increases, with most of the achievable improvement obtained within approximately 8 to 12
990 turbulent realizations. Beyond this range, additional seeds provide only marginal reductions in NRMSE.

For the shaft, yaw-bearing, and tower-base DELs, the training and validation NRMSE values continue to decrease gradually up to approximately 14 to 16 seeds. However, the corresponding testing NRMSE values reach a plateau much earlier, typically after about 10 seeds. This observation is particularly interesting because the testing dataset contains independent operating conditions and turbulent realizations that were not used during surrogate training. The results therefore suggest that increasing
995 the number of turbulent seeds beyond approximately 10 does improve the agreement with the specific realizations used for model development, but provides little additional benefit when predicting unseen cases. From a practical perspective, this indicates that averaging over approximately 10 seeds is already sufficient to capture most of the load-relevant stochastic variability required for surrogate generalization.

The blade-root DEL exhibits an even faster convergence, with the testing NRMSE reaching its asymptotic value after approximately 6 to 8 turbulent realizations. This behavior is consistent with the conclusions of the previous sections, which showed that blade-root fatigue loads are governed primarily by rotor-scale flow asymmetries and mean operating conditions rather than by turbulence intensity alone. Indeed, the projected blade-root DEL is largely dominated by the blade in-plane bending moment, whose fatigue accumulation is strongly linked to mean rotor speed, mean wind speed, and wake-induced in-flow heterogeneity. Consequently, averaging over a relatively small number of turbulent realizations is sufficient to characterize
1000 the relevant loading trends.

It is important to emphasize that the present analysis should be interpreted qualitatively rather than as a strict convergence proof. Ideally, a much larger number of turbulent realizations would be required to establish a true statistical reference solution. Here, the analysis uses averages over 24 turbulent seeds as the best available approximation of such a reference. As a consequence, the uncertainty bands naturally collapse at 24 seeds because only one possible realization exists when all available
1010 seeds are included.

The width and asymmetry of the uncertainty envelopes provide additional insight. For a given number of turbulent seeds, the upper envelope is generally further from the mean than the lower envelope. In other words, a random selection of seeds is more likely to degrade surrogate performance than to improve it. Obtaining an unusually favorable seed selection offers only limited gains, whereas an unfavorable selection can noticeably increase the NRMSE. Encouragingly, the upper envelopes also
1015 converge rapidly with increasing numbers of seeds. Even in the worst-performing realizations, using approximately 10 to 12 turbulent seeds results in substantially lower prediction errors than training on only one or a few realizations. This observation further supports the conclusion that a relatively modest number of turbulent seeds is sufficient to construct accurate and robust load surrogate models.

Based on the convergence trends observed in Fig. 13, a value of approximately 12 turbulent realizations appears to provide a robust compromise between predictive accuracy and computational cost for the development of wind-farm-level load surrogate models. Although the exact convergence behavior will depend on the considered turbine, load channel, inflow conditions, and surrogate formulation, little additional benefit is observed beyond this range in the present study. Interestingly,



this recommendation is approximately twice the number of turbulent realizations commonly considered sufficient for turbine-level fatigue-load assessment in design standards (International Electrotechnical Commission, 2019), reflecting the additional variability introduced by wake interactions and partial-wake conditions in a wind-farm environment.

4.6 Influence of the training database size

The final question addressed in this surrogate-design study concerns the amount of training data required to construct a robust and generalizable load surrogate model.

By denoting f_{train} as the fraction of the full training and validation dataset ($N = 30783$) allocated to training, the convergence of surrogate performance with increasing training database size can be investigated. Figure 14 reports the resulting NRMSE values as a function of f_{train} . For each value of f_{train} , the remaining fraction ($1 - f_{\text{train}}$) is used for validation, while the independent testing dataset remains unchanged throughout the analysis.

All results presented so far used the conventional machine-learning choice of $f_{\text{train}} = 80\%$, corresponding to 24 626 training samples. An important remaining question is whether this database size is sufficient to achieve convergence of the surrogate performance, and conversely whether comparable predictive accuracy could be obtained using substantially fewer training samples.

More fundamentally, the required database size is closely linked to the representativeness of the selected cases. The objective of the training dataset is not necessarily to maximize the number of samples, but rather to provide sufficient coverage of the relevant inflow, control, and wake-interaction patterns that govern turbine loads. In principle, once a given pattern is adequately represented, adding additional samples with highly similar characteristics may provide only limited benefit. Investigating dedicated sampling and dataset-design strategies therefore constitutes a complementary research direction. In the present work, however, the focus is placed on evaluating surrogate performance for datasets generated from broad and physically representative ranges of operating conditions and wind-farm configurations, while more advanced sampling methodologies are discussed in the related study of Vad et al. (2026a).

To statistically quantify the influence of case selection, the complete dataset was randomly shuffled ten times prior to generating training and validation subsets of different sizes. The solid lines in Fig. 14 correspond to the mean NRMSE obtained across these ten realizations, while the shaded envelopes indicate the associated minimum and maximum values.

The results show a clear and well-behaved convergence trend consistently across all considered load channels. As additional training cases are introduced, the prediction errors gradually decrease before approaching a plateau. This behavior indicates that the surrogate progressively learns the dominant relationships between the reduced-order inflow descriptors, the control variables, and the corresponding fatigue loads. Beyond a certain database size, additional training cases provide only limited improvements, suggesting that the surrogate has already captured most of the relevant patterns contained within the considered operating space.

In the present study and for the considered datasets, the surrogate prediction errors have effectively converged for all load channels once approximately 20% of the available training and validation dataset is used for training, corresponding to roughly 6 000 training samples. It should be emphasized, however, that the exact number of cases required for convergence is not a

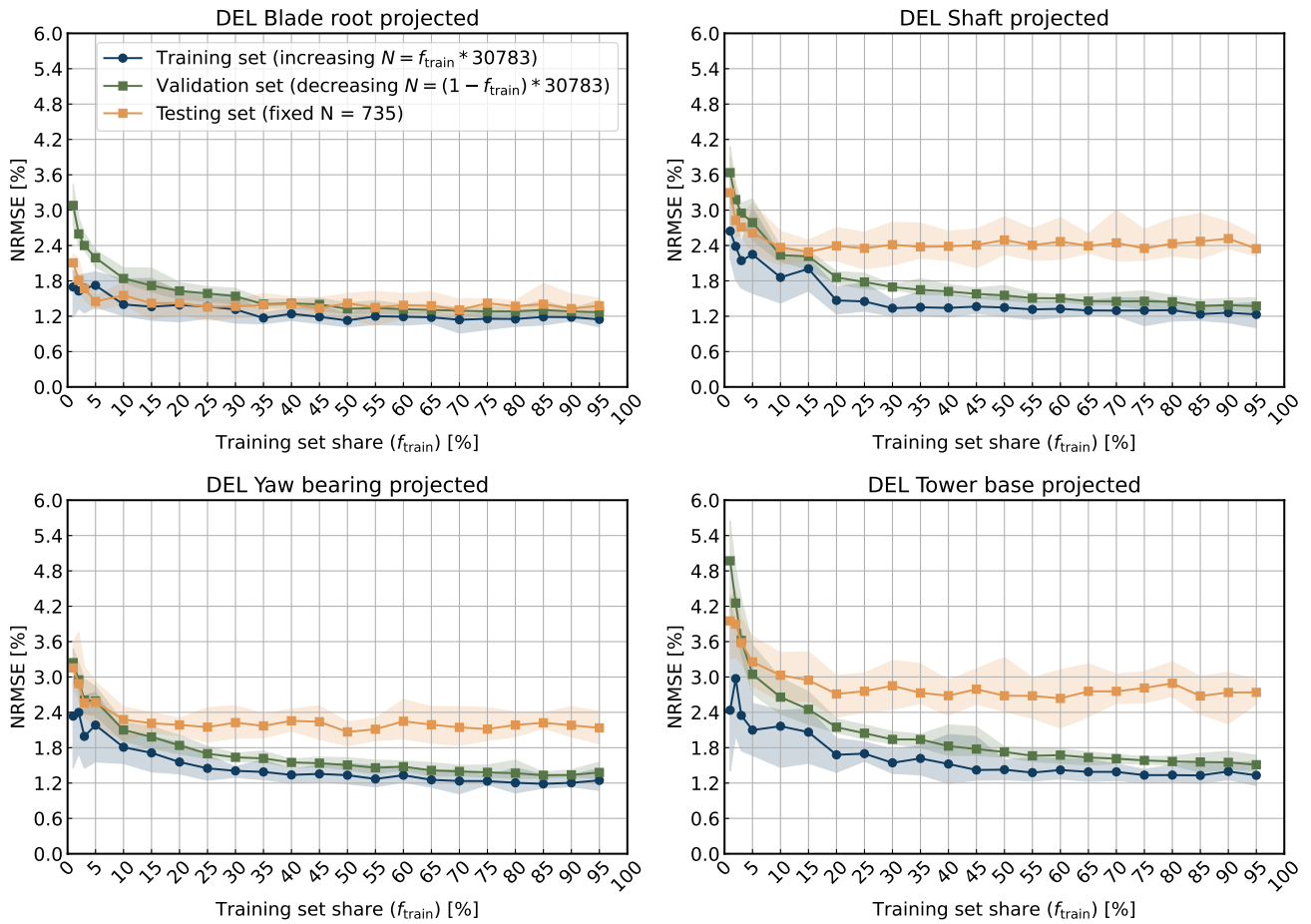


Figure 14. Performance of the surrogates as a function of the training fraction f_{train} , using the SAWS 4A + RATI inflow representation. For each value of f_{train} , the full training and validation dataset is randomly shuffled ten times before generating the corresponding training and validation subsets. The solid lines indicate the mean NRMSE across the ten realizations, while the shaded areas denote the associated minimum and maximum values.



universal quantity. The specific value depends strongly on the extent of the input space covered by the database, including the considered ranges of wind speed, turbulence intensity, wake conditions, control operating points, and their combined interactions. Consequently, the convergence level observed in Fig. 14 should not be interpreted as a general requirement for load surrogate modeling. Rather, the important conclusion is that a finite and comparatively moderate number of training cases is sufficient for the present surrogate formulation to generalize accurately to previously unseen operating conditions.

A second noteworthy observation concerns the distinct convergence behavior of the training, validation, and testing datasets. For all considered channels, the independent testing set reaches its performance plateau earlier, typically between 15 % and 20 % of the available training database, whereas the training and validation errors continue to improve more gradually until approximately 30 % to 40 %. This result suggests that only a relatively limited number of representative cases is required when the primary objective is generalization to new wind-farm layouts and operating conditions. While increasing the database size beyond this point may continue to improve performance on cases drawn from the same underlying distribution, it provides comparatively little additional benefit when predicting entirely unseen wind-farm configurations.

The spread associated with the random dataset shuffling also provides useful insight. In contrast to the convergence study with respect to the number of turbulent seeds, where asymmetric envelopes revealed the existence of particularly unfavorable seed selections, the minimum and maximum envelopes in Fig. 14 remain approximately symmetric around the mean. Furthermore, the total variation introduced by different random selections of training cases remains relatively small, typically on the order of only 0.7 % NRMSE. This narrow spread indicates that no particular subset of cases is indispensable for achieving good predictive performance. Instead, the dominant load-generating patterns appear to be represented throughout the dataset, making the surrogate relatively insensitive to the precise realization obtained through random shuffling. Naturally, this conclusion applies only within the context of the present database construction strategy, in which the full range of operating conditions is sampled. Restricting the training data to a narrow subset of operating regions would be expected to produce significantly poorer extrapolation capabilities.

An equally important observation emerges when comparing the present results with the surrogate-design studies presented earlier on in this article. Increasing the size of the training database certainly improves surrogate performance, but its impact remains comparatively modest relative to the choice of surrogate inputs. Figure 7 showed that replacing the simplest RAWs inflow representation by the more informative SAWS 4A + RATI description reduced the testing-set NRMSE from approximately 20 % to 2 %. Figure 11 further demonstrated that an appropriate sectoral representation of the inflow can reduce the testing-set NRMSE from approximately 10 % to 2 %. In comparison, the present convergence study indicates that increasing the training database size from only 1 % of the available dataset (roughly 300 training samples) to the default 80 % configuration (approximately 25 000 training samples) reduces the testing-set NRMSE from about 4 % to about 2 %. While the latter improvement is certainly beneficial, it remains substantially smaller than the gains achieved through physics-informed choices of surrogate inputs.

This comparison highlights one of the central conclusions of the present work. The predictive capability of a load surrogate depends far more strongly on providing the model with informative and physically meaningful descriptors than on brute-force increases in dataset size. In other words, selecting an appropriate reduced-order representation of the underlying physics is



considerably more valuable than simply providing a larger number of examples associated with an inadequate representation. Once such informative descriptors have been identified, only a moderate number of training cases is required to achieve robust predictive performance and good generalization to unseen wind-farm operating conditions.

1095 5 Implementation of the load surrogate model in different environments

Having established the surrogate formulation and its key design choices in Sect. 4, the next step is to demonstrate its practical deployment in realistic applications. In particular, the surrogate was formulated using reduced-order inflow and control descriptors, namely the SAWS 4A + RATI inflow representation contained in $\mathbf{i}$ together with the yaw-misalignment and power-demand descriptors contained in $\mathbf{c}$. While these quantities are directly available during the surrogate-design phase, practical
1100 deployment requires suitable approaches to obtain them from the information available in operational wind-farm environments.

An important outcome of the design study is that the selected control descriptors remain extremely simple. As demonstrated in Sect. 4.3, the effects of turbine control actions can be represented accurately using only the yaw-misalignment and power-demand setpoints. Unlike more sophisticated control parameters based on turbine operating states or time-series statistics, these quantities are directly prescribed by the wind-farm controller and are therefore readily available in practice. Consequently, the
1105 deployment challenge does not primarily lie in the estimation of the control vector $\mathbf{c}$.

As illustrated by the transition from the left to the right side of Fig. 1, the remaining challenge is therefore the estimation of the local inflow descriptor $\mathbf{i}$ required as input to the surrogate. The present section investigates two complementary implementation pathways for this purpose.

As discussed in Sect. 2.3.3, the practical performance of a load surrogate depends not only on the choice of inflow representation, but also on the consistency between the inflow quantities used during training and those available during deployment.
1110 Different sources of inflow information may represent the flow field with different levels of fidelity and therefore yield different values of the reduced-order descriptors. For this reason, the objective is not necessarily to recover the exact inflow quantities used during surrogate design, but rather to assess whether a consistent training-to-deployment strategy can preserve the predictive accuracy of the surrogate.

1115 This section is organized as follows. First, Sect. 5.1 considers the use of wind-observer technology to estimate the required inflow quantities from turbine measurements. Second, Sect. 5.2 investigates the use of steady-state engineering wake models to provide the corresponding inflow descriptors from flow-model predictions. Together, these examples cover two representative deployment scenarios, namely measurement-based estimation and model-based inflow prediction.

5.1 Implementation using wind-observer estimates

1120 The wind-observer technology (Bottasso et al., 2018; Schreiber et al., 2020; Bertelè et al., 2024; Onnen et al., 2026) is used to estimate the local inflow at the rotor disk based on blade-load measurements. The underlying principle is to use the rotor itself as a distributed wind sensor. Each rotating blade acts as a scanning sensor, sampling different regions of the inflow as it



traverses the rotor disk. By combining the measured blade-root bending moments with suitable observer algorithms, estimates of the local wind conditions can be reconstructed across the rotor-swept area.

1125 In the present work, time series of local inflow estimates are obtained using this observer and subsequently transformed into the reduced-order inflow quantities required by the surrogate model. Although the detailed implementation of the observer itself is outside the scope of the present work, the key point is that it provides a practical pathway for estimating the surrogate inflow vector $\hat{i}$ directly from standard turbine measurements, without requiring dedicated flow-measurement devices. While wind-observer technology is commonly applied to measurements from operating turbines, the present study adopts the
 1130 same methodology in a simulation environment in order to assess the integration of observer-based inflow estimates with the proposed load surrogate.

The observer framework also provides a natural motivation for the distinction between the averaged and effective quantities defined in Sect. 2.3. The effective formulation is particularly well aligned with the observer approach, since the rotating blades inherently perform a spatial averaging of the inflow before temporal statistics are computed from the resulting signals.
 1135 Conversely, the averaged formulation can be obtained by first reconstructing local inflow time series and subsequently applying the corresponding temporal and spatial averaging operations. For this reason, the observer-based implementation is assessed for both the SAWS/RATI and SEWS/RETI formulations in the results that follow. This distinction is especially relevant in the present subsection, whereas the steady-state engineering wake models considered in Sect. 5.2 naturally provide only averaged quantities due to the absence of time-varying inflow information.

1140 The resulting estimates of RAWs, SAWS 4A, RATI, and SATI 4A obtained from the wind observers (denoted OBS) are directly compared with the corresponding quantities extracted from virtual sensors (VS) in the simulation environment. These VS-based inflow quantities were used throughout Sect. 4 for the design and evaluation of the surrogate formulation. The comparison between OBS and VS estimates for the averaged inflow quantities is presented in Fig. D1 in Appendix D, while the corresponding comparison for the effective quantities, namely REWS, SEWS 4A, RETI, and SETI 4A, is shown in Fig. D2.

1145 As shown in Appendix D, the local inflow quantities derived from VS and OBS exhibit strong agreement and correlation. Although the numerical values do not match perfectly, which is expected given the fundamentally different physical processes used to estimate them, all comparisons yield R^2 values above 0.89. This provides strong evidence supporting the key hypothesis introduced in Sect. 2.3.3. In particular, the results suggest that the mapping $\mathcal{G}$ relating OBS- and VS-based inflow quantities remains sufficiently close to a one-to-one correspondence for the two representations to be considered practically equivalent
 1150 from the perspective of surrogate deployment.

Another important observation from Appendix D is that wind-speed quantities generally exhibit stronger agreement between OBS and VS than turbulence-intensity quantities. This is consistent with the greater difficulty of estimating local turbulence intensity, which is inherently more sensitive to the underlying estimation methodology. For the sector-resolved turbulence descriptors, systematic differences can be observed between the two turbulence formulations: the averaged quantities SATI 4A
 1155 tend to be underestimated by OBS relative to VS, whereas the effective quantities SETI 4A are generally overestimated. The comparison therefore reveals a formulation-dependent bias rather than a systematic agreement across all turbulence descriptors. In contrast, the corresponding rotor-averaged quantities, RATI and RETI, exhibit substantially stronger agreement and corre-



lation, with all associated R^2 values exceeding 0.97. This suggests that turbulence intensity can be estimated more robustly at the rotor scale than at the sector scale, where local discrepancies between estimation approaches become more pronounced.

1160 This observation is particularly encouraging in light of the conclusions from Sect. 4, where RATI and RETI were shown to provide sufficient turbulence-related information for accurate DEL prediction, while inflow heterogeneity was primarily captured through SAWS and SEWS. The fact that the rotor-averaged turbulence descriptors are also the most robust quantities with respect to the choice of inflow estimation method therefore provides additional support for this modelling choice and further strengthens its practical relevance for deployment.

1165 According to the considerations developed in Sect. 2.3.3, two deployment strategies are evaluated. In the first approach, the surrogates are trained using VS-derived inflow quantities, as in Sect. 4, and subsequently deployed using OBS-derived estimates. In the second approach, both training and deployment rely exclusively on OBS-derived inflow quantities.

Together with the comparison between averaged and effective inflow representations, the evaluation of these two deployment strategies constitutes the main objective of the performance analysis presented in Fig. 15. As in the surrogate design
 1170 study, performance is primarily assessed using the NRMSE metric, with the objective of identifying the most suitable inflow representation and deployment strategy.

As argued in Sect. 2.3.3, the results in Fig. 15 clearly show that surrogates trained using VS-derived inflow quantities cannot be directly deployed using OBS-derived inflow quantities without a substantial loss of accuracy. Depending on the load channel, the NRMSE increases from approximately 1.5–3% to 6–12% when the inflow source used during deployment differs from that
 1175 used during training. This degradation is consistently observed across all four load channels and for both the averaged and effective inflow formulations.

On the other hand, as expected, when the surrogates are trained directly using OBS-derived inflow quantities and subsequently deployed with OBS-derived inflow quantities, the prediction accuracy returns to similarly low NRMSE values of approximately 1.5–3%. Across all load channels, and for both averaged and effective formulations, the performance obtained
 1180 with OBS-based surrogates is therefore comparable to the one obtained with VS-based surrogates, provided that the same inflow source is used consistently during both training and deployment. This result highlights two important findings. First, the surrogate formulation introduced in Sect. 2 is sufficiently generic to achieve good predictive performance regardless of whether the local inflow descriptors originate from virtual sensors or wind observers. Second, surrogate deployment should ideally rely on the same type of inflow estimate used during training, or on an inflow representation that remains consistent with the learned
 1185 surrogate mapping. While the present results demonstrate the importance of consistency between training and deployment, the transferability of a surrogate also depends on whether the deployed inflow descriptors adequately represent the load-relevant characteristics encountered in application. A broader discussion of these aspects, including the interaction between surrogate models and different inflow-modeling frameworks, is provided by Vad et al. (2026a).

Although this requirement limits the possibility of training a surrogate in one environment and applying it unchanged in another, it is physically intuitive and provides a clear practical guideline: train the surrogate using the same inflow representation
 1190 that will be available during deployment. This consistency is important not only from a machine-learning perspective, where training and inference inputs should follow the same distribution, but also from a wind-physics perspective, as the spatio-

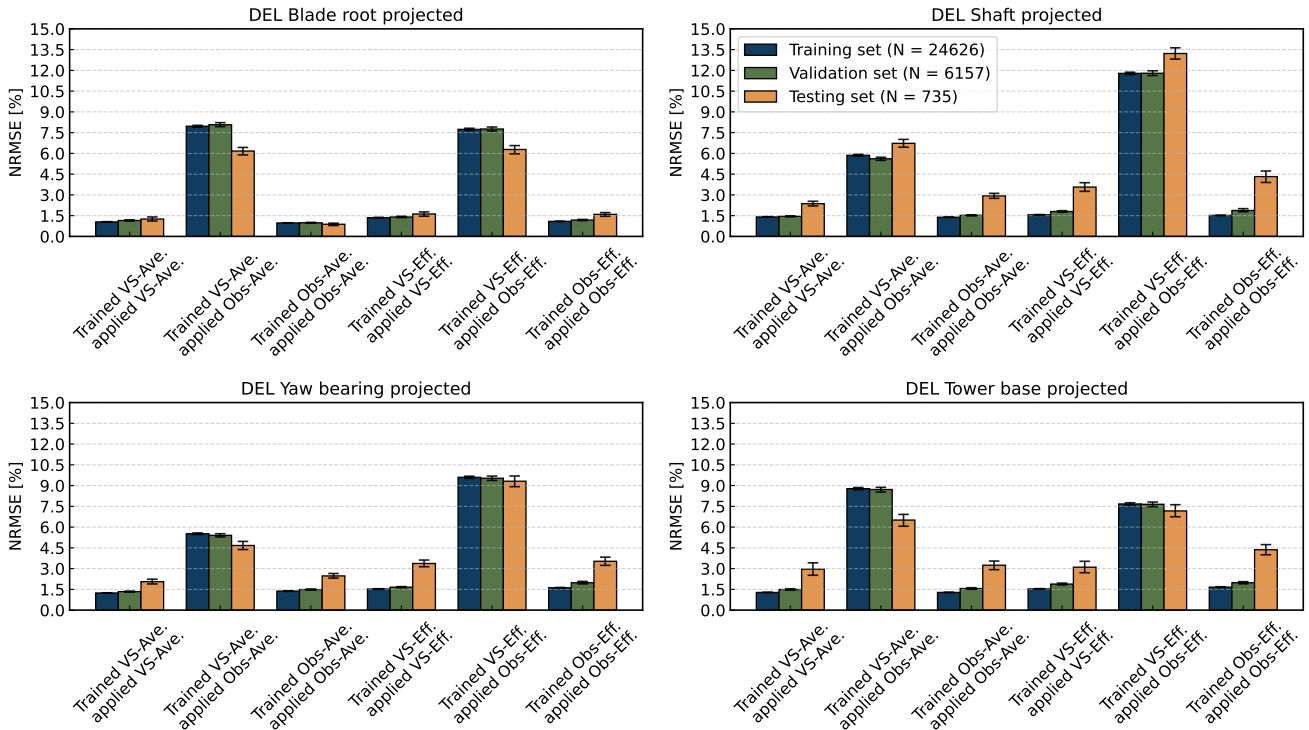


Figure 15. Performance of load surrogates using wind observers (denoted Obs) compared to virtual sensors in the dynamic simulations (denoted VS). Both sector-averaged inflow quantities (denoted Ave.) and sector-effective inflow quantities (denoted Eff.) are included for comparison. In each case, the option SAWS 4A + RATI for Ave. (respectively, SEWS 4A + RETI for Eff.) is used here to characterize the inflow. The bars show the deterministic NRMSE computed on the full training, validation, and test sets, respectively. The whiskers indicate the corresponding 95% confidence intervals estimated using non-parametric bootstrapping (1000 resamples with replacement).

temporal characteristics of the local inflow must be represented in a consistent manner. These observations therefore provide strong support for the arguments formulated in Sect. 2.3.3.

1195 Finally, similarly to the results presented in Sect. 4.2, the averaged and effective inflow formulations yield nearly identical predictive performance when training and deployment rely on the same inflow source. This again suggests that both formulations contain essentially the same level of information for the considered surrogate task. In contrast, when surrogates are trained using VS-derived inflow quantities and deployed using OBS-derived inflow quantities, the performance degradation is generally more pronounced for the effective formulation than for the averaged formulation, particularly for the shaft and yaw
 1200 bearing load channels.

To complement the aggregate NRMSE metrics presented above, Fig. 16 provides a point-wise view of the surrogate predictions through scatter plots of predicted versus reference DEL values. Focusing on the averaged inflow formulation, three



configurations are compared: (i) training and deployment using VS-derived inflow quantities, (ii) training using VS-derived inflow quantities and deployment using OBS-derived inflow quantities, and (iii) training and deployment using OBS-derived inflow quantities. For clarity, all available training, validation, and testing cases are combined in Fig. 16.

The trends identified from the NRMSE analysis become particularly evident in these scatter plots. When the inflow source used during deployment differs from that used during training, the predictions exhibit a substantially larger spread around the one-to-one line, with both under- and over-predictions occurring across the DEL range. In addition, the surrogate appears less capable of reproducing the full range of observed DEL values, leading to a noticeable compression of the predictions toward intermediate values. In contrast, when training and deployment rely on the same inflow source, the predicted values remain tightly clustered around the one-to-one line. In particular, the OBS-based surrogates recover a level of agreement comparable to the one achieved with VS-based surrogates, providing a direct visual confirmation that the loss of accuracy originates primarily from the inconsistency between training and deployment inflow representations rather than from any intrinsic limitation of the observer-derived inflow quantities themselves.

Beyond the levels of prediction accuracy reported above, it is important to clarify the intended role of the combined wind-observer and load-surrogate framework. The objective is clearly not to reconstruct the same blade-root loads used by the observer, but rather to use a limited set of measurements to estimate structural loading quantities that are otherwise unavailable. In the present study, blade-root measurements are first used to estimate local inflow conditions through the wind observer, and these inflow estimates are subsequently used by the surrogate model to predict DELs at multiple turbine locations, including the shaft, yaw bearing and tower base. The proposed methodology therefore enables virtual sensing of fatigue-related loads across turbine components that may not be directly instrumented. A second potential application is the integration of wind observers and load surrogates within a closed-loop control framework. Wind observers provide estimates of the instantaneous local inflow and wake conditions, while the surrogate model provides a computationally inexpensive estimate of the associated fatigue loading over longer time periods. Since control actions can generally be updated on timescales shorter than those associated with changes in ambient inflow and wake conditions, the combined framework may provide a predictive capability whereby emerging loading conditions are identified and mitigated before their full fatigue consequences develop.

5.2 Implementation using engineering wake models

The second implementation investigated in this work consists of coupling the load surrogate model with steady-state engineering wake models. In contrast to the observer-based implementation presented in Sect. 5.1, the reduced-order inflow quantities are here obtained directly from wake model predictions rather than reconstructed from turbine measurements.

Two wake model configurations are considered within the FLORIS framework (NREL, 2025). The first, denoted GCH, combines the Gaussian velocity-deficit model of Bastankhah and Porté-Agel (2016); Niayifar and Porté-Agel (2016) with the turbulence model of Crespo and Hernández (1996). The second, denoted GIQ, combines the same Gaussian velocity-deficit model with the wake-added turbulence formulation of Ishihara and Qian (2018). While GCH is available in the standard FLORIS implementation, GIQ was implemented specifically for the present work following the formulation proposed by Ishihara and Qian (2018). For both approaches, the default model parameters were retained, including the standard FLORIS

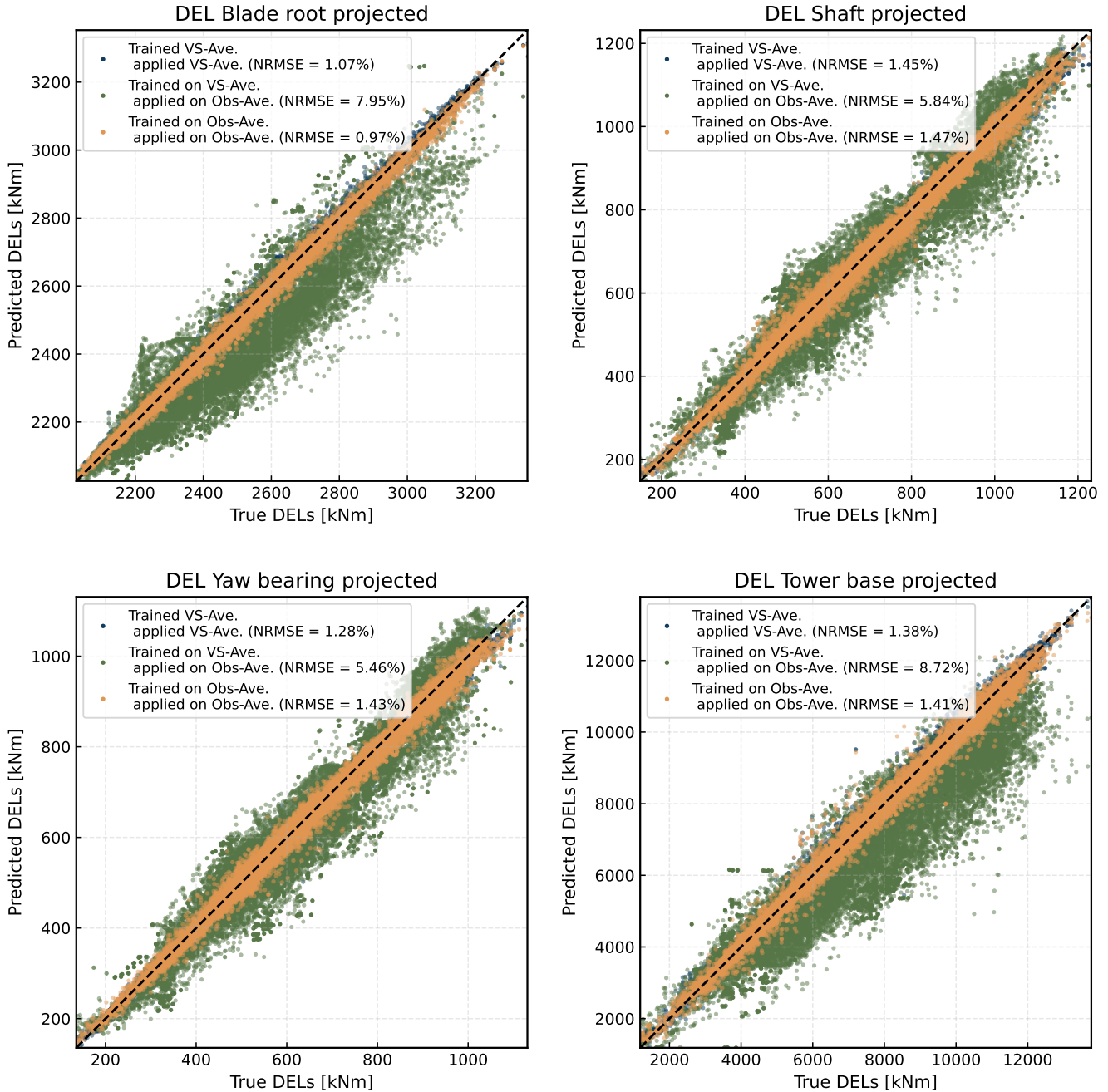


Figure 16. Scatter plots from load surrogates using wind observers (denoted Obs) compared to virtual sensors in the dynamic simulations (denoted VS). Here, all three (training, validation + testing) sets are merged together (31518 cases). In each case, the option SAWS 4A + RATI is used here to characterize the inflow.



parameters for GCH (NREL, 2025) and the parameter values reported by Ishihara and Qian (2018) for GIQ. Additional details on the adopted parameterizations are provided in Appendix E. This choice avoids introducing a dataset-specific calibration that could bias the comparison and ensures that the robustness and generalization capability of the proposed load-surrogate framework are evaluated under representative out-of-the-box steady-state wake-model conditions.

Among the two configurations, GIQ is expected to provide a more suitable foundation for the load surrogate model because it represents wake-added turbulence through both lateral and vertical distribution functions, thereby capturing its spatial variability more realistically than GCH, which describes wake-added turbulence only as a streamwise function. Although previous results in this article suggest that the surrogate can effectively map fatigue loads using the compact rotor-averaged turbulence parameter RATI, a more detailed representation of wake-added turbulence is still expected to improve the prediction of this quantity, particularly under conditions of partial wake overlap. Nevertheless, both models are considered here to assess whether such additional complexity is required in practice.

To quantify the consistency between dynamic and steady-state wake descriptions, the reduced-order inflow quantities obtained from DWM, GCH, and GIQ are compared for the combined training and validation datasets. One-to-one scatter plots are provided in Appendix E, with GCH versus DWM shown in Fig. E1 and GIQ versus DWM shown in Fig. E2.

Similarly to the observer-based implementation, the comparisons in Fig. E1 and Fig. E2 show that the reduced-order inflow quantities do not coincide exactly, particularly for TI. Nevertheless, a strong correlation is observed, with R^2 values exceeding 0.95 for wind speed parameters and 0.85 for TI parameters. This confirms the reasoning of Sect. 2.3.3 that the mapping $\mathcal{G}$ relating dynamic and steady-state wake-model inflow quantities cannot be regarded as an identity mapping, yet remains sufficiently close to a one-to-one correspondence for both representations to be considered practically equivalent from the perspective of surrogate deployment.

A comparison of Fig. E1 and Fig. E2 indicates that both GCH and GIQ exhibit a similar overall level of agreement with the DWM-derived inflow quantities. In particular, the distributions and corresponding R^2 values are comparable for both wind-speed and TI parameters. A closer inspection nevertheless reveals a systematic difference in the TI predictions: the default GCH configuration tends to underestimate the local TI levels, whereas the default GIQ configuration tends to overestimate them. These opposite biases partially compensate when considering aggregate correlation metrics, resulting in broadly similar levels of agreement with the DWM reference. From this perspective, neither wake-model configuration appears to provide a clear advantage over the other in terms of reproducing the reduced-order inflow quantities required by the surrogate framework.

Notably, both GCH and GIQ exhibit a noticeably stronger agreement with DWM for the rotor-averaged turbulence parameter RATI than for the four sector-resolved SATI quantities. This observation further supports the choice of SAWS 4A + RATI as the preferred inflow representation for the surrogate framework. In this formulation, the spatial variability of the inflow is primarily captured by the four sector-averaged wind-speed parameters, while a single rotor-averaged turbulence parameter appears sufficient to characterize the overall turbulence level, whether originating from ambient conditions or wake-added turbulence. As shown in Sect. 4.1, this reduced-order representation already provides sufficient surrogate accuracy. The present results suggest that this choice is beneficial not only from a surrogate-modeling perspective, but also from an implementation per-



spective, as the rotor-averaged turbulence descriptor appears inherently more robust to differences between inflow-estimation methodologies than its sector-resolved counterparts.

Following the same methodology as for the observers, three deployment strategies are evaluated: (i) training and application using DWM-based inflow quantities, (ii) training using DWM-based inflow quantities and deployment using GCH- or GIQ-
 1275 based inflow quantities, and (iii) training and deployment performed consistently using the same steady-state wake model.

The resulting surrogate performance, quantified through the NRMSE across all datasets, is summarized in Fig. 17, which also enables a direct comparison between the GCH and GIQ wake-model implementations.

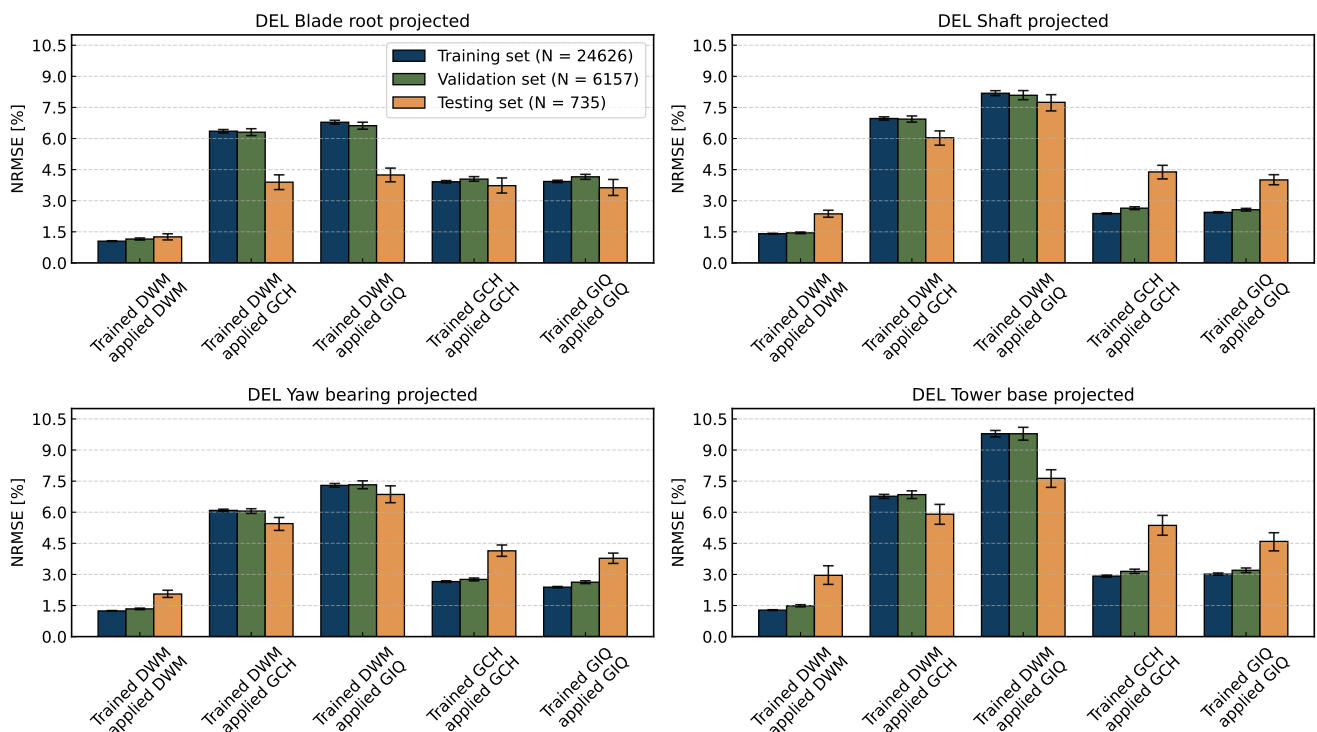


Figure 17. Performance of load surrogates coupled with steady-state wake models. In all cases, the SAWS 4A + RATI inflow representation is used.

Remarkably, the surrogate performance obtained with GCH and GIQ is nearly identical, both when the surrogates are deployed inconsistently and when they are trained and applied using the same wake-model representation. It is important
 1280 to note that the primary difference between these two wake models lies in their treatment of the spatial heterogeneity of the wake-added turbulence. This observation is in line with the results presented earlier in this article, which showed that a rotor-averaged turbulence representation is generally sufficient for accurate load prediction. Although the spatial distribution of wake-added turbulence remains important for determining the resulting RATI value, differences in the detailed turbulence



patterns predicted by GCH and GIQ appear to have only a limited influence on the final surrogate accuracy. This further suggests that the surrogate framework does not rely on a uniquely "correct" representation of the local inflow. Rather, it learns the relationship between a given inflow description and the associated structural loads. Nevertheless, successful deployment still requires the chosen inflow representation to retain the load-relevant characteristics governing a structural response. The good performance obtained with both GCH and GIQ indicates that these engineering wake models capture the dominant inflow trends required for load prediction, even though they rely on different turbulence formulations.

The same fundamental conclusion as for the observer-based implementation, and as predicted by the theoretical considerations of Sect. 2.3.3, therefore emerges. Surrogates trained using DWM-derived inflow quantities cannot be directly deployed using inflow quantities derived from steady-state wake models without a substantial loss of accuracy. Depending on the load channel, the NRMSE increases from approximately 1.5–3% to 6–9% when the inflow source used during deployment differs from that used during training. This degradation is consistently observed across all four load channels and the two steady-state wake models tested.

Conversely, when the surrogates are trained directly using inflow quantities derived from the corresponding steady-state wake model and deployed using the same representation, the prediction accuracy improves markedly compared to the inconsistent deployment strategy. Depending on the load channel, NRMSE values are reduced from approximately 6–9% to about 2–4.5%. Although this performance remains slightly worse than the one achieved when both training and deployment rely on DWM-derived inflow quantities (approximately 1.5–3%), it nevertheless indicates that a large fraction of the apparent loss of accuracy originates from the inconsistency between training and deployment rather than from the steady-state wake models themselves.

This distinction is particularly informative when compared to the observer-based implementation discussed previously. In that case, training and deployment using observer-derived inflow quantities recovered virtually the same performance level as the original DWM-based surrogates. In contrast, a moderate performance gap remains here even when the steady-state wake-model implementation is fully consistent. This result is physically plausible, as steady-state wake models do not resolve the full spatio-temporal variability of the local inflow experienced by the turbines. Consequently, part of the load-driving information available in the dynamic simulations is inevitably lost when constructing the reduced-order inflow descriptors from steady-state wake predictions. The remaining performance gap therefore suggests that further improvements may be achieved through future engineering wake models that better capture the unsteady inflow features governing structural loading.

Nevertheless, the remaining prediction errors remain remarkably low considering the simplicity and computational efficiency of the underlying wake models. From an application perspective, this is an encouraging result: while steady-state wake models cannot reproduce all sources of load variability captured by dynamic simulations, they still provide sufficiently informative inflow descriptors to enable accurate load-surrogate predictions. The results therefore suggest that a practical compromise can be achieved, whereby a modest reduction in surrogate accuracy is accepted in exchange for the computational tractability required in applications such as wind farm control, layout design and co-design, and others.

To provide a more detailed view of the error patterns underlying the NRMSE results, Fig. 18 presents predicted versus reference DEL values for the three deployment strategies considered above. For clarity, only the GIQ implementation is shown, as very similar trends were obtained with GCH. All available training, validation, and testing cases are combined in the figure.

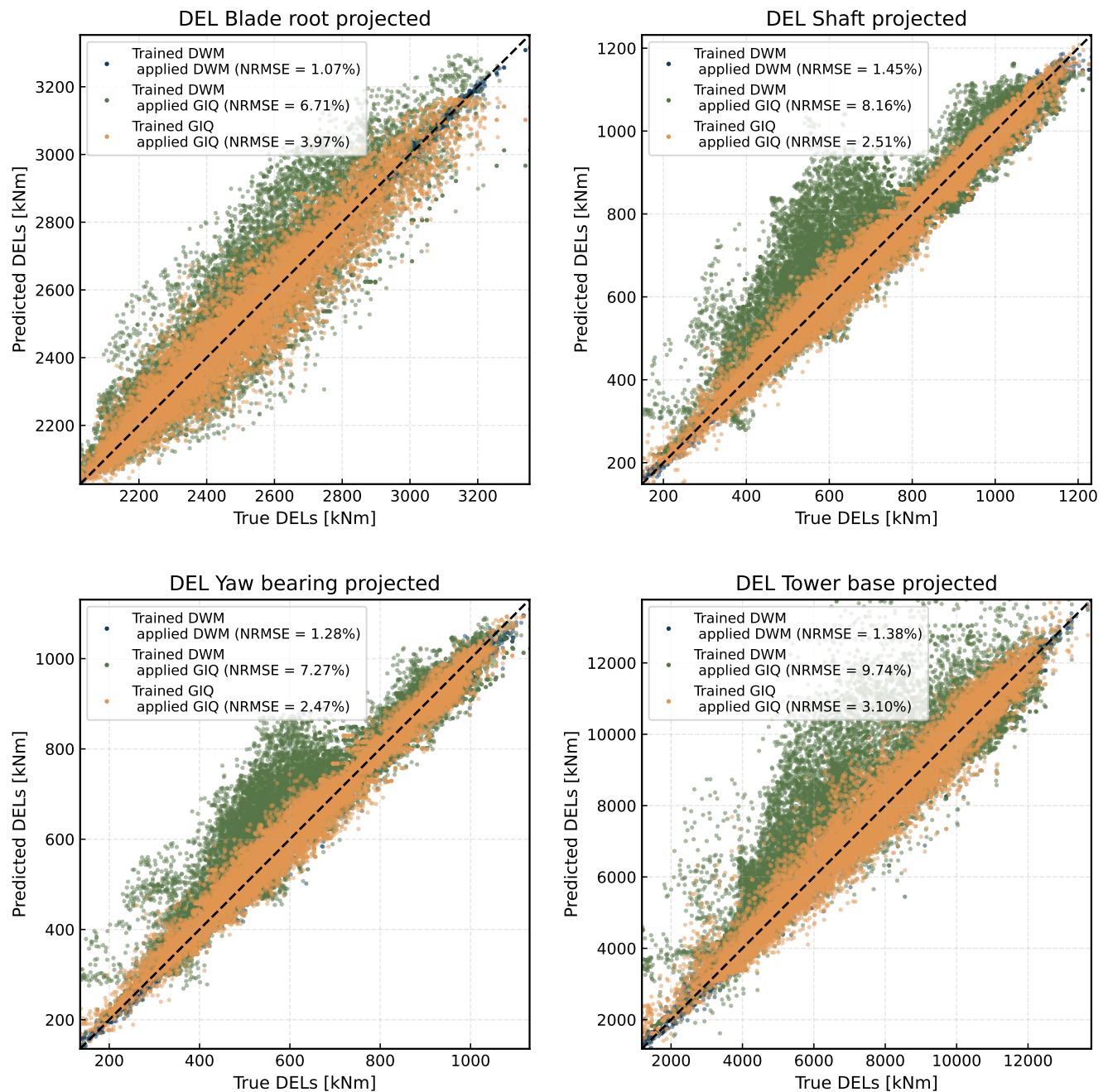


Figure 18. Predicted versus reference load quantities for surrogates coupled with the GIQ steady-state wake model. Training, validation, and testing datasets are combined (31 518 cases). The SAWS 4A + RATI inflow representation is used in all cases.



The conclusions drawn from the aggregate performance metrics are clearly reflected in the scatter plots. When a surrogate trained using DWM-derived inflow quantities is deployed using GIQ-derived inflow quantities, a pronounced degradation in prediction quality is observed across all load channels. Unlike the observer-based implementation, where the mismatch primarily manifested as an increased spread of the predictions, the present case also exhibits a clear systematic overprediction of DEL values. This behavior is consistent with the tendency of GIQ to overestimate the local RATI values relative to DWM, as previously observed in Fig. E2.

When training and deployment are performed consistently using GIQ-derived inflow quantities, the agreement improves substantially and the systematic bias largely disappears. Although the resulting scatter remains slightly wider than for the reference DWM-trained and DWM-deployed surrogate, it is nevertheless much narrower than for the inconsistent deployment strategy. This observation is in line with the NRMSE analysis and further confirms that most of the performance degradation originates from the mismatch between the inflow representation used during training and deployment. The remaining difference relative to the DWM-based implementation can then be attributed to the lower-fidelity description of the inflow provided by the steady-state wake model itself.

An additional question naturally arises from these results. Since the analyses of Sect. 4 were primarily conducted using DWM-derived inflow quantities, one may wonder whether the same conclusions regarding the optimal inflow representation remain valid when the surrogates are instead trained and deployed directly using steady-state wake-model inflow quantities.

To address this question, the complete convergence analyses were repeated using inflow quantities derived from the steady-state wake models. The detailed results are reported in Appendix G to avoid distracting from the main discussion. Remarkably, the same trends are recovered. Figure G1 shows that the convergence with increasing spatio-temporal inflow complexity closely mirrors that previously obtained using DWM-derived inflow quantities in Fig. 7. Likewise, the convergence with increasing numbers of sectors (from 1 to 8) obtained using GIQ-derived inflow quantities, shown in Fig. G2, leads to the same conclusions as those drawn from the DWM-based analysis in Fig. 11: four sectors are sufficient to achieve converged surrogate accuracy, with little benefit from further increasing the spatial resolution. In both cases, the conclusions remain unchanged: SAWS 4A + RATI represents the simplest inflow description capable of reproducing the converged surrogate accuracy.

These additional results strengthen the generality of the proposed framework. Beyond the requirement of maintaining consistency between training and deployment, the main design recommendations established in Sect. 4 remain applicable across all investigated inflow sources and wake-model fidelities. In particular, SAWS 4A + RATI emerges as the most robust inflow representation identified in this work, regardless of whether the reduced-order quantities are derived from DWM simulations, wind-observer estimates, or steady-state wake models.

Taken together, the observer- and wake-model-based implementations provide strong empirical support for the theoretical considerations introduced in Sect. 2.3.3. The reduced-order inflow quantities should not be viewed as universal physical variables that can be exchanged freely between different models and estimation approaches. Instead, they serve as an intermediate representation whose exact definition is tied to the method used to generate them. The surrogate therefore learns the relationship between loads and a particular inflow representation, rather than loads and a unique "true" inflow state. Consistency between training and deployment is consequently of primary importance. When this consistency is maintained, high predictive



accuracy is achieved across all investigated implementations, with the remaining differences mainly reflecting the amount of
1355 load-relevant information retained by the underlying inflow representation itself.

6 Conclusions

This work addressed a central question in the development of wind-turbine load surrogate models: what level of physical complexity is actually required to obtain accurate and robust predictions of fatigue loads across a broad range of operating conditions, wake interactions, and deployment environments? Although recent studies have demonstrated the potential of
1360 machine-learning-based load surrogates for a variety of engineering applications, considerably less attention has been devoted to understanding how such surrogates should be designed and deployed in practice. Moreover, previous studies have primarily focused on surrogate accuracy within a given simulation or estimation framework, whereas the present work additionally investigated how the resulting surrogate formulation transfers across different practical implementations.

The objective of this paper was therefore not merely to demonstrate surrogate accuracy, but rather to identify the simplest
1365 physically meaningful formulation capable of delivering converged predictive performance and to assess its robustness across different sources of inflow information. In doing so, this work addressed both the design of load surrogate models and their practical deployment, thereby bridging the gap between methodological developments and real-world implementation.

The first contribution of this work concerns the design of the surrogate formulation itself. The results demonstrate that fatigue-load prediction does not require arbitrarily detailed inflow descriptions. Instead, a finite and relatively low-complexity
1370 representation is sufficient to capture the dominant effects of local inflow heterogeneity and control actions on fatigue loading. The principal findings can be summarized as follows:

- A clear convergence of surrogate performance is observed with increasing inflow-description complexity, demonstrating that only a limited amount of inflow information is required to accurately predict fatigue loads.
- Among all investigated inflow representations, SAWS 4A + RATI emerges as the simplest formulation capable of repro-
1375 ducing converged surrogate accuracy. The results indicate that four sector-averaged wind-speed parameters are sufficient to capture the dominant spatial heterogeneity of rotor inflow, while a single rotor-averaged turbulence parameter provides an adequate description of the load-relevant turbulence effects.
- The influence of turbine control actions can be represented directly through the steady-state control setpoints themselves. More detailed temporal descriptions of the control system do not provide measurable improvements in surrogate
1380 performance and therefore appear unnecessary for practical applications.
- Approximately 10–12 turbulent realizations are required and sufficient to achieve robust convergence with respect to turbulence variability for a location-agnostic load surrogate intended to generalize across different wake conditions and unseen turbulent realizations. Beyond this level, only marginal gains in predictive accuracy are observed.



- Surrogate performance converges with a finite and relatively modest number of training cases. More importantly, the results demonstrate that physically meaningful input selection plays a significantly greater role than simply increasing dataset size. In other words, understanding the physics of the problem is more valuable than brute-force expansion of the training database.

Taken together, these findings reveal an important property of fatigue-load surrogate modeling: accurate prediction does not require an exhaustive description of the inflow field. Instead, a carefully selected set of reduced-order descriptors can retain essentially all information required to estimate load responses with high accuracy. This result provides both a physical interpretation of the surrogate-learning process and a practical framework for constructing computationally efficient surrogate models.

The second contribution of this work concerns implementation. A theoretical framework was introduced in Sect. 2.3.3 to describe how surrogate performance is affected when reduced-order inflow quantities are derived from different models, estimation methods, or sources of information. The practical implications of this framework were then evaluated using multiple inflow-estimation approaches spanning different levels of fidelity. The principal findings are:

- Across all investigated implementations, a single conclusion emerges: load surrogates must be trained and deployed using a consistent inflow representation. The observed performance degradation originates primarily from mismatches between training and deployment inflow-description spaces rather than from the inflow-estimation methods themselves.
- The main design recommendations established in this work remain valid across all investigated implementations. In particular, SAWS 4A + RATI consistently emerges as the most robust reduced-order inflow representation, providing accurate predictions regardless of the source used to derive the inflow quantities.

These results provide strong empirical support for the theoretical considerations developed in Sect. 2.3.3. Reduced-order inflow quantities should not be viewed as universal physical variables that can be exchanged freely between different models or estimation frameworks. Instead, they represent a model-dependent description of the local inflow environment. From this perspective, the objective of the surrogate is not to reconstruct a unique “true” inflow state, but rather to learn the functional relationship between a chosen inflow representation and the structural loads. Consistency between training and deployment is therefore a fundamental requirement for successful implementation.

From a broader perspective, the results of this work demonstrate that fatigue-load surrogate models can achieve a desirable balance between physical interpretability, predictive accuracy, and practical applicability. The proposed formulation relies on a small number of physically meaningful inputs, exhibits converged performance with manageable training requirements, and remains robust across different inflow-estimation environments. These characteristics make the approach particularly attractive for applications requiring rapid load assessment, including wind-farm design, layout optimization, control-oriented studies, and operational decision support.

Based on the findings of this work, the following practical recommendations are proposed for the development of fatigue-load surrogate models:



- Use SAWS 4A + RATI as the default reduced-order inflow representation.
- Represent control actions using the corresponding steady-state control setpoints.
- Include at least 10–12 turbulent realizations during database generation.
- Prioritize physically meaningful inflow descriptors over simply increasing the number of training cases.
- Train and deploy the surrogate within the same inflow-description space.
- Select the inflow-estimation framework according to application requirements, while maintaining consistency between training and deployment.

Future work may extend the present framework in several directions.

First, the database considered in this work focused primarily on normal atmospheric and operating conditions, with ambient wind speeds close to rated operation. Future studies should investigate a broader range of inflow characteristics, including strong veer, low-level jets, gust events, extreme atmospheric conditions, and other non-standard wind profiles. While the recommended SAWS 4A + RATI formulation proved sufficient for all conditions considered here, more complex inflow structures may require a finer spatial discretization of the rotor to adequately capture their influence on fatigue loads. Nevertheless, extending the present methodology to such conditions would be relatively straightforward through the inclusion of additional representative training cases.

Second, the control-oriented demonstrations presented here were intentionally limited to the most widely deployed wind-farm control actions, namely wake steering through yaw misalignment and induction control through power-demand setpoints. Future work should investigate additional control strategies, including alternative derating approaches, noise-reduction operating modes, and more advanced load-aware and wake-mixing control concepts. Such applications may require extending the surrogate formulation with additional control descriptors, while preserving the simplicity principles identified in the present work. A first step in this direction has already been demonstrated by van der Hoek et al. (2026), where the sector-based surrogate formulation was coupled with a Helix wake-mixing strategy.

Third, the present study focused on establishing the optimal surrogate formulation and implementation strategy using a dynamic wake-model environment as the primary source of training and validation data. Future work should continue evaluating the robustness of the proposed framework across a wider hierarchy of modeling fidelities and experimental conditions. In particular, further validation against higher-fidelity numerical simulations, controlled wind-tunnel experiments, and full-scale field measurements would provide valuable insight into the practical limits and generality of the proposed methodology. A first step toward experimental validation of the surrogate framework and its practical application was presented by Braunbehrens et al. (2024). Nevertheless, substantially broader validation campaigns will be required to fully assess the robustness and transferability of the proposed methodology under realistic operating conditions. The framework itself is fundamentally independent of the specific source of inflow and load information, making such validations conceptually straightforward whenever the required inflow descriptors and load measurements are available.



Fourth, although the proposed surrogate formulation was demonstrated to be location-agnostic within a wind farm, it remains turbine- and controller-specific. Extending the methodology toward turbine-agnostic surrogates represents an important long-term research challenge. Such a capability would improve transferability across turbine platforms and reduce the need to generate new training databases whenever a different turbine design or control architecture is considered. While the present work establishes the surrogate-design principles required for such future developments, the question of cross-turbine generalization remains open and will likely require new forms of reduced-order parameterization capable of representing turbine-specific characteristics.

Finally, the present work provides a practical framework for the construction and deployment of fatigue-load surrogate models. By identifying a converged reduced-order inflow representation and demonstrating its robustness across multiple inflow-estimation environments, this study establishes a set of design principles that can be applied independently of a particular application. By establishing and evaluating a set of practical design and implementation principles, the present work provides a foundation upon which future application-oriented developments can build, while further improvements to the surrogate formulation itself remain an active area of research. Potential applications include load-aware wind-farm control, wind-farm layout optimization, operational load monitoring, lifetime assessment, and large-scale design-space exploration. The combination of physical interpretability, computational efficiency, and demonstrated robustness presented in this work makes such applications a particularly promising direction for future research. More broadly, the results suggest that future engineering wake-model development may benefit from placing greater emphasis on the accurate representation of load-driving inflow heterogeneity, in addition to the traditional objective of predicting power and wake losses.

Appendix A: Nomenclature and abbreviations

ANN	Artificial neural network
DEL	Damage equivalent load
1470 DWM	Dynamic wake meandering
GCH	Gaussian-based velocity model with Crespo-Hernandez turbulence model
GIQ	Gaussian-based velocity model with Ishihara–Qian turbulence model
HPO	Hyperparameter optimization
IEA	International Energy Agency
1475 LES	Large eddy simulation
MISO	Multiple-input single-output
NRMSE	Normalized root mean square error
OBS	Wind observer
RAHS	Rotor-averaged horizontal shear
1480 RATI	Rotor-averaged turbulence intensity
RAVS	Rotor-averaged vertical shear



	RAWS	Rotor-averaged wind speed
	RETI	Rotor-effective turbulence intensity
	REWS	Rotor-effective wind speed
1485	SATI	Sector-averaged turbulence intensity
	SAWS	Sector-averaged wind speed
	SETI	Sector-effective turbulence intensity
	SEWS	Sector-effective wind speed
	VS	Virtual sensor (in simulation environment)
1490	WFC	Wind farm control
	WT	Wind turbine
	$\mathbf{c}$	Vector of turbine control parameters
1495	C_P	Power coefficient
	C_T	Thrust coefficient
	D	Rotor diameter
	f_{train}	Fraction of samples allocated to training (with regard to the training + validation dataset)
	$\mathbf{i}$	Vector of reduced-order local inflow parameters
1500	j	Index variable for ranges of load cycle amplitude
	k	Index variable for samples in a dataset
	l	Index variable for structural load channels
	L_j	Load cycle amplitude
	n_j	Number of load cycles with amplitude L_j
1505	N	Total number of samples in a dataset
	N_{eq}	Equivalent number of load cycles
	N_{seed}	Number of turbulent seed realizations included in the training quantities
	r	Radial position on the rotor disk
	R	Rotor radius
1510	s	Index variable for geometrical sectors across the rotor disk
	t	Time
	T	Duration of the considered time window
	u	Instantaneous streamwise wind velocity component
	y	True value of a surrogate output
1515	z_H	Hub height



	Ω	Rotor speed
	ψ	Azimuthal position on the rotor disk (null when the blade points vertically up)
1520	ψ_s^-	Lower azimuthal boundary of sector s
	ψ_s^+	Upper azimuthal boundary of sector s
	$\mathcal{G}(\cdot)$	Mapping function between two different inflow descriptions
1525	$\mathcal{L}_l(\cdot)$	Surrogate mapping function for structural load channel l
	$\square_{\text{eq}}$	Equivalent
	$\square_{\text{max}}$	Maximum value
	$\square_{\text{min}}$	Minimum value
1530	$\square_{\text{ref}}$	Reference value
	$\hat{\square}$	Predicted value
	$\overline{\square}$	Mean or temporally averaged value
1535	$\square_{4A}$	Quantity corresponding to the four-sector configuration with sectors centered on the vertical axis
	A	Sector configuration in which a sector is centered on the vertical axis
	B	Sector configuration in which a sector boundary aligns with the vertical axis
1540	<i>Note:</i> The notation 2A, 2B, 4A, 4B, etc., denotes the number of azimuthal sectors followed by the corresponding sector-placement convention (A or B), as illustrated in detail in Fig. 2 and Fig. 5.	



Appendix B: Implementation of turbine curtailment for the IEA 3.4 MW reference turbine

This appendix summarizes the turbine curtailment strategy considered throughout the present work, which is representative of applications such as load mitigation and wind farm induction control. The implementation is illustrated through the steady-state regulation trajectories obtained for different power-demand levels. For reduced power demand, the transition to rated operation occurs at lower wind speeds, whereas the below-rated operating regime remains unchanged. In the FAST.Farm simulations, these trajectories are enforced through a modified DISCON.dll controller specifically adapted for curtailment operation.

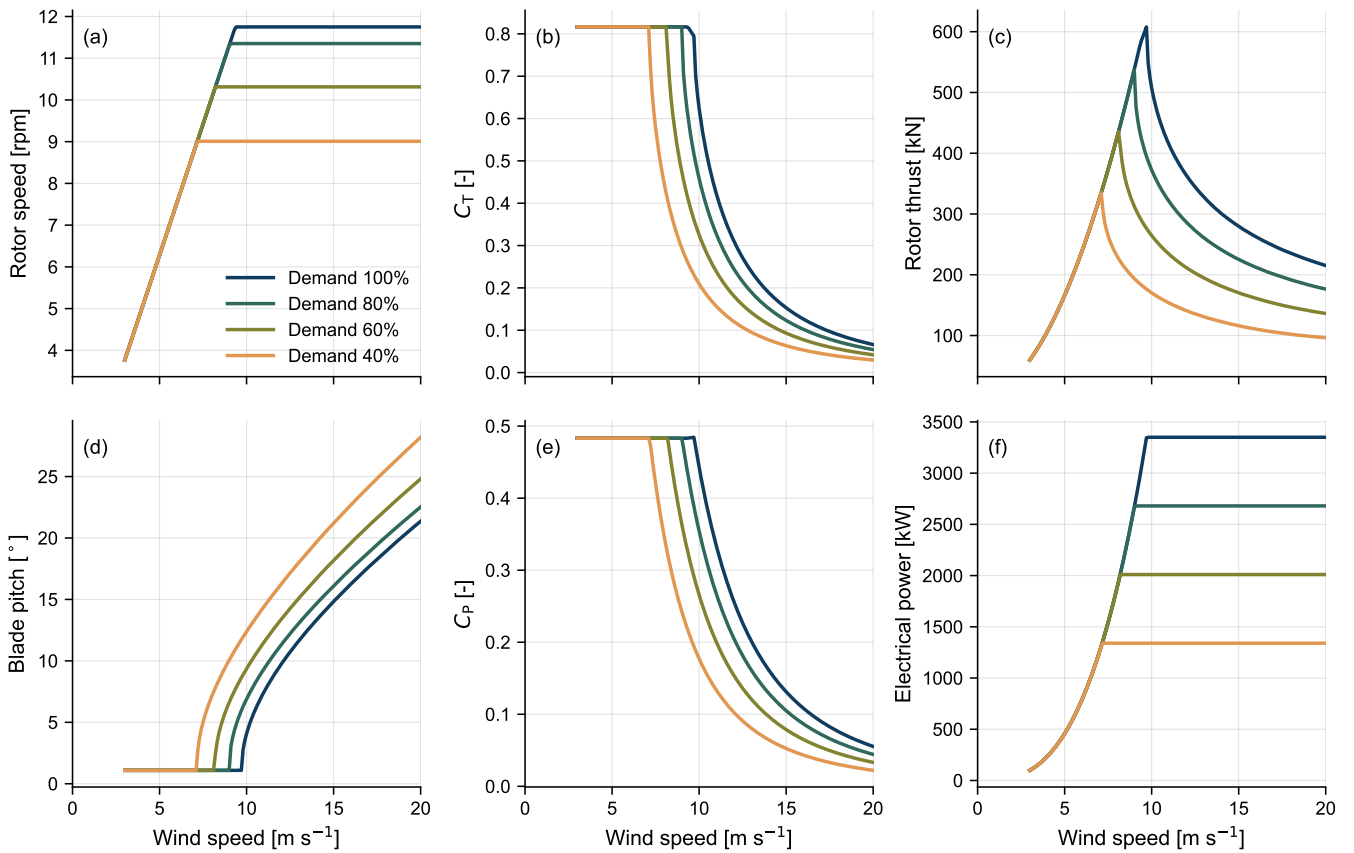


Figure B1. Steady-state regulation trajectories of the IEA 3.4 MW reference turbine under different power-demand levels, expressed as a percentage of rated power. Each curve corresponds to a distinct curtailment setting. The controller adjusts the rotor-speed (a) and blade-pitch (d) setpoints to track the corresponding trajectories of thrust coefficient C_T (b), rotor thrust (c), power coefficient C_P (e), and electrical power output (f).



Appendix C: Inflow mapping technique: averaged and effective quantities, all from virtual sensors in DWM simulations

1550 This appendix provides the detailed comparison between the “averaged” and “effective” reduced-order inflow quantities discussed in Sect. 4.2. The figure illustrates the relationship between the two inflow representations for the combined training and validation datasets. These results support the conclusion that both formulations contain largely equivalent information for the purpose of fatigue-load surrogate modeling, despite the different physical averaging procedures used to derive the quantities.

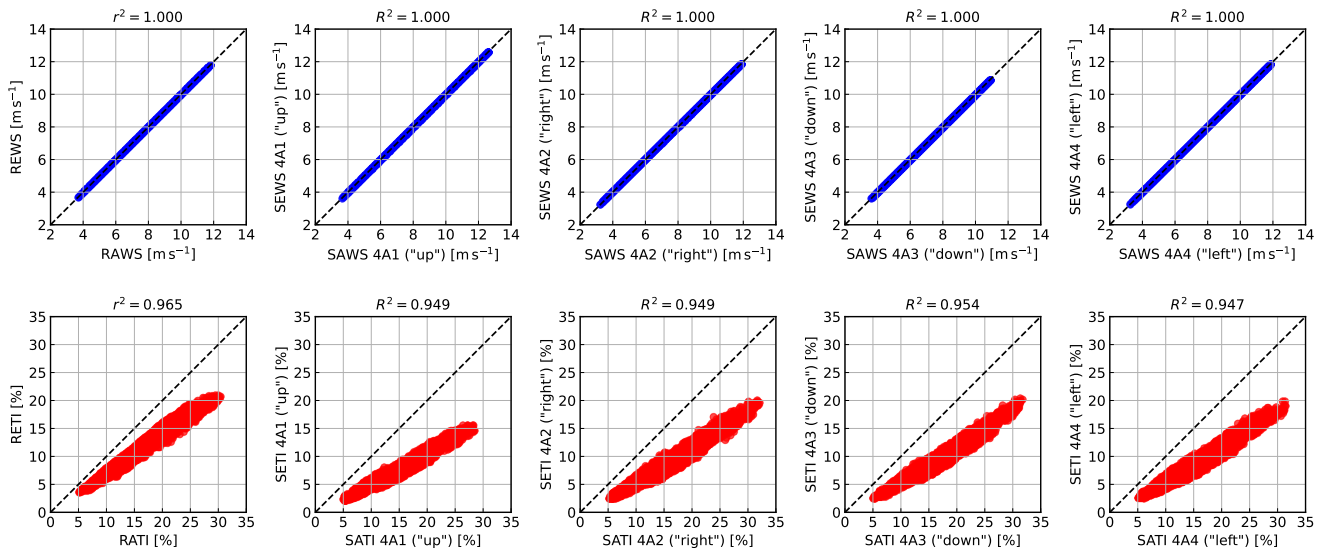


Figure C1. Comparison of averaged quantities (RAWS, SAWS, RATI and SATI) and effective quantities (REWS, SEWS, RETI and SETI) from virtual sensors in dynamic environment (DWM implemented in FAST.Farm) for the whole training and validation datasets combined.

Appendix D: Inflow mapping technique: virtual sensors vs. wind observers

1555 This appendix provides the complete comparison between inflow quantities obtained from virtual sensors (VS) and those estimated using wind observers (OBS). These results complement the discussion presented in Sect. 5.1 by providing the detailed point-by-point comparison for all investigated reduced-order inflow quantities. Figures D1 and D2 respectively show the averaged- and effective-inflow formulations.

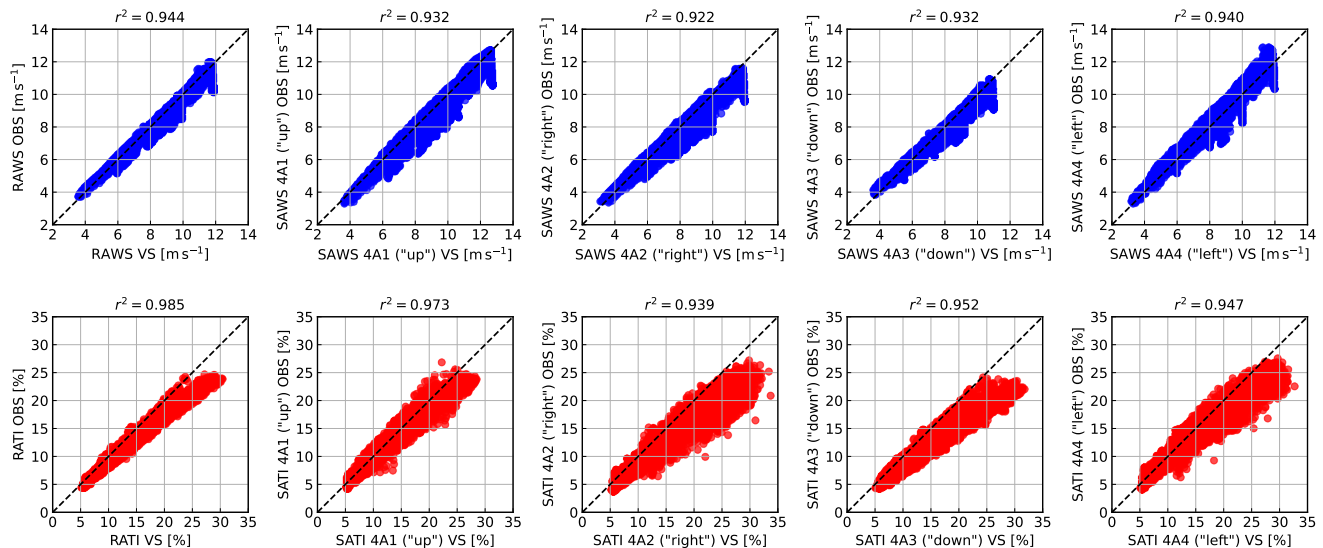


Figure D1. Comparison of RAWs, SAWS, RATI and SATI measured by virtual sensors (VS) and estimated by the wind observers (OBS) for the whole training and validation datasets combined.

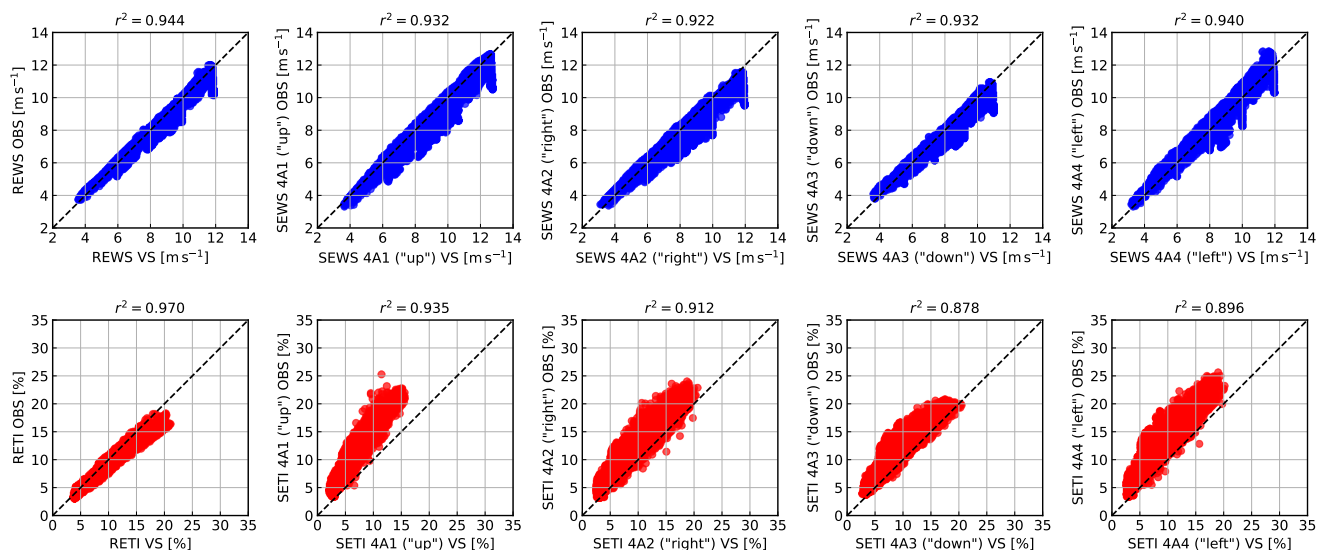


Figure D2. Comparison of REWS, SEWS, RETI and SETI measured by virtual sensors (VS) and estimated by the wind observers (OBS) for the whole training and validation datasets combined.



Appendix E: Inflow mapping technique: dynamic vs. steady-state wake models

1560 This appendix provides the detailed comparisons supporting the discussion of Sect. 5.2. The figures illustrate the relationship
 between reduced-order inflow quantities derived from the dynamic wake meandering model and those obtained from the two
 steady-state wake-model formulations investigated in this work. Together, they provide additional insight into the mapping
 between inflow-description spaces discussed in the main text.

For the GCH configuration, the default FLORIS (NREL, 2025) wake-added turbulence parameters were used (initial =
 1565 0.1, constant = 0.5, $a_i = 0.8$, and downstream = -0.32). For the GIQ configuration, the parameter values recommended by
 Ishihara and Qian (2018) were retained.

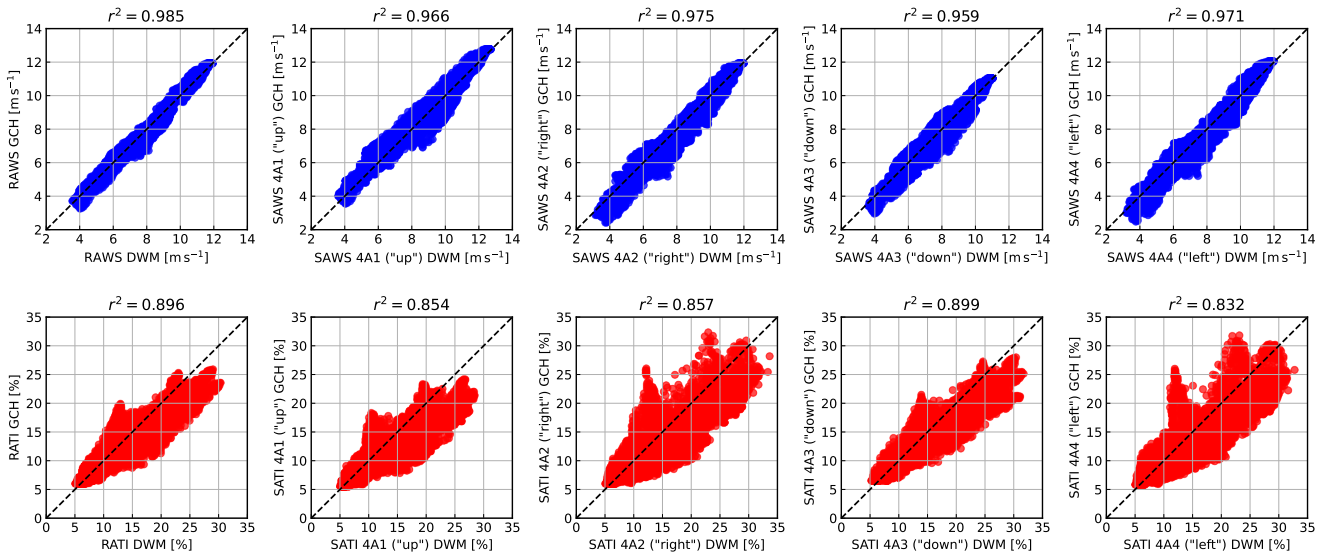


Figure E1. Comparison of RAWs, SAWS, RATI and SATI using dynamic wake meandering model (DWM) and steady-state wake model (GCH, Gaussian-based velocity model + Crespo-Hernandez turbulence model) inflow conditions across the combined training and validation datasets.

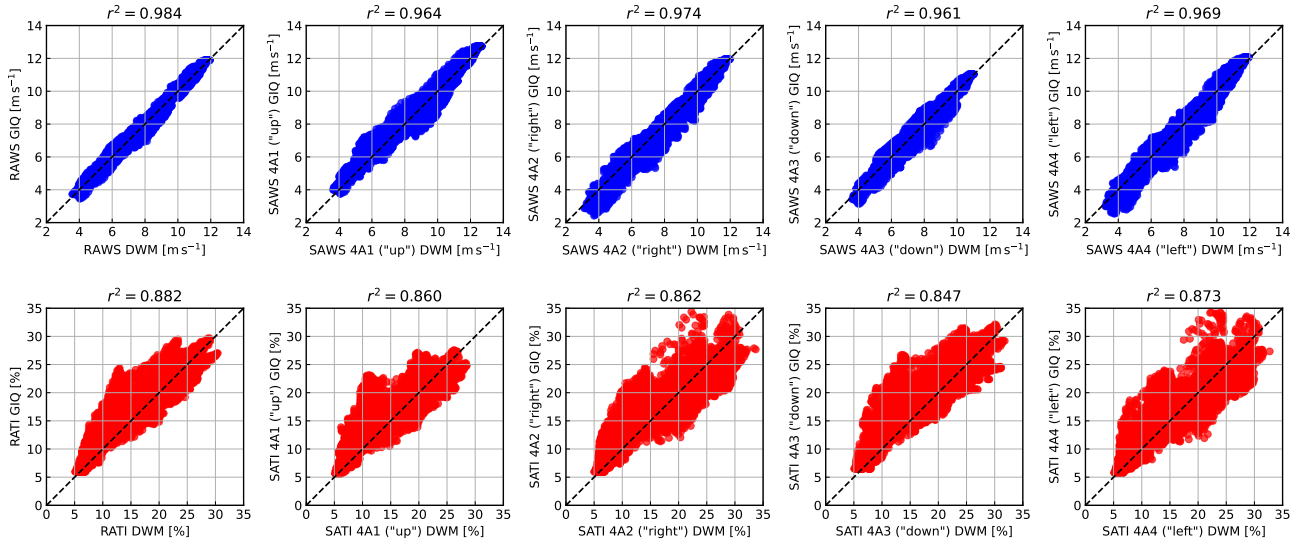


Figure E2. Comparison of RAWs, SAWS, RATI and SATI using dynamic wake meandering model (DWM) and steady-state wake model (GIQ, Gaussian-based velocity model + Ishihara–Qian turbulence model) inflow conditions across the combined training and validation datasets.

Appendix F: ANN grid search results for all sector combinations

This appendix presents the complete results of the ANN hyperparameter optimization performed for all investigated sector configurations. The figure complements the analysis of Sect. 4.4 by providing the projected grid-search results used to identify the surrogate architectures retained throughout this work. The objective of this appendix is not to discuss individual optimal architectures, but rather to document the robustness of the selected ANN configurations across the different inflow representations considered.

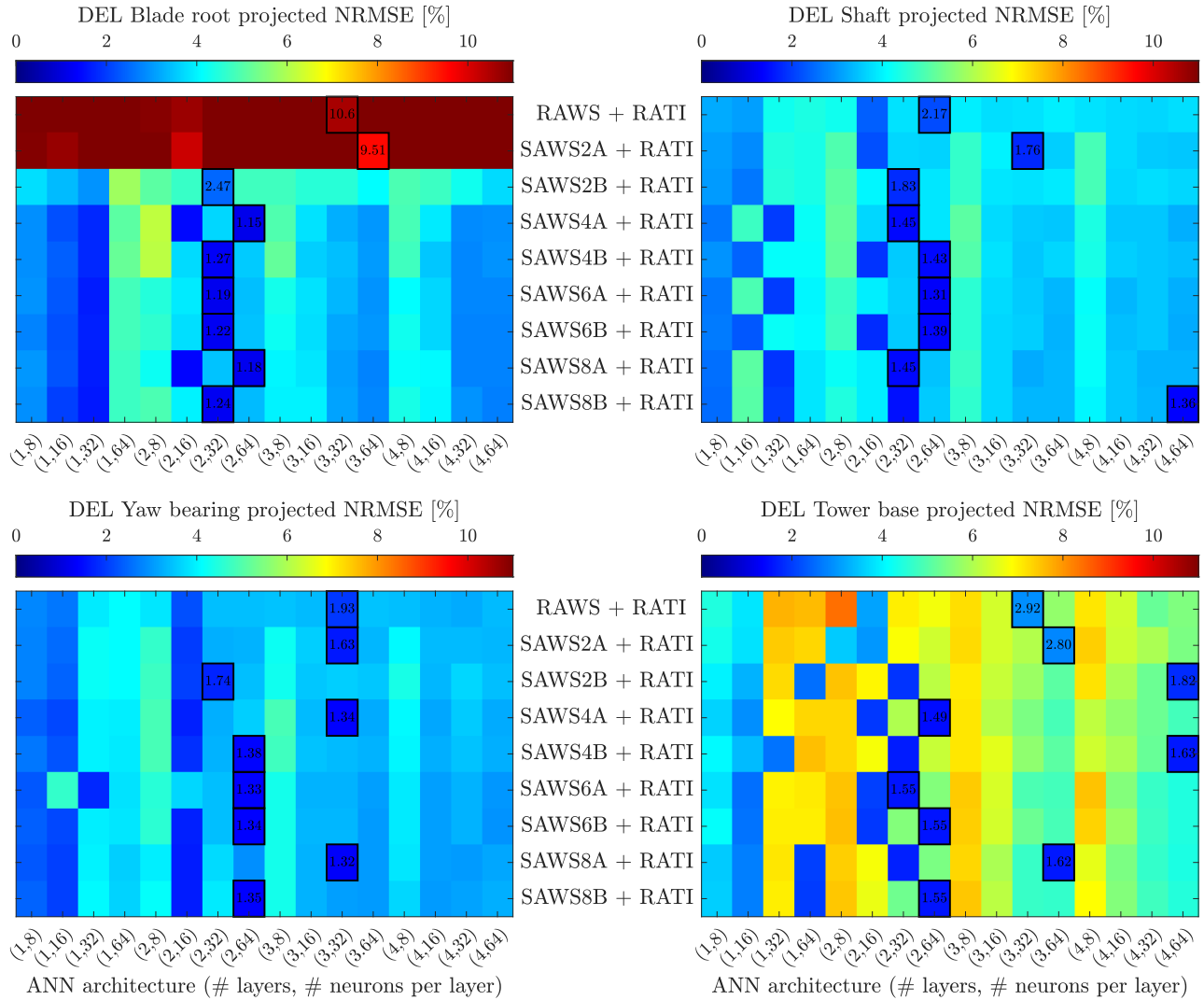


Figure F1. Projected ANN hyperparameter optimization (HPO) grid-search results for the investigated sector configurations. All reported NRMSE are computed on the validation dataset used for model selection. The optimal hyperparameter configuration for each sector setup is highlighted by a square, and the corresponding NRMSE value is reported.

Appendix G: Surrogate results for various levels of inflow mapping when linked with steady-state wake models

The analyses presented in Sect. 4 were primarily conducted using inflow quantities derived from the dynamic wake meandering model. A legitimate question is therefore whether the same conclusions regarding the optimal inflow representation remain



valid when the surrogate is instead trained and deployed directly using inflow quantities obtained from steady-state wake models.

To address this question, the complete convergence analyses were repeated using inflow quantities derived from the steady-state wake-model formulations considered in this work. The results presented in this appendix show that the main conclusions remain unchanged. In particular, the convergence with increasing spatio-temporal inflow-description complexity closely follows the trends observed using dynamic-wake-model inflow quantities, while the convergence with increasing sector discretization leads to the same recommended surrogate configuration.

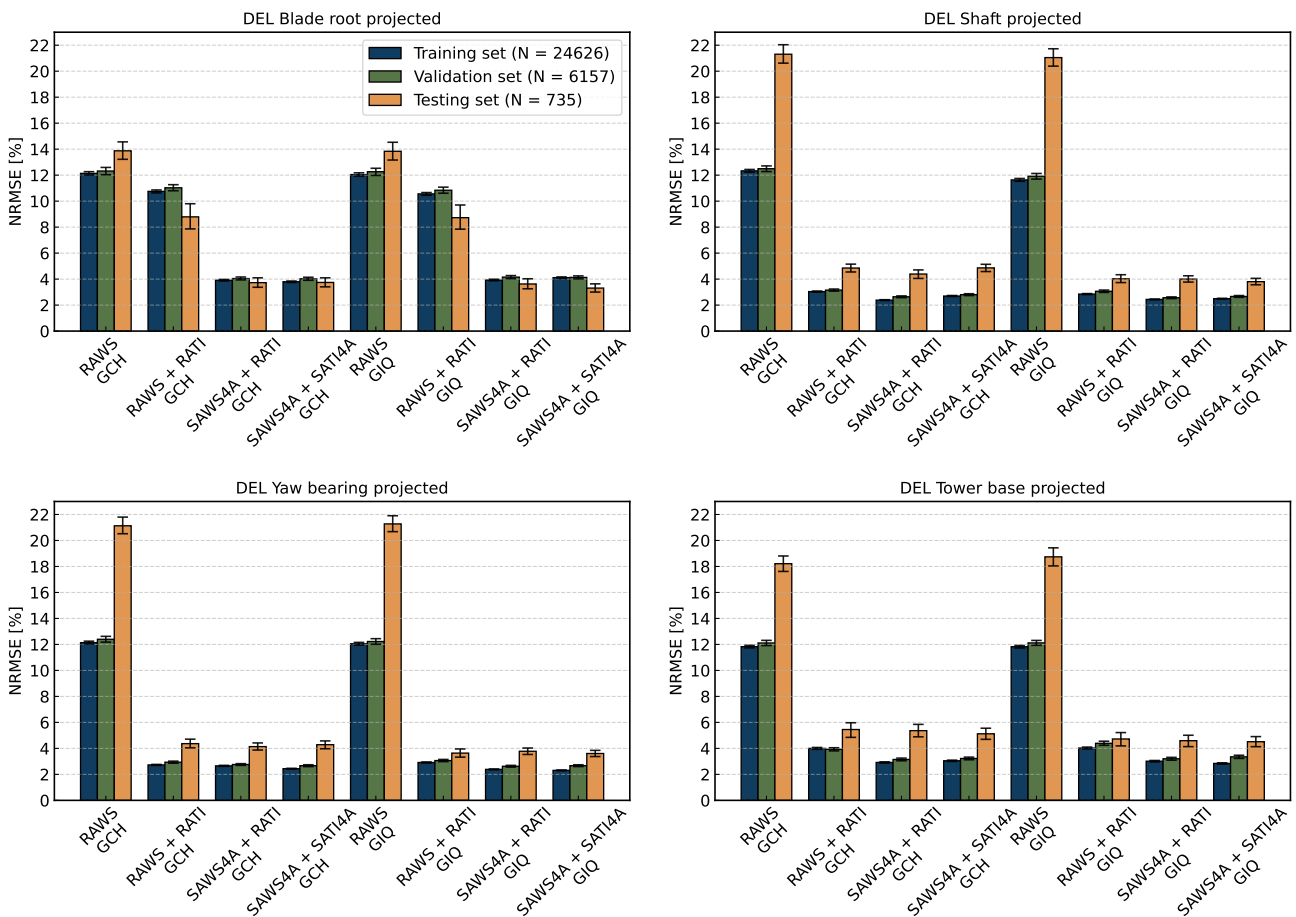


Figure G1. Performance of surrogates by choice of spatio-temporal reduced parameters, using the local inflow derived from steady-state wake models (FLORIS Crespo-Hernandez & Ishihara) for training and application.

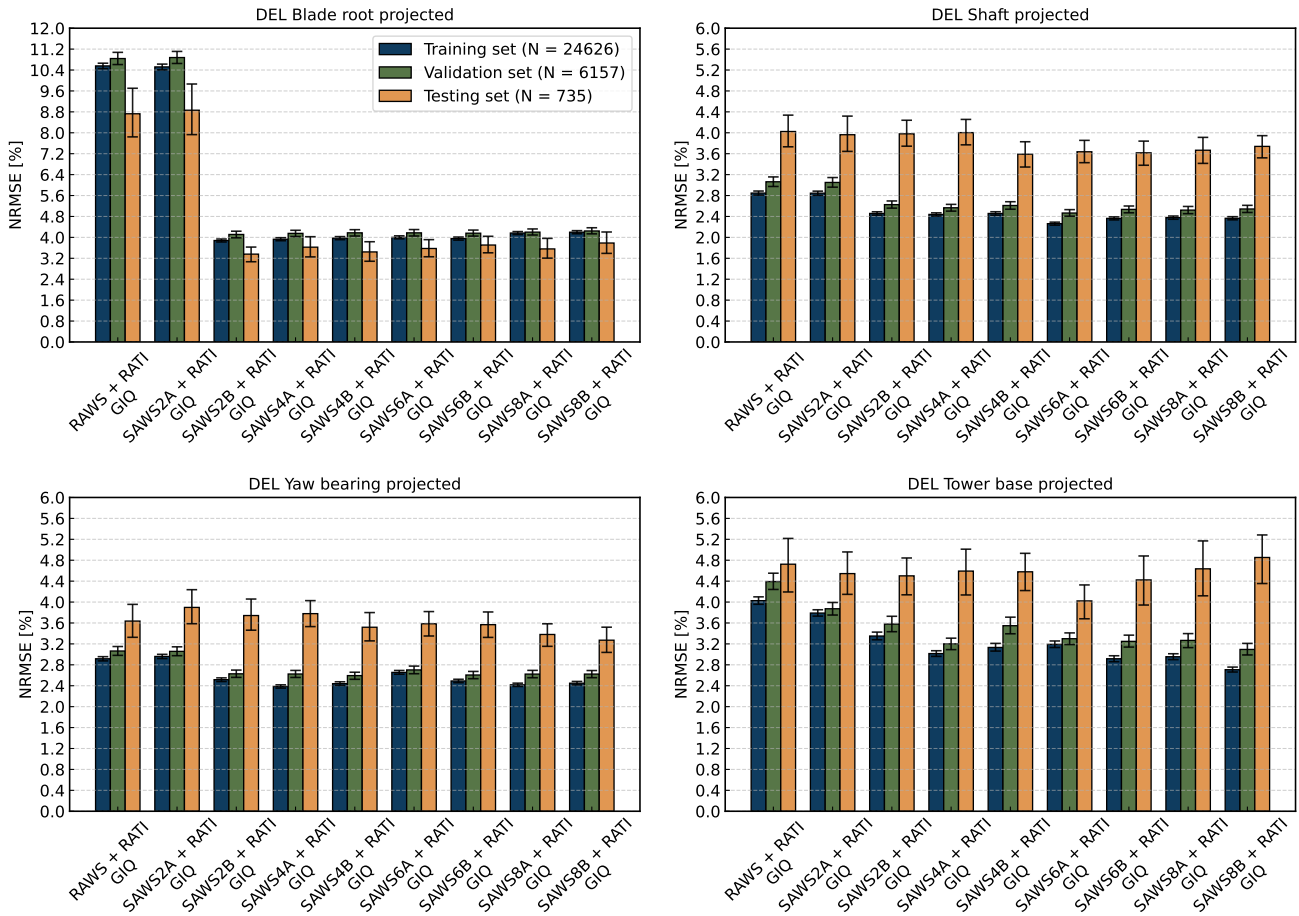


Figure G2. Performance of surrogates by increasing number and position of sectors, using inflows derived from steady-state wake model (Ishihara) for training and application. Here the sectors are applied to wind speeds (SAWS), while fixed rotor-averaged turbulence intensities (RATI) are used.

Code and data availability. The datasets, trained TensorFlow surrogate models and the associated source codes are publicly available on Zenodo at <https://doi.org/10.5281/zenodo.23372270> (Guilloré et al., 2026).

1585 *Author contributions.* AG led the technical development of the study, including the development of the surrogate-model formulation, generation of the datasets, training of the surrogate models, and analysis of surrogate design choices and implementation strategies. AG prepared the initial manuscript, which was subsequently reviewed and revised with input from the co-authors. CLB developed the concept of a location-agnostic load surrogate model based on local inflow quantities and played a key role in shaping and refining the methodology throughout its development. CLB provided scientific guidance and supervision and contributed to the preparation and revision of the manuscript. AV



1590 derived the local inflow quantities from the steady-state engineering flow models and contributed to writing the corresponding section of
the manuscript. FC contributed to the conceptualization and development of the surrogate-model formulation through technical discussions
and feedback, particularly regarding its applicability to observers and steady-state wake models. All authors provided valuable input to this
research work through discussions, feedback, and improvement of the manuscript.

Competing interests. The authors have the following competing interests: At least one of the (co-)authors is a member of the editorial board
1595 of Wind Energy Science.

Financial support. This work has been supported by the MERIDIONAL and the SUDOCO projects, which receive funding from the Euro-
pean Union's Horizon Europe Programme under the grant agreement No. 101084216 and No. 101122256, respectively.

Acknowledgements. During the preparation of this manuscript, the authors used M365 Copilot, powered by the GPT-5 chat model and
provided by Microsoft, to improve the grammar, readability, and style of the text. The tool was not used to generate scientific content,
1600 perform data analysis, interpret results, or draw conclusions. All scientific content, including the methodology, analyses, results, discussion,
and conclusions, was developed by the authors before the tool was used for language editing. The authors reviewed and edited all suggested
changes and take full responsibility for the content of this publication.

The authors thank Marta Bertelè for implementing the wind observers for the IEA 3.4 MW wind turbine within the simulation environment.



References

- 1605 Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, <https://doi.org/10.48550/arXiv.1603.04467>, 2016.
- 1610 Abbas, N. J., Zalkind, D. S., Pao, L., and Wright, A.: A reference open-source controller for fixed and floating offshore wind turbines, *Wind Energy Science*, 7, 53–73, <https://doi.org/10.5194/wes-7-53-2022>, 2022.
- Anand, A., Braunbehrens, R., Guilloré, A., and Bottasso, C. L.: Economic lifetime-aware wind farm control, *Wind Energy Science Discussions*, 2025, 1–39, <https://doi.org/10.5194/wes-2025-67>, 2025.
- Ayachit, U.: The ParaView Guide: A Parallel Visualization Application, Kitware, Inc., Clifton Park, NY, USA, ISBN 978-1-930934-30-6, <https://www.paraview.org>, 2015.
- 1615 Baas, P., Verzijlbergh, R., van Dorp, P., and Jonker, H.: Investigating energy production and wake losses of multi-gigawatt offshore wind farms with atmospheric large-eddy simulation, *Wind Energy Science*, 8, 787–805, <https://doi.org/10.5194/wes-8-787-2023>, 2023.
- Bastankhah, M. and Porté-Agel, F.: Experimental and theoretical study of wind turbine wakes in yawed conditions, *Journal of Fluid Mechanics*, 806, 506–541, <https://doi.org/10.1017/jfm.2016.595>, 2016.
- 1620 Bay, C. J., Guilloré, A., Bortolotti, P., Frontin, C., Sinner, M., Thedin, R., and Bottasso, C. L.: Impacts on damage equivalent loads from plant optimization using artificial neural network surrogate models, *Journal of Physics: Conference Series*, 3224, 032 125, <https://doi.org/10.1088/1742-6596/3224/3/032125>, 2026.
- Bertelè, M., Meyer, P. J., Sucameli, C. R., Fricke, J., Wegner, A., Gottschall, J., and Bottasso, C. L.: The rotor as a sensor – observing shear and veer from the operational data of a large wind turbine, *Wind Energy Science*, 9, 1419–1429, <https://doi.org/10.5194/wes-9-1419-2024>, 2024.
- 1625 Bodini, N., Zardi, D., and Lundquist, J. K.: Three-dimensional structure of wind turbine wakes as measured by scanning lidar, *Atmospheric Measurement Techniques*, 10, 2881–2896, <https://doi.org/10.5194/amt-10-2881-2017>, 2017.
- Bortolotti, P., Tarrés, H. C., Dykes, K., Merz, K., Sethuraman, L., Verelst, D., and Zahle, F.: IEA Wind TCP Task 37: Systems Engineering in Wind Energy – WP2.1 Reference Wind Turbines, Tech. Rep. NREL/TP-5000-73492, National Renewable Energy Laboratory, <https://doi.org/10.2172/1529216>, 2019.
- 1630 Bossanyi, E.: Surrogate model for fast simulation of turbine loads in wind farms, *Journal of Physics: Conference Series*, 2265, 042 038, <https://doi.org/10.1088/1742-6596/2265/4/042038>, 2022.
- Bottasso, C. L., Cacciola, S., and Schreiber, J.: Local wind speed estimation, with application to wake impingement detection, *Renewable Energy*, 116, 155–168, <https://doi.org/10.1016/j.renene.2017.09.044>, 2018.
- 1635 Braunbehrens, R., Anand, A., Campagnolo, F., and Bottasso, C. L.: First experimental results on lifetime-aware wind farm control, *Journal of Physics: Conference Series*, 2767, 032 042, <https://doi.org/10.1088/1742-6596/2767/3/032042>, 2024.
- Calaf, M., Meneveau, C., and Meyers, J.: Large eddy simulation study of fully developed wind-turbine array boundary layers, *Physics of Fluids*, 22, 015 110, <https://doi.org/10.1063/1.3291077>, 2010.
- Churchfield, M. J. and Lee, S.: NWTC Design Codes – SOWFA, Technical report, National Renewable Energy Laboratory, Golden, CO, USA, <https://www.nrel.gov/wind/nwtc/sowfa>, 2012.
- 1640



- Churchfield, M. J., Lee, S., Moriarty, P. J., Martínez, L. A., Leonardi, S., Vijayakumar, G., and Brasseur, J. G.: A Large-Eddy Simulation of Wind-Plant Aerodynamics, in: 50th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, AIAA 2012-0537, Nashville, Tennessee, USA, <https://doi.org/10.2514/6.2012-537>, 2012.
- Crespo, A. and Hernández, J.: Turbulence characteristics in wind-turbine wakes, *Journal of Wind Engineering and Industrial Aerodynamics*, 1645 61, 71–85, [https://doi.org/10.1016/0167-6105\(95\)00033-X](https://doi.org/10.1016/0167-6105(95)00033-X), 1996.
- Dimitrov, N., Kelly, M. C., Vignaroli, A., and Berg, J.: From wind to loads: wind turbine site-specific load estimation with surrogate models trained on high-fidelity load databases, *Wind Energy Science*, 3, 767–790, <https://doi.org/10.5194/wes-3-767-2018>, 2018.
- Göçmen, T., van der Laan, P., Réthoré, P.-E., Peña Díaz, A., Larsen, G. C., and Ott, S.: Wind turbine wake models developed at the Technical University of Denmark: A review, *Renewable and Sustainable Energy Reviews*, 60, 752–769, <https://doi.org/10.1016/j.rser.2016.01.113>, 1650 2016.
- Guilloré, A., Vad, A., Campagnolo, F., and Bottasso, C. L.: Datasets as supplement for the journal publication "Design and implementation of a location-agnostic and control-oriented load surrogate model for wind turbines", <https://doi.org/10.5281/zenodo.23372270>, dataset., 2026.
- Guilloré, A., Campagnolo, F., and Bottasso, C. L.: A control-oriented load surrogate model based on sector-averaged inflow quantities: capturing damage for unwaked, waked, wake-steering and curtailed wind turbines, *Journal of Physics: Conference Series*, 2767, 032 019, 1655 <https://doi.org/10.1088/1742-6596/2767/3/032019>, 2024.
- He, R., Yang, H., Sun, S., Lu, L., Sun, H., and Gao, X.: A machine learning-based fatigue loads and power prediction method for wind turbines under yaw control, *Applied Energy*, 326, 120 013, <https://doi.org/10.1016/j.apenergy.2022.120013>, 2022.
- Hornik, K., Stinchcombe, M., and White, H.: Multilayer Feedforward Networks are Universal Approximators, *Neural Networks*, 2, 359–366, 1660 [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8), 1989.
- International Electrotechnical Commission: IEC 61400–1: Wind energy generation systems – Part 1: Design requirements, Tech. rep., International Electrotechnical Commission, 2019.
- Ishihara, T. and Qian, G.-W.: A new Gaussian-based analytical wake model for wind turbines considering ambient turbulence intensities and thrust coefficient effects, *Journal of Wind Engineering and Industrial Aerodynamics*, 177, 275–292, 1665 <https://doi.org/10.1016/j.jweia.2018.04.010>, 2018.
- Jensen, N. O.: A Note on Wind Generator Interaction, Technical Report Risø-M-2411, Risø National Laboratory, Roskilde, Denmark, ISBN 87-550-0971-9, https://orbit.dtu.dk/files/55857682/ris_m_2411.pdf, 1983.
- Jonkman, B. J.: TurbSim User's Guide: Version 1.50, Tech. Rep. NREL/TP-500-46198, National Renewable Energy Laboratory (NREL), <https://doi.org/10.2172/965520>, 2009.
- 1670 Jonkman, J., Doubrava, P., Hamilton, N., Annoni, J., and Fleming, P.: Validation of FAST.Farm Against Large-Eddy Simulations, *Journal of Physics: Conference Series*, 1037, 062 005, <https://doi.org/10.1088/1742-6596/1037/6/062005>, 2018.
- Kainz, S., Zhu, Z., Guilloré, A., Vad, A., and Bottasso, C. L.: Balancing energy yield and turbine fatigue: a load-aware approach to wind farm layout optimization, *Journal of Physics: Conference Series*, 3224, 032 041, <https://doi.org/10.1088/1742-6596/3224/3/032041>, 2026.
- Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, in: 3rd International Conference on Learning Representations (ICLR 2015), Conference Track Proceedings, <https://doi.org/10.48550/arXiv.1412.6980>, 2015.
- 1675 Kretschmer, M., Jonkman, J., Pettas, V., and Cheng, P. W.: FAST.Farm load validation for single wake situations at alpha ventus, *Wind Energy Science*, 6, 1247–1262, <https://doi.org/10.5194/wes-6-1247-2021>, 2021.



- Kuhn, M., Henry de Frahan, M., Mohan, P., Deskos, G., Churchfield, M., Cheung, L., Sharma, A., Almgren, A., Ananthan, S., Brazell, M., Martinez-Tossas, L., Thedin, R., Rood, J., Sakievich, P., Vijayakumar, G., Zhang, W., and Sprague, M.: AMR-Wind: A Performance-Portable, High-Fidelity Flow Solver for Wind Farm Simulations, *Wind Energy*, 28, e70 010, <https://doi.org/10.1002/we.70010>, 2025.
- Larsen, G. C.: Dynamic Wake Meandering Modeling, Tech. Rep. Risø-R-1607(EN), Risø National Laboratory for Sustainable Energy, Technical University of Denmark, Roskilde, Denmark, <https://orbit.dtu.dk/en/publications/dynamic-wake-meandering-modeling>, 2007.
- Letizia, S., Brugger, P., Bodini, N., Krishnamurthy, R., Scholbrock, A., Simley, E., Porté-Agel, F., Hamilton, N., Doubrawa, P., and Moriarty, P.: Characterization of Wind Turbine Flow Through Nacelle-Mounted Lidars: A Review, *Frontiers in Mechanical Engineering*, 9, 1261 017, <https://doi.org/10.3389/fmech.2023.1261017>, 2023.
- Li, C., Abraham, A., Li, B., and Hong, J.: Incoming flow measurements of a utility-scale wind turbine using super-large-scale particle image velocimetry, *Journal of Wind Engineering and Industrial Aerodynamics*, 197, 104 074, <https://doi.org/10.1016/j.jweia.2019.104074>, 2020.
- Liew, J., Göçmen, T., Lio, A. W. H., and Larsen, G. C.: Extending the dynamic wake meandering model in HAWC2Farm: a comparison with field measurements at the Lillgrund wind farm, *Wind Energy Science*, 8, 1387–1402, <https://doi.org/10.5194/wes-8-1387-2023>, 2023a.
- Liew, J., Riva, R., and Göçmen, T.: Efficient Mann turbulence generation for offshore wind farms with applications in fatigue load surrogate modelling, *Journal of Physics: Conference Series*, 2626, 012 050, <https://doi.org/10.1088/1742-6596/2626/1/012050>, 2023b.
- Liew, J., Riva, R., Friis-Møller, M., and Göçmen, T.: Wind Farm Control Optimisation Under Load Constraints Via Surrogate Modelling, *Journal of Physics: Conference Series*, 2767, 092 039, <https://doi.org/10.1088/1742-6596/2767/9/092039>, 2024.
- Lokken, T. S., Baricchio, M., Gebraad, P. M. O., Guilloré, A., and van Wingerden, J.-W.: Tuning of a holistic synchronous multi-objective framework for wind farm flow control, *Journal of Physics: Conference Series*, 3224, 032 117, <https://doi.org/10.1088/1742-6596/3224/3/032117>, 2026.
- Meyers, J., Bottasso, C. L., Dykes, K., Fleming, P., Gebraad, P., Giebel, G., Göçmen, T., and van Wingerden, J.-W.: Wind farm flow control: prospects and challenges, *Wind Energy Science*, 7, 2271–2306, <https://doi.org/10.5194/wes-7-2271-2022>, 2022.
- Miner, M. A.: Cumulative Damage in Fatigue, *Journal of Applied Mechanics, Transactions ASME*, 12, A159–A164, <https://doi.org/10.1115/1.4009458>, 1945.
- Niayifar, A. and Porté-Agel, F.: Analytical Modeling of Wind Farms: A New Approach for Power Prediction, *Energies*, 9, <https://doi.org/10.3390/en9090741>, 2016.
- NREL: FLORIS v4.5, <https://doi.org/10.5281/zenodo.17186844>, 2025.
- NREL: OpenFAST: Open-source wind turbine simulation software, GitHub repository, <https://github.com/OpenFAST/openfast>, 2026.
- Onnen, D., Larsen, G. C., Lio, A. W. H., Hulsman, P., Kühn, M., and Petrović, V.: Field comparison of load-based wind turbine wake tracking with a scanning lidar reference, *Wind Energy Science*, 11, 175–193, <https://doi.org/10.5194/wes-11-175-2026>, 2026.
- Pedersen, M. M., Steiner, J., Nilsen, M. B., Lohmann, J., Hodgson, E. L., Riva, R., Troldborg, N., Andersen, S. J., Larsen, G., Verelst, D. R., and Réthoré, P.-E.: Dynamiks: An open-source Dynamic Wind System Simulator, GitLab [code], <https://gitlab.windenergy.dtu.dk/DYNAMIKS/dynamiks>, last access: June 2026, 2023a.
- Pedersen, M. M., van der Laan, P., Friis-Møller, M., Meyer Forsting, A., Riva, R., Alcayaga Romàn, L. A., Criado Risco, J., Quick, J., Christiansen, J. P. S., Olsen, B. T., Rodrigues, R. V., and Réthoré, P.-E.: DTUWindEnergy/PyWake: PyWake (v2.5.0), <https://doi.org/10.5281/zenodo.6806136>, 2023b.
- Pettas, V. and Cheng, P. W.: Surrogate modeling and aeroelastic analysis of a wind turbine with down-regulation, power boosting, and IBC capabilities, *Energies*, 17, 1284, <https://doi.org/10.3390/en17061284>, 2024.



- 1715 Porchetta, S., Munters, W., Lejeune, M., Borgers, R., Buckingham, S., and Howland, M. F.: Annual wake impacts in and between wind farm clusters – Part 1: WRF-simulated wake losses for different atmospheric conditions, *Wind Energy Science*, 11, 3273–3294, <https://doi.org/10.5194/wes-11-3273-2026>, 2026.
- Ramaswamy, H., Guilloiré, A., Riva, R., Quick, J., Pettas, V., Murcia, J. P., Campagnolo, F., Bottasso, C. L., and Réthoré, P. E.: Comparison of wind turbine load surrogate model performances with multi-scale local inflow parameterization, *Journal of Physics: Conference Series*, 3224, 032 073, <https://doi.org/10.1088/1742-6596/3224/3/032073>, 2026.
- 1720 Schreiber, J., Bottasso, C. L., and Bertelè, M.: Field testing of a local wind inflow estimator and wake detector, *Wind Energy Science*, 5, 867–884, <https://doi.org/10.5194/wes-5-867-2020>, 2020.
- Shaler, K. and Jonkman, J.: FAST.Farm development and validation of structural load prediction against large eddy simulations, *Wind Energy*, 24, 428–449, <https://doi.org/10.1002/we.2581>, 2021.
- 1725 Shaler, K., Jasa, J., and Barter, G. E.: Efficient loads surrogates for waked turbines in an array, *Journal of Physics: Conference Series*, 2265, 032 095, <https://doi.org/10.1088/1742-6596/2265/3/032095>, 2022.
- Simley, E., Fleming, P., Girard, N., Alloin, L., Godefroy, E., and Duc, T.: Results from a wake-steering experiment at a commercial wind plant: investigating the wind speed dependence of wake-steering performance, *Wind Energy Science*, 6, 1427–1453, <https://doi.org/10.5194/wes-6-1427-2021>, 2021.
- 1730 Steiner, J., Hodgson, E. L., van der Laan, M. P., Alcayaga, L., Pedersen, M., Andersen, S. J., Larsen, G., and Réthoré, P.-E.: A multi-fidelity model intercomparison for wake steering of a large turbine in a conventionally neutral atmospheric boundary layer, *Wind Energy Science*, 11, 1679–1703, <https://doi.org/10.5194/wes-11-1679-2026>, 2026.
- Sutherland, H. J.: On the Fatigue Analysis of Wind Turbines, Tech. Rep. SAND99-0089, Sandia National Laboratories, Albuquerque, NM, USA, 1999.
- 1735 Vad, A., Guilloiré, A., Anand, A., Pettas, V., Shah, A. H., Lizarraga-Saenz, I., Aparicio-Sanchez, M., Eguinoa, I., Conti Gost, N., Tsaklis, I., Frère, A., Hermans, K. W., Kamau, J. K., Dimitrov, N., Göçmen, T., and Bottasso, C. L.: Modeling wind farm response: a modular, integrated, and multi-stakeholder approach, *Wind Energy Science Discussions*, 2026, 1–49, <https://doi.org/10.5194/wes-2026-45>, 2026a.
- Vad, A., Shah, A. H., Guilloiré, A., Dröse, M., and Bottasso, C. L.: Turbine-specific lifetime consumption estimation by integrating environmental data with a farm flow model, *Journal of Physics: Conference Series*, 3224, 032 064, <https://doi.org/10.1088/1742-6596/3224/3/032064>, 2026b.
- 1740 van der Hoek, D., Dammann, T., and van Wingerden, J.-W.: Gaussian process surrogate modeling for efficient controller tuning and fatigue load prediction of the helix wake-mixing method, *Wind Energy Science*, 11, 3171–3192, <https://doi.org/10.5194/wes-11-3171-2026>, 2026.
- Wetz, T. and Wildmann, N.: Multi-point in situ measurements of turbulent flow in a wind turbine wake and inflow with a fleet of uncrewed aerial systems, *Wind Energy Science*, 8, 515–534, <https://doi.org/10.5194/wes-8-515-2023>, 2023.
- 1745 Witter, S., Roth, J., Schelenz, R., and Jacobs, G.: Creating a real-time capable surrogate model of a wind turbine using machine learning, *Wind Energy*, 28, <https://doi.org/10.1002/we.70070>, 2025.