https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

SCADA-based calibration of analytical wake models: uncertainty-aware generalisation across offshore wind farms and the role of atmospheric stability

Diederik van Binsbergen[1], Pieter-Jan Daems[1], Timothy Verstraeten[1], Konstantinos Vratsinis[1], Amir Nejad[2], and Jan Helsen[1,3]

[1]Department of Mechanical Engineering, Vrije Universiteit Brussel (VUB), Brussels, Belgium
[2]Department of Marine Technology, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[3]Flanders Make@VUB, Flanders Make, Brussels, Belgium

Correspondence: Diederik van Binsbergen (dirk.willem.van.binsbergen@vub.be)

Abstract. This study presents a Bayesian framework for generalising SCADA-calibrated wake-model tuning parameters across offshore wind farms. The framework infers cluster-wide, farm-specific, and new-farm parameter distributions while accounting for calibration uncertainty, intra-farm variability, inter-farm variability, and residual model–data mismatch. It is applied to the Jensen and Gaussian TurbOPark wake models and extended to condition the central tuning parameter on atmospheric stability, represented by the bulk Richardson number and the inverse Monin–Obukhov length. For both wake models, unstable conditions are associated with larger tuning parameters, indicating faster wake expansion and recovery, whereas stable conditions yield smaller tuning parameters and more persistent wakes. The results show that the globally pooled stability-independent parameter should not be interpreted as a neutral-stability parameter because it reflects the stability mix at the case-study offshore site. Propagating representative stability-class-specific parameters to wake-loss estimates shows that atmospheric stability substantially affects predicted wake losses, with model differences becoming more pronounced when external wakes from surrounding wind farms are included.

1 Introduction

Turbine wakes have long been recognized as a key driver of wind farm performance (Lissaman, 1979; Jensen, 1983; Katić et al., 1987; Porté-Agel et al., 2019; Ørsted, 2019). Their importance has grown further with the rapid upscaling of the wind energy sector, where larger turbines, increasing installed capacities, and denser offshore clusters increase losses due to wake interactions (Porté-Agel et al., 2019; Ørsted, 2019; Nygaard et al., 2020, 2022). Consequently, wake-induced production losses remain a central challenge in wind farm design, operation, and modelling.

A wide range of analytical wake models have been developed to represent wake behaviour under steady-state conditions (Jensen, 1983; Katić et al., 1987; Bastankhah and Porté-Agel, 2014; Niayifar and Porté-Agel, 2016; Zong and Porté-Agel, 2020; Nygaard et al., 2020; Bastankhah et al., 2021; J G Pedersen and Nygaard, 2022). In contrast to large-eddy simulations (LES), which can capture complex flow physics through numerical solutions of the Navier-Stokes equations, analytical

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

wake models are typically based on simplified formulations of mass and momentum conservation (Bastankhah and Porté-Agel, 2014). This makes them computationally efficient and attractive for wind farm applications, including yield assessment (Sukhman et al., 2025), layout optimization (Pedersen and Larsen, 2020), and control-oriented modelling (Meyers et al., 2022). However, because analytical wake models cannot fully resolve complex nonlinear flow phenomena, they rely on calibration of key model parameters, such as the wake expansion rate (van Beek et al., 2021; Göçmen et al., 2022; Aerts et al., 2023; van Binsbergen et al., 2024).

Calibration and validation of these model parameters have been performed using Computational Fluid Dynamics (CFD) simulations (Ishihara and Qian, 2018), wind tunnel experiments (Campagnolo et al., 2022), field measurements (Zhan et al., 2020; Daems et al., 2025), and operational wind farm data, such as supervisory control and data acquisition (SCADA) data (Barthelmie et al., 2010). A key advantage of SCADA data is that it inherently reflects site-specific conditions, including topography, wind resource, surface roughness, atmospheric stability, and turbulence intensity, as well as wind-farm-specific characteristics such as layout, size, and turbine spacing, and turbine-specific properties such as power and thrust curves (van Binsbergen, 2025). Additionally, atmospheric stability is particularly relevant because it affects wake recovery and wake-induced power deficits, with stable conditions generally associated with stronger and more persistent wake losses than unstable conditions (Hansen et al., 2011; Barthelmie et al., 2015; Dörenkämper et al., 2015). Calibration based on SCADA data can therefore account for uncertainties associated with these factors more directly than controlled or idealized datasets.

A common SCADA-based calibration strategy is to bin the data by wind direction and wind speed prior to parameter estimation. This binned approach reduces the influence of sensor noise and stochastic variability in the inflow conditions. Depending on the bin size, it also smooths the wake, which tends to reduce peak wake losses in the aggregated data (Barthelmie et al., 2009, 2010; Gaumond et al., 2013). Binned calibration requires observations within each bin to represent a consistent operating state. Periods of curtailment, faults, downtime, or other abnormal turbine behaviour should therefore preferably not be mixed with normal operation, as this would bias the power production reference used for calibration. However, in large wind farms the probability that at least one turbine operates outside its normal regime increases with the number of turbines. Enforcing normal operation across the full farm can therefore lead to substantial data filtering and, consequently, data scarcity.

An alternative is time-series-based calibration, in which model parameters are estimated for each timestamp in the time series (van Binsbergen et al., 2024). Unlike the binned approach, this method allows the wind farm layout and the calibration cost function to be adapted for each timestamp. For instance, turbines that are non-operational can be removed from the modelled layout and excluded from the cost function. This can substantially increase the usable fraction of the dataset and yield a time series of calibrated parameter values. However, time-series-based calibration does not benefit from the smoothing and reduction in variability obtained by aggregating multiple observations. Individual timestamps may therefore reflect transient or otherwise non-representative conditions that are not fully consistent with the steady-state assumptions of analytical wake models. Although such effects are often averaged out in binned analyses, they remain visible in a time-series-based framework. This can reduce the representativeness of some timestamp-level parameter estimates, but it also enables these cases to be identified and quantified post hoc using quantitative metrics.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

As a result, direct generalisation through simple averaging of timestamp-level parameter estimates may lead to misleading estimates of site-specific parameters. Although this approach retains a larger fraction of SCADA data by adapting the modelled wind farm layout and calibration objective for each timestamp, it remains sensitive to measurement uncertainty and stochastic inflow variability. Therefore, it remains an open question how time-series-based calibration results can be generalised into representative site-specific model parameters without introducing bias due to uneven data exclusion across inflow conditions. This work addresses that question. In addition, given the known influence of atmospheric stability on wake recovery and wake-induced power deficits, this study examines how stability regimes affect the resulting calibrated parameter distributions.

1.1 Objectives and paper structure

This work performs time-series-based calibration of analytical wake-model parameters using SCADA data and investigates how the resulting calibration outputs can be generalised into robust site-specific and cluster-level parameter estimates. Two analytical wake models are considered: the Jensen wake model (Jensen, 1983; Katić et al., 1987) and the Gaussian TurbOPark wake model (Nygaard et al., 2020; J G Pedersen and Nygaard, 2022; Nygaard et al., 2022). The Jensen model is included because it remains widely used in industry due to its simplicity and computational efficiency. The Gaussian TurbOPark model is included as a more recent analytical formulation with a Gaussian wake-deficit profile and a different far-wake recovery formulation. Both models are particularly suitable for the present analysis because their wake-recovery behaviour can be controlled through a single primary tuning parameter. This enables a direct comparison of parameter generalisation and uncertainty quantification across wake-model formulations, while avoiding the additional parameter non-identifiability that can occur in multi-parameter calibration when different parameter combinations produce similarly optimal model predictions.

Particular emphasis is placed on avoiding systematic bias caused by the disproportionate exclusion of data associated with specific inflow conditions. To this end, the calibrated parameters from individual 10-minute timestamps are retained and analysed within a Bayesian framework, rather than being reduced to heavily filtered or bin-averaged parameter estimates. This enables the estimation of farm-specific tuning parameters, a cluster-wide central tuning parameter, and the posterior predictive distribution for the tuning parameter of a wind farm outside the calibration dataset. The same framework also quantifies the expected spread in new timestamp-level estimates of the tuning parameter through posterior prediction. In this sense, the present study extends the work of van Binsbergen et al. (2024).

The study further examines the role of atmospheric stability in wake-model calibration. Two stability proxies are considered: the inverse Monin–Obukhov length, L^{-1}, and the bulk Richardson number, Ri_b. While unstable atmospheric boundary layers are generally associated with faster wake recovery than neutral or stable layers, the magnitude of this dependence remains insufficiently quantified for large offshore wind-farm clusters (Hansen et al., 2011; Peña et al., 2014; Abkar and Porté-Agel, 2015; Platis et al., 2021; Syed et al., 2023). The stability proxies are first analysed in terms of their diurnal and seasonal variability, as well as their dependence on flow origin by distinguishing between wind directions associated with marine and land-influenced inflow. The calibrated wake-model tuning parameters are then analysed as a function of these stability proxies.

The paper is structured as follows. First, Section 2 describes the case-study concession zone and the data sources used in this work. Section 3 then summarises the data filtering procedure and the time-series-based calibration of the Jensen and Gaussian

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

TurbOPark wake models. This is followed by Section 4, where the Bayesian framework used to generalise the calibrated tuning parameter estimates is described and the corresponding results are presented. Section 5 introduces the atmospheric-stability proxies, describes how they are computed, and compares the different data sources used to derive them. Finally, Section 6 examines the effect of atmospheric stability on wake expansion. This section first analyses the continuous relationship between the calibrated wake expansion parameters and the atmospheric-stability proxies, and then evaluates the effect of stability-class-specific parameter values on yield estimates for both wake models.

2 Case study and data sources

SCADA data from four wind farms (WFs) within the Belgian–Dutch (BE–NL) concession zone is used as the primary reference data for this work. The full concession zone comprises 12 densely spaced WFs, as shown in Figure 1. The SCADA variables used are turbine-level 10-minute values of active power, nacelle-measured wind speed, wind direction, and wind-speed standard deviation, with the latter computed over each 10-minute interval from 1 Hz wind-speed measurements. The concession zone is surrounded by three meteorological masts (met masts). The met masts provide the measurement data required to compute the bulk Richardson number, Ri_b, which is used as the first observational proxy for atmospheric stability. In addition, the Obukhov length, L, is computed from ERA5 reanalysis data and subsequently expressed as the inverse Monin–Obukhov length, L^{-1}, which is used as a second atmospheric-stability proxy (Stull, 1988; ERA5; Hersbach, 2020). A two-year measurement period is used for the SCADA data, ERA5 reanalysis data, and met mast data.

3 Calibration

In the following section, the filtering and calibration procedure used as a reference to the Bayesian generalisation framework are described in Subsection 3.1 and 3.2, respectively.

3.1 Filtering procedure

The aim of the filtering procedure is to identify, for each 10-minute timestamp, whether each turbine operates consistently with normal power-curve behaviour. First, sensor faults are addressed by removing timestamps for which the variance of the 10-minute SCADA wind-speed or wind-direction signal is equal to zero, or for which multiple consecutive measurements are identical. A power-curve-based filtering procedure, adopted from Doekemeijer et al. (2019), is then applied, as illustrated in Figure 2. Turbine operation within the manufacturer's power-curve bounds is classified as accepted. Operation outside these bounds is classified as rejected, except for two specific cases: turbines producing less than 1% of the expected power-curve power are classified as inactive, while turbines with nearly constant active power within a 10-minute interval, or over consecutive timestamps, are classified as curtailed.

The resulting operational classification is subsequently used to adjust the calibration setup for each 10-minute timestamp. Turbines classified as accepted are included in both the wake-model simulation and the calibration cost function. Turbines

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

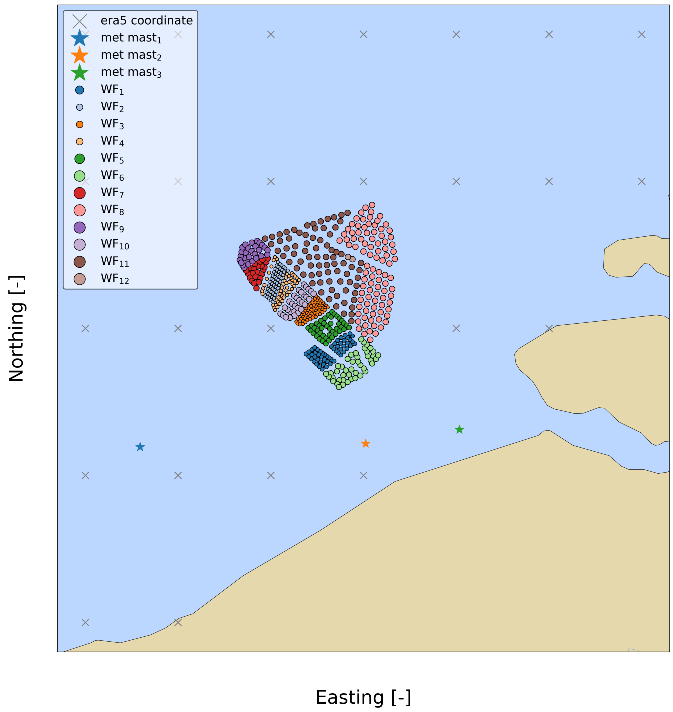


Figure 1. Layout of the BE–NL concession zone, showing WF$_1$–WF$_{12}$ and met mast$_1$–met mast$_3$, together with the ERA5 grid points on a 0.25° latitude–longitude grid. The sea is shown in blue, while land is shown in beige.

classified as inactive are removed from both the simulation and the cost function, as they are assumed not to contribute meaningfully to thrust production.

If one or more turbines are classified as rejected, the corresponding timestamp is excluded from the calibration dataset. For these turbines, the observed power production is inconsistent with normal power-curve behaviour, and the corresponding thrust production is unknown. Their wakes therefore cannot be represented consistently in the wake-model simulation. In the present work, turbines classified as curtailed are treated equivalently to rejected turbines and are therefore also excluded from calibration, since no curtailment-specific thrust correction is applied. Nevertheless, the explicit identification of curtailed operating states may be useful in future work. If a suitable thrust representation is available, curtailed turbines could be retained as wake-generating turbines in the simulation while being excluded from the calibration cost function.

This timestamp-specific treatment is an important difference from conventional binned calibration. In the time-series-based approach, the cost function, wind-farm layout, and turbine power- and thrust-curve representation can be adjusted for every 10-minute window according to the operational state of each turbine. The filtering procedure is therefore applied to each turbine in the wind farm being calibrated before constructing the timestamp-specific cost function and wind-farm layout.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

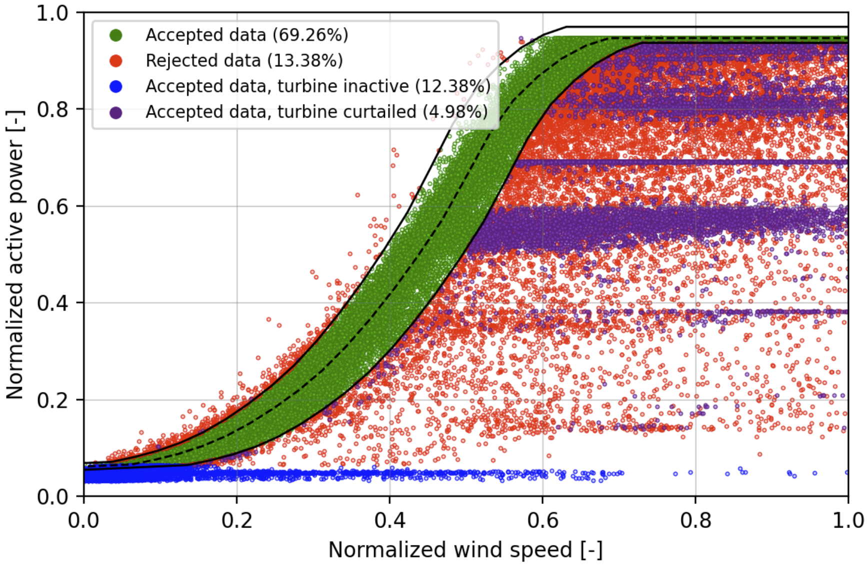


Figure 2. Example of the operational classification for one wind turbine. Green points indicate accepted data. Red and purple points indicate rejected data, for which the corresponding timestamp is excluded from calibration. Blue points indicate inactive operation: the timestamp is retained, but the turbine is removed from the simulation layout and calibration cost function because it produces insufficient power and is assumed not to contribute meaningfully to thrust production.

3.2 Calibration procedure

Calibration was performed using the Optuna framework (Akiba et al., 2019) in a three-stage procedure. In the first stage, the unwaked wind speed U is explored using a quasi-Monte Carlo (QMC) sampler (Bergstra and Bengio, 2012) with wide search bounds. In the second stage, both U and the unwaked wind direction ϕ are explored using the same QMC-based approach, with narrower bounds for the wind speed and wide bounds for the wind direction. In the final stage, U, ϕ, and the set of wake-model tuning parameters $\mathbf{\Omega}$ are jointly calibrated using the tree-structured Parzen estimator (TPE) algorithm (Bergstra et al., 2012). In this final stage, wide bounds are used for the wake-model tuning parameters, while narrower bounds are used for U and ϕ to retain some flexibility in the inflow conditions. The calibration procedure is described in detail in previous studies, and therefore only the essential elements are discussed here. For further details, the reader is referred to van Binsbergen et al. (2024); van Binsbergen (2025). The calibration objective is defined as the mean absolute error (MAE) between the SCADA active power, $\bar{P}$, and the modelled active power, $\hat{P}$, evaluated over the turbines included in the timestamp-specific calibration objective, as given by Equation 1.

$$\mathrm{MAE} = \frac{1}{N_{\mathrm{WT}}} \sum_{i=1}^{N_{\mathrm{WT}}} \left| \bar{P}_i - \hat{P}_i(U, \phi, \mathbf{\Omega}) \right| \tag{1}$$

Here, N_{WT} denotes the number of turbines included in the timestamp-specific calibration objective. The wake-model power estimates are computed using the PyWake framework (Pedersen et al., 2023) as a function of U, ϕ, and $\mathbf{\Omega}$. For the wake

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

models considered in this work, the set of wake-model tuning parameters contains a single scalar parameter, denoted by $\omega \in \mathbf{\Omega}$. Specifically, $\omega = k$ for the Jensen wake model, with $k \in [0, 0.2]$, and $\omega = A$ for the Gaussian TurbOPark model, with $A \in [0, 1]$. In both cases, wake deficits are combined using squared-sum superposition. For the Jensen wake model, rotor averaging is performed using Gaussian quadrature. Using the MAE simplifies the cost-function definition compared with the L_2-norm formulation used in van Binsbergen et al. (2024); Binsbergen et al. (2024); Van Binsbergen et al. (2024), where a weighting factor was required to balance turbine-level and farm-level error contributions. For more details, van Binsbergen (2025) can be consulted.

The reference turbulence intensity (TI) is modelled as a function of wind speed. The TI fit is derived from two years of SCADA data from unwaked turbines. As shown in Figure 3, the fitted TI as a function of the unwaked wind speed agrees well with the reference values proposed in IEC 61400-1 (IEC, 2019), and is therefore used as the reference TI in the wake-model simulations.

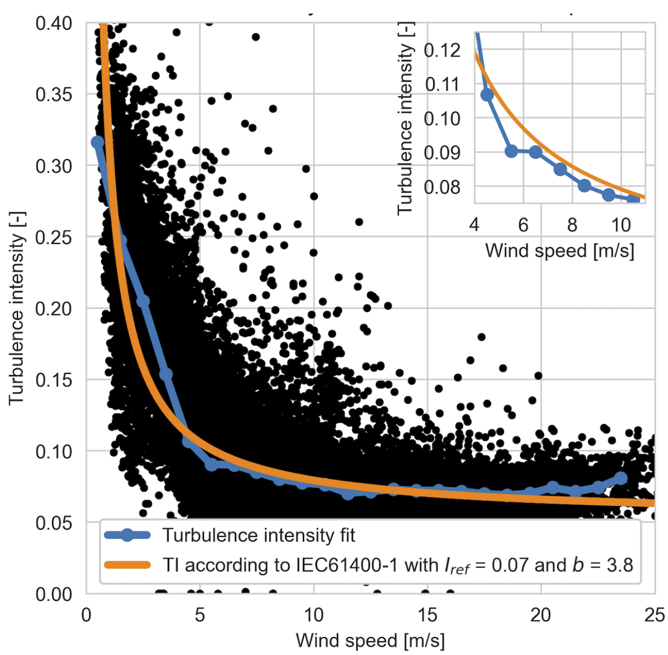


Figure 3. TI as a function of unwaked wind speed. The blue curve shows the fit derived from two years of SCADA data from unwaked turbines and is compared with the IEC 61400-1 (IEC, 2019) formulation using $I_{\text{ref}} = 0.07$ and $b = 3.8$.

4 Generalisation of wake-model tuning parameters

In earlier work, analytical wake models were calibrated using deterministic optimisation, yielding point estimates of model tuning parameters for individual farms, time windows, or operating regimes. Repeating this procedure over many 10-minute SCADA timestamps provides an empirical distribution of calibrated values, from which summary statistics can be derived.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

However, such empirical summaries do not explicitly separate the different processes that contribute to the observed variability. The spread in the calibrated parameters should therefore not be interpreted as a direct measure of wake-expansion variability alone. Instead, it reflects the combined effects of physical variability, measurement uncertainty, and residual model mismatch, where the latter may arise from non-representative flow states or flow physics not captured by the analytical wake model.

In this work, a Bayesian hierarchical framework is introduced to generalise the calibrated wake-model tuning parameters across wind farms while accounting for these uncertainty sources. The interpretation of calibrated parameters as uncertain quantities is related to recent Bayesian uncertainty-quantification work for engineering wind-farm flow models (Aerts et al., 2026), although the present work uses a simpler representation and does not explicitly model inadequacy in the same way. The framework is used to infer representative farm-level and cluster-level tuning parameters, to quantify inter-farm variability, and to construct posterior predictive distributions for additional wind farms assumed to belong to the same offshore wind-farm cluster.

4.1 Conceptual framework

The time-series calibration procedure produces a large number of calibrated wake-model tuning parameters for each wind farm. As discussed above, these timestamp-level values reflect a combination of physical variability, calibration uncertainty, model mismatch, and non-representative 10-minute timestamps, rather than repeated measurements of a single fixed constant.

The aim of the generalisation framework is to use these timestamp-level calibrated values to infer representative wake-model tuning parameters at different aggregation levels. At the wind-farm level, the objective is to estimate a characteristic tuning parameter for each calibrated wind farm. At the cluster level, the objective is to estimate the central tendency and variability of the tuning parameter across the offshore wind-farm cluster. Finally, this cluster-level variability is used to estimate the expected range of representative tuning parameters for an additional wind farm assumed to belong to the same offshore wind-farm cluster, but for which no calibration data are available.

This distinction between aggregation levels is important. Uncertainty in the cluster-wide central parameter describes how well the central behaviour of the observed farms is constrained, whereas uncertainty for a new wind farm also includes the estimated variability between farms. The latter therefore represents predictive uncertainty for a new member of the same offshore-cluster population, rather than only uncertainty in the mean behaviour of the calibrated farms.

4.2 Probabilistic formulation

The conceptual framework described above is implemented as a Bayesian hierarchical model. The calibrated timestamp-level tuning parameters are treated as noisy observations of latent farm-level central parameters. These farm-level parameters are modelled as draws from a common offshore-cluster distribution, whose mean and variability represent the cluster-wide central tendency and inter-farm variability. A posterior predictive distribution for an additional wind farm assumed to belong to the same offshore-cluster population is obtained by drawing a new farm-level parameter from this inferred cluster-level distribution.

The calibration framework is formulated for a general set of wake-model tuning parameters, denoted by $\boldsymbol{\Omega}$. For the wake models considered in this work, this set contains a single scalar tuning parameter, denoted by $\omega \in \boldsymbol{\Omega}$. Specifically, $\omega = k$ for

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

the Jensen wake model (Jensen, 1983; Katić et al., 1987) and $\omega = A$ for the Gaussian TurbOPark model (Nygaard et al., 2020; J G Pedersen and Nygaard, 2022; Nygaard et al., 2022). The probabilistic formulation below is therefore written for the scalar parameter ω.

The Bayesian formulation is introduced in steps. First, the calibration-level observations are defined from the repeated time-series calibration results. The observation model then links these calibrated values to latent farm-level parameters and specifies the uncertainty terms associated with each observation. A residual mismatch extension is introduced to account for systematic model-data mismatch, after which the farm-level parameters are connected through a hierarchical offshore-cluster model. Finally, the prior distributions and posterior predictive formulation are specified.

Figure 4 provides a schematic overview of the complete generalisation framework. The figure is intended as a roadmap for the probabilistic formulation. It shows how the timestamp-level calibration results are first converted into precomputed model inputs, how these inputs enter the hierarchical Bayesian model, and how posterior predictive quantities are derived from the inferred cluster-level distribution. The individual quantities, uncertainty terms, and probability distributions shown in the figure are introduced in the following subsections as the formulation is developed.

4.2.1 Calibration-level observations

The precomputed run-to-run statistics shown in the decoupled-preprocessing block of Figure 4 are derived from repeated calibration runs. For each valid 10-minute calibration instance, the time-series calibration is performed $R = 5$ times. This number of repeated runs is chosen as a compromise between computational cost and the need to obtain a basic diagnostic of run-to-run calibration stability. It is not intended to represent a physical source of variability. The resulting values are indexed by wind farm, calibration instance, and calibration run. Let $\hat{\omega}_{p,i,r}$ denote the calibrated value of the scalar tuning parameter for wind farm p, calibration instance i, and repeated calibration run r, with $p = 1, \dots, P$, $i = 1, \dots, I_p$, and $r = 1, \dots, R$. Here, P is the number of wind farms, I_p is the number of valid 10-minute calibration instances for wind farm p, and $R = 5$.

For each valid 10-minute calibration instance, the repeated calibration results are summarised by their mean,

$$\hat{\omega}_{p,i} = \frac{1}{R}\sum_{r=1}^{R} \hat{\omega}_{p,i,r}, \tag{2}$$

and their standard deviation,

$$s_{p,i} = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R} \left(\hat{\omega}_{p,i,r} - \hat{\omega}_{p,i}\right)^2}. \tag{3}$$

The mean $\hat{\omega}_{p,i}$ is used as the calibrated tuning parameter for calibration instance i at wind farm p. The standard deviation $s_{p,i}$ is interpreted as an empirical, timestamp-specific diagnostic of run-to-run calibration stability under the chosen optimisation setup and parameter space, rather than as a direct estimate of physical wake-parameter variability. Because only $R = 5$ repeated calibration runs are available, $s_{p,i}$ can be sensitive to individual outlying runs and is therefore used only as a calibration-stability diagnostic.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

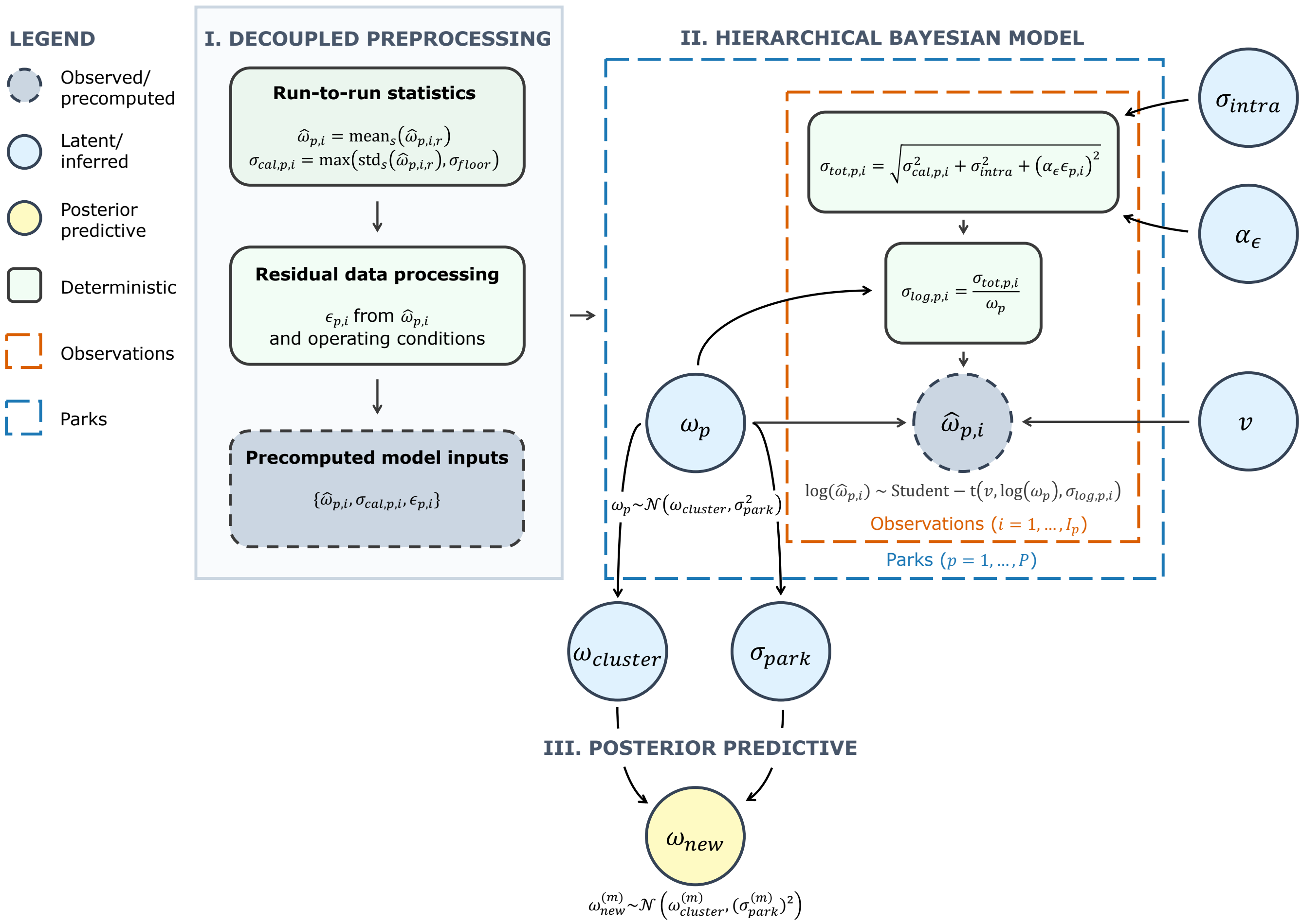


Figure 4. Schematic overview of the Bayesian generalisation framework for wake-model tuning parameters. The framework consists of three parts: decoupled preprocessing of timestamp-level calibration results, the hierarchical Bayesian model used to infer farm-level and cluster-level tuning parameters, and the posterior predictive distribution for an additional wind farm. Dashed nodes indicate observed or precomputed quantities, blue nodes indicate latent or inferred quantities, yellow nodes indicate posterior predictive quantities, and green boxes indicate deterministic transformations. The quantities and distributions shown in the figure are defined in the following subsections.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

To avoid unrealistically small uncertainty estimates in the subsequent Bayesian model, a lower bound is imposed on this calibration-related uncertainty,

$$\sigma_{\mathrm{cal},p,i} = \max\left(s_{p,i}, \sigma_{\mathrm{floor}}\right), \qquad \sigma_{\mathrm{floor}} = 10^{-4}. \tag{4}$$

Here, $\sigma_{\mathrm{cal},p,i}$ represents the calibration-related uncertainty contribution assigned to calibration instance i at wind farm p.

For faster applications, the repeated calibration runs can be omitted and each calibration instance can instead be calibrated once. In that case, no run-to-run stability statistics are computed, $s_{p,i}$ is omitted, and $\hat{\omega}_{p,i}$ is taken as the tuning-parameter value returned by the single calibration run. This alternative removes the run-to-run calibration-stability contribution from the uncertainty model and is not used in the present analysis.

4.2.2 Observation model

The observation model and uncertainty terms shown in the hierarchical Bayesian model block of Figure 4 are defined below.

The mean calibrated value $\hat{\omega}_{p,i}$ is modelled as a noisy observation of a latent farm-level central tuning parameter ω_p. Since the tuning parameter is strictly positive, the observation model is specified on the logarithmic scale. A log-Student-t likelihood is used to allow for heavy-tailed deviations in the calibrated values,

$$\log \hat{\omega}_{p,i} \sim \text{Student-}t\left(\nu, \log \omega_p, \sigma_{\log,p,i}\right). \tag{5}$$

Equivalently, Eq. (5) can be written as

$$\hat{\omega}_{p,i} = \omega_p \exp\left(\eta_{p,i}\right), \qquad \eta_{p,i} \sim \text{Student-}t\left(\nu, 0, \sigma_{\log,p,i}\right). \tag{6}$$

This formulation defines deviations additively on the logarithmic scale and multiplicatively on the original tuning-parameter scale, while retaining a heavy-tailed residual model.

The Student-t likelihood is chosen because individual 10-minute calibration results may deviate strongly from the latent farm-level central parameter. Such deviations can arise from measurement uncertainty, non-representative flow conditions, wake-model mismatch, or calibration artefacts, which are not separately identifiable at the level of an individual calibrated value. The degrees-of-freedom parameter ν controls the heaviness of the tails, while $\log \omega_p$ and $\sigma_{\log,p,i}$ define the location and scale of the distribution on the logarithmic scale. In this formulation, large timestamp-specific deviations are therefore down-weighted probabilistically rather than removed by an explicit outlier filter.

The uncertainty contributions are first defined on the original tuning-parameter scale. In the reference model, which does not include the residual mismatch term, the total original-scale observation uncertainty is defined as

$$\sigma_{\mathrm{tot},p,i} = \sqrt{\sigma^2_{\mathrm{cal},p,i} + \sigma^2_{\mathrm{intra}}}. \tag{7}$$

The first term, $\sigma_{\mathrm{cal},p,i}$, represents the timestamp-specific calibration-related uncertainty derived from the repeated calibration runs, as defined in Equation 4. The second term, σ_{intra}, is an inferred residual scale representing additional intra-farm variability among calibration instances that is not explained by the timestamp-specific calibration uncertainty alone.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

For the log-Student-t observation model, this original-scale uncertainty is converted to an approximate log-scale residual width,

$$\sigma_{\log,p,i} = \frac{\sigma_{\mathrm{tot},p,i}}{\omega_p}. \tag{8}$$

This conversion expresses the residual uncertainty relative to the farm-level central tuning parameter. The observation model therefore preserves the original-scale interpretation of the uncertainty components, while ensuring that posterior predictive tuning-parameter values remain positive.

4.2.3 Residual mismatch model

The residual data-processing step shown in the decoupled preprocessing block of Figure 4 defines a post-calibration mismatch diagnostic, $\epsilon_{p,i}$, for each wind farm and calibration instance. The diagnostic is subsequently used in the residual-mismatch-aware model as an additional observation-scale contribution, assigning larger uncertainty to calibration instances with larger residual power mismatch.

The choice of residual mismatch indicator requires care, because the diagnostic should not systematically favour particular stability regimes or tuning-parameter ranges. Metrics based on correlation or explained variance, such as the concordance correlation coefficient or coefficient of determination, are sensitive to the variance of the turbine-power values within the calibration instance. Atmospheric stability can alter this variance by changing the spread of turbine-power values across the wind farm. A similar issue can occur when the wind farm operates in the wake of an upstream wind farm, in which case turbine powers may become more similar across the farm and the variance of the power signal is reduced. In such cases, correlation- or variance-based diagnostics can change for reasons unrelated to the representativeness of the calibrated wake-model parameter. Using these diagnostics directly as uncertainty weights could therefore unintentionally reweight the posterior towards specific atmospheric regimes or operating conditions.

For this reason, $\epsilon_{p,i}$ is defined as a normalised absolute power-error metric. For each wind farm p and 10-minute calibration instance i, the residual mismatch indicator $\epsilon_{p,i}$ is computed after the repeated calibration runs have been completed. The wake model is evaluated using the mean calibrated tuning parameter over the $R = 5$ calibration runs, together with the corresponding unwaked wind speed and wind direction. The residual mismatch indicator is defined as the turbine-averaged L_1-norm of the post-calibration power mismatch, normalised by the maximum modelled turbine power. This normalisation is used to reduce the dependence of the diagnostic on the overall power level or spread within a calibration instance, which could otherwise vary systematically with atmospheric stability, operating conditions, or upstream-wake effects.

$$\epsilon_{p,i} = \frac{1}{n_{\mathrm{WT},p,i}} \frac{L_{1,p,i}}{P_{\mathrm{max,model},p,i}}, \tag{9}$$

where $L_{1,p,i}$ is the turbine-level L_1-norm of the power mismatch,

$$L_{1,p,i} = \sum_{j=1}^{n_{\mathrm{WT},p,i}} \left| P_{\mathrm{model},p,i,j}\left(\hat{\omega}_{p,i}, U_{p,i}, \phi_{p,i}\right) - P_{\mathrm{SCADA},p,i,j} \right|. \tag{10}$$

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

WIND ENERGY SCIENCE DISCUSSIONS

Here, $n_{\mathrm{WT},p,i}$ is the number of turbines included in the calibration instance, $P_{\mathrm{SCADA},p,i,j}$ is the observed SCADA power of turbine j, and $P_{\mathrm{model},p,i,j}$ is the corresponding modelled power. The mean calibrated tuning parameter $\hat{\omega}_{p,i}$, the mean unwaked wind speed $U_{p,i}$, and the mean unwaked wind direction $\phi_{p,i}$ are determined as the means over the five calibration runs.

The quantity $P_{\mathrm{max,model},p,i}$ denotes the maximum modelled turbine power among the turbines included in the calibration instance. The resulting $\epsilon_{p,i}$ is therefore a non-negative post-calibration diagnostic of the mean normalised absolute power mismatch for each 10-minute calibration instance.

In the residual-mismatch-aware model, this diagnostic is used to define an additional observation-scale contribution. A global coefficient α_ϵ is inferred to convert the residual mismatch indicator into this scale contribution,

$$\sigma_{\epsilon,p,i} = \alpha_\epsilon \epsilon_{p,i}. \tag{11}$$

Here, $\epsilon_{p,i}$ varies across wind farms and calibration instances, while α_ϵ is shared across the dataset and inferred from the data. The posterior distribution of α_ϵ quantifies the uncertainty in how strongly residual mismatch is associated with additional observation-scale variability.

The total observation scale is then extended as

$$\sigma_{\mathrm{tot},p,i} = \sqrt{\sigma^2_{\mathrm{cal},p,i} + \sigma^2_{\mathrm{intra}} + \sigma^2_{\epsilon,p,i}}. \tag{12}$$

For the residual-mismatch-aware model, this extended value of $\sigma_{\mathrm{tot},p,i}$ is used in Equation 8 to obtain the corresponding log-scale residual width $\sigma_{\log,p,i}$.

The ϵ-dependent term is included as a heteroscedastic observation-scale contribution. It is informative only to the extent that variation in $\epsilon_{p,i}$ explains variation in the residual scatter of the calibrated tuning parameters. If $\epsilon_{p,i}$ is approximately constant across calibration instances, the term provides little observation-specific weighting and becomes difficult to distinguish from the common residual scale σ_{intra}. If $\epsilon_{p,i} = 0$, the term has no effect. This term should therefore be interpreted as a modelled residual-mismatch-dependent uncertainty contribution, rather than as a directly measured physical uncertainty source.

4.2.4 Hierarchical model across wind farms

The wind-farm-level hierarchy shown in the hierarchical Bayesian model block of Figure 4 links the latent farm-level tuning parameters through a common offshore-cluster distribution. This hierarchy provides partial pooling across wind farms: information is shared across the cluster, while each wind farm is still allowed to deviate from the common parameter level. The hierarchy is specified as

$$\omega_p = \omega_{\mathrm{global}} + \sigma_{\mathrm{park}} z_p. \tag{13}$$

Here, p indexes the wind farm, ω_p is the farm-level central tuning parameter for wind farm p, and ω_{global} is the population-level central tuning parameter of the offshore-cluster distribution. The parameter σ_{park} represents the inter-farm standard deviation on the original tuning-parameter scale, while z_p is the standardised farm-level deviation.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

The posterior distribution of ω_{global} represents uncertainty in the central value of the offshore-cluster population, not the full spread of farm-level parameters. Individual farm-level parameters also depend on their farm-specific deviations from this central value. Therefore, posterior distributions for ω_p include both uncertainty in the cluster-wide central value and uncertainty in the farm-specific deviation.

4.2.5 Prior specification

Weakly informative priors are assigned to the unknown model parameters. The global tuning parameter is assigned a normal prior centred on an empirical reference value,

$$\omega_{\text{global}} \sim \mathcal{N}\left(\omega_{\text{ref}}, 0.05^2\right). \tag{14}$$

Here, ω_{ref} is defined, for each wake model separately, as the mean of the calibrated tuning parameters after preprocessing and is used only to centre the weakly informative prior for ω_{global}.

The standardised farm-level deviations are assigned standard-normal priors,

$$z_p \sim \mathcal{N}(0,1), \tag{15}$$

so that farm-level deviations are centred around the offshore-cluster mean before observing the data.

The scale parameters are assigned weakly informative half-normal priors,

$$\sigma_{\text{park}} \sim \text{HalfNormal}(0.05), \tag{16}$$

$$\sigma_{\text{intra}} \sim \text{HalfNormal}(0.05), \tag{17}$$

$$\alpha_\epsilon \sim \text{HalfNormal}(0.3). \tag{18}$$

Here, σ_{park} represents the inter-farm standard deviation, σ_{intra} represents residual intra-farm variability, and α_ϵ controls the strength of the residual-mismatch-dependent observation-scale contribution. The parameter α_ϵ is only included in the residual-mismatch-aware model. This allows comparison between a reference model with $\sigma_{\epsilon,p,i} = 0$ and a residual-mismatch-aware model with $\sigma_{\epsilon,p,i} = \alpha_\epsilon \epsilon_{p,i}$, thereby assessing the influence of residual-mismatch-dependent uncertainty on the posterior distributions and observation-scale uncertainty.

For the log-Student-t degrees of freedom, the following prior is used:

$$\nu = 2 + \nu', \qquad \nu' \sim \text{Exponential}(1). \tag{19}$$

The offset ensures that $\nu > 2$, such that the log-Student-t likelihood has finite variance.

Conditioning the model on the calibrated values $\hat{\omega}_{p,i}$ yields posterior distributions for the unknown parameters. The posterior distribution of ω_{global} is used to summarise the cluster-wide central tuning parameter, while the posterior distributions of ω_p describe the farm-specific central tuning parameters. The posterior distributions of σ_{park}, σ_{intra}, and, where included, α_ϵ, describe the uncertainty components governing inter-farm variability, intra-farm residual variability, and residual-mismatch-dependent observation-scale variability, respectively.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

WIND ENERGY SCIENCE DISCUSSIONS

Posterior samples were obtained using Markov chain Monte Carlo (MCMC) (Metropolis et al., 1953; Hastings, 1970) with the No-U-Turn Sampler (NUTS) (Hoffman and Gelman, 2014), as implemented in NumPyro (Phan et al., 2019). The number of warm-up and posterior samples was chosen such that posterior summaries were stable and standard MCMC diagnostics indicated satisfactory convergence.

The uncertainty decomposition should be interpreted as a statistical representation of variability in the calibrated parameters, rather than as a complete physical separation of all uncertainty sources. The components may overlap because the same underlying cause can affect multiple terms. For example, calibration instances with large residual model-data mismatch may also have larger run-to-run calibration variability, causing both σ_{cal} and σ_ϵ to increase. Similarly, flow states that are not well represented by the analytical wake model may contribute both to residual intra-farm scatter and to larger post-calibration power mismatch. Therefore, σ_{cal}, σ_{intra}, and σ_ϵ should be interpreted as structured statistical contributions to the observation scale, rather than as uniquely identifiable physical uncertainty sources.

4.2.6 Posterior prediction for a new wind farm

The posterior predictive quantities shown in the posterior predictive block of Figure 4 are defined here.

The hierarchical model can also be used to estimate a representative tuning parameter for a wind farm that was not included in the calibration dataset. This is relevant when calibrated wake-model parameters are transferred to a planned wind farm, or to another wind farm in the same offshore cluster. This prediction assumes that the new wind farm is exchangeable with the observed wind farms, meaning that it can be regarded as belonging to the same offshore-cluster population. This assumption should be considered when interpreting ω_{new}, since calibrated wake-recovery parameters may depend on wind-farm characteristics such as average turbine spacing, as also observed in previous work (Nygaard et al., 2020, 2022). The prediction for ω_{new} therefore assumes that such layout-dependent effects are sufficiently represented by the inter-farm variability inferred from the observed farms.

The observed wind farms are used to infer two population-level quantities: the central tuning parameter of the offshore cluster, ω_{global}, and the typical variability between wind farms, represented by σ_{park}. A new wind farm is then treated as an additional farm drawn from the same offshore-cluster population, with its central tuning parameter denoted by ω_{new}.

For the observed wind farms, the farm-specific deviations from the cluster-wide value are informed by their calibration data. For a new wind farm, no such calibration data are available. Therefore, a new farm-level deviation is drawn from the same standardised deviation distribution as the observed farm-level deviations,

$$\omega_{\text{new}} = \omega_{\text{global}} + \sigma_{\text{park}} z_{\text{new}}, \qquad z_{\text{new}} \sim \mathcal{N}(0,1). \tag{20}$$

In practice, this posterior predictive distribution is constructed draw-wise from the posterior samples. For each posterior draw m, the corresponding posterior draws $\omega^{(m)}_{\text{global}}$ and $\sigma^{(m)}_{\text{park}}$ are combined with a newly drawn farm-level deviation $z^{(m)}_{\text{new}} \sim \mathcal{N}(0,1)$, giving

$$\omega^{(m)}_{\text{new}} = \omega^{(m)}_{\text{global}} + \sigma^{(m)}_{\text{park}} z^{(m)}_{\text{new}}. \tag{21}$$

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

The collection of values $\{\omega_{\text{new}}^{(m)}\}$ forms the posterior predictive distribution for the representative tuning parameter of a new wind farm.

The resulting distribution of ω_{new} represents the expected range of representative tuning parameters for a wind farm outside the calibration dataset, assuming that it belongs to the same offshore-cluster population. This distribution is generally wider than the posterior distribution of ω_{global}. The reason is that ω_{global} only describes uncertainty in the central value of the offshore-cluster population, whereas ω_{new} also includes the expected variability between wind farms.

In practical terms, ω_{new} can be interpreted as the central tuning parameter of a hypothetical new wind farm. If a new series of calibrated parameters were generated for this wind farm, the individual 10-minute values, denoted by $\hat{\omega}_{\text{new},i}$, would vary around ω_{new}. Since the calibrated tuning-parameter values are positive, these values are described using the same log-Student-t residual structure as in the observation model,

$$\log \hat{\omega}_{\text{new},i} \sim \text{Student-}t\left(\nu, \log \omega_{\text{new}}, \sigma_{\text{log,pred}}\right). \tag{22}$$

Equivalently, this can be written as

$$\hat{\omega}_{\text{new},i} = \omega_{\text{new}} \exp\left(\eta_{\text{new},i}\right), \qquad \eta_{\text{new},i} \sim \text{Student-}t\left(\nu, 0, \sigma_{\text{log,pred}}\right). \tag{23}$$

For the reference model without the residual-mismatch term, the residual scale on the original tuning-parameter scale is given by

$$\sigma_{\text{pred}} = \sigma_{\text{intra}}. \tag{24}$$

This residual scale is converted to the log-scale predictive width as

$$\sigma_{\text{log,pred}} = \frac{\sigma_{\text{pred}}}{\omega_{\text{new}}}. \tag{25}$$

For the reference model, this conversion is evaluated draw-wise from the posterior samples. For each posterior draw m, the predictive residual scale is

$$\sigma_{\text{pred}}^{(m)} = \sigma_{\text{intra}}^{(m)}, \tag{26}$$

and the corresponding log-scale predictive width is

$$\sigma_{\text{log,pred}}^{(m)} = \frac{\sigma_{\text{pred}}^{(m)}}{\omega_{\text{new}}^{(m)}}. \tag{27}$$

The posterior predictive values $\hat{\omega}_{\text{new},i}^{(m)}$ are then generated using the corresponding posterior predictive draw $\omega_{\text{new}}^{(m)}$, the corresponding log-scale predictive width $\sigma_{\text{log,pred}}^{(m)}$, and a draw from the Student-t residual distribution with the corresponding posterior draw $\nu^{(m)}$.

The resulting posterior predictive distribution represents the range of new 10-minute calibrated tuning-parameter values expected for a new wind farm when only the common intra-farm residual variability is considered. It can therefore be used

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

to assess whether the inferred model reproduces the observed spread of the calibrated values $\hat{\omega}_{p,i}$, while ensuring that model-implied tuning-parameter values remain positive on the original scale.

For the residual-mismatch-aware model, the instance-level residual scale can additionally depend on an assumed residual mismatch level for the future calibration instance, ϵ_{new}. The corresponding residual-mismatch-dependent scale contribution is defined as

$$\sigma_{\epsilon,\text{new}} = \alpha_\epsilon \epsilon_{\text{new}}. \tag{28}$$

The total instance-level residual scale on the original tuning-parameter scale is then

$$\sigma_{\text{pred}} = \sqrt{\sigma_{\text{intra}}^2 + \sigma_{\epsilon,\text{new}}^2}. \tag{29}$$

For each posterior draw m, this becomes

$$\sigma_{\text{pred}}^{(m)} = \sqrt{\left(\sigma_{\text{intra}}^{(m)}\right)^2 + \left(\alpha_\epsilon^{(m)} \epsilon_{\text{new}}\right)^2}. \tag{30}$$

The corresponding log-scale predictive width is again obtained using Equation 25, evaluated draw-wise with $\sigma_{\text{pred}}^{(m)}$ and $\omega_{\text{new}}^{(m)}$.

Here, ϵ_{new} represents the residual mismatch level assumed for the future calibration instance, and $\sigma_{\epsilon,\text{new}}$ is the corresponding additional observation-scale contribution. Evaluating the posterior predictive distribution at $\epsilon_{\text{new}} = 0$ gives a low-mismatch prediction, for which $\sigma_{\epsilon,\text{new}} = 0$. Comparing this conditional prediction with the reference posterior predictive distribution indicates how much of the observed spread is attributed to high-ϵ calibration instances rather than to the common intra-farm residual variability.

The run-to-run calibration-uncertainty term $\sigma_{\text{cal},p,i}$ is not included in σ_{pred}, because it is an observation-specific quantity estimated from the repeated calibration runs of the observed 10-minute instances. The posterior predictive distribution for a new wind farm instead uses the inferred residual variability terms, σ_{intra} and, where applicable, $\sigma_{\epsilon,\text{new}}$, to describe the expected spread of new 10-minute calibrated tuning-parameter values.

4.3 Generalised tuning-parameter results

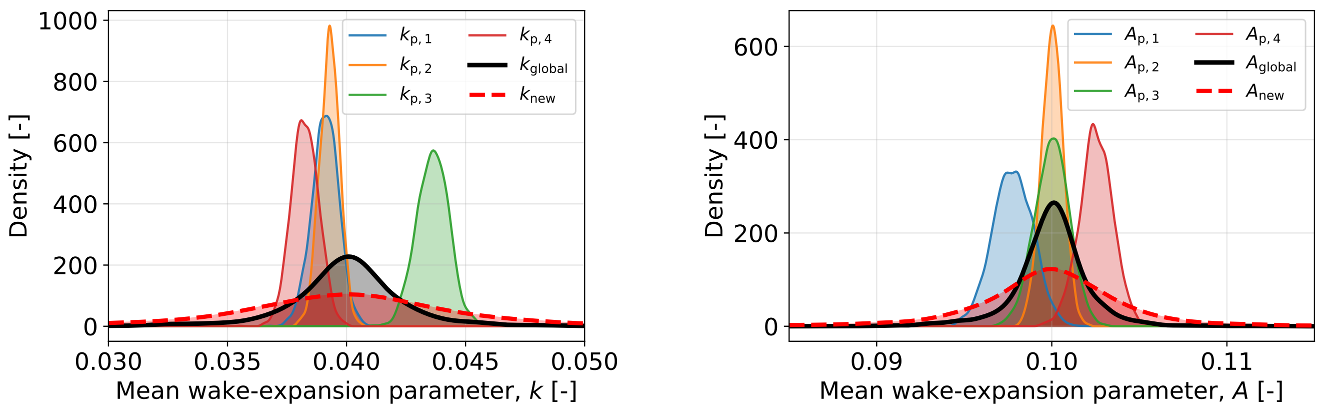
In this section, the Bayesian hierarchical model is evaluated without explicitly conditioning the tuning parameter on atmospheric stability. This provides a baseline estimate of the central wake-model tuning parameter across the offshore cluster and quantifies the expected variability when transferring the calibrated tuning parameter to a wind farm not included in the calibration dataset.

When comparing the two wake models, it should be noted that the calibration datasets used for the Jensen and Gaussian TurbOPark analyses are not identical. Therefore, cross-model differences in the generalised tuning-parameter results cannot be interpreted as resulting solely from the wake-model formulation, but may also partly reflect differences in the retained calibration data. Using identical calibration data would more directly isolate the effect of the wake-model formulation.

Figure 5 shows the posterior distributions of the farm-level central tuning parameters ω_p, the cluster-wide central parameter ω_{global}, and the posterior predictive distribution of the central tuning parameter for a new wind farm, ω_{new}, for the Jensen and

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

Gaussian TurbOPark wake models. The farm-level distributions represent the inferred representative tuning parameter for each observed wind farm after partial pooling. The cluster-wide distribution describes the common central value inferred from the observed wind farms, whereas the new-farm distribution represents the expected range of central tuning parameters for a wind farm drawn from the same offshore-cluster population but not included in the calibration dataset.



(a) Jensen wake model with tuning parameter k.

(b) Gaussian TurbOPark model with tuning parameter A.

Figure 5. Posterior and posterior predictive distributions of the latent central tuning parameters for the Jensen wake model (a) and the Gaussian TurbOPark model (b). The coloured distributions show the farm-level central parameters, ω_p, with $p = 1, \ldots, P$ indexing the observed wind farms. The black distribution shows the cluster-wide central parameter, ω_{global}, while the red dashed distribution shows the posterior predictive distribution of the central tuning parameter for a wind farm not included in the calibration dataset, ω_{new}.

The distribution of ω_{new} is generally wider than that of ω_{global}, because it includes both uncertainty in the cluster-wide central parameter and the inferred inter-farm variability. This distinction is important for model transferability. The posterior distribution of ω_{global} describes how well the central parameter of the observed offshore cluster is constrained, while the posterior predictive distribution of ω_{new} describes the uncertainty in the representative central tuning parameter expected when applying the generalised calibration to a new wind farm.

To interpret the inferred uncertainty structure, Table 1 reports posterior summaries of the main scale and tail parameters. Each entry gives the posterior median and the 5–95% posterior interval. The parameter σ_{park} describes the inferred inter-farm variability, σ_{intra} describes the residual intra-farm variability, and α_ϵ, when included, describes the strength of the residual-mismatch-dependent observation-scale contribution.

Comparing the reference and residual-mismatch-aware models indicates how the residual scatter is distributed between the common intra-farm residual scale and the residual-mismatch-dependent contribution.

The posterior summaries of σ_{park} indicate that the inferred inter-farm variability is small compared to the common intra-farm residual scale. For the Jensen model, the median inter-farm standard deviation is approximately 0.003 in the reference model and 0.004 in the residual-mismatch-aware model. For Gaussian TurbOPark, the corresponding values are approximately 0.006 and 0.003. The inferred inter-farm variability is therefore small in absolute terms, although the Gaussian TurbOPark

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Table 1. Posterior summaries of the main scale and tail parameters in the generalised tuning-parameter model. Values are reported as posterior median with the 5–95% posterior interval in brackets.

Wake model	Model variant	σ_{park}	σ_{intra}	α_{ϵ}	ν
Jensen	Reference	0.003 [0.001, 0.010]	0.025 [0.024, 0.025]	–	8.6 [7.5, 10.2]
Jensen	Residual-mismatch-aware	0.004 [0.002, 0.012]	0.016 [0.016, 0.016]	0.13 [0.13, 0.14]	9.9 [8.6, 11.5]
Gaussian TurbOPark	Reference	0.006 [0.003, 0.017]	0.035 [0.034, 0.036]	–	4.1 [3.8, 4.4]
Gaussian TurbOPark	Residual-mismatch-aware	0.003 [0.001, 0.009]	0.018 [0.016, 0.019]	0.28 [0.27, 0.29]	8.5 [7.4, 9.8]

reference model shows a somewhat larger inter-farm spread than the other cases. This suggests that most of the variability in the calibrated tuning parameters is not explained by systematic differences between farms, but by residual variability between calibration instances.

The posterior summaries of σ_{intra} show that including the residual mismatch term reduces the common intra-farm residual scale for both wake models. For Jensen, σ_{intra} decreases from approximately 0.025 to 0.016, while for Gaussian TurbOPark it decreases from approximately 0.035 to 0.018. This indicates that part of the residual scatter in the reference model is reassigned to the residual-mismatch-dependent observation-scale contribution. Because σ_{intra} is expressed on the calibrated-parameter scale, its absolute value is not directly comparable between Jensen k and Gaussian TurbOPark A.

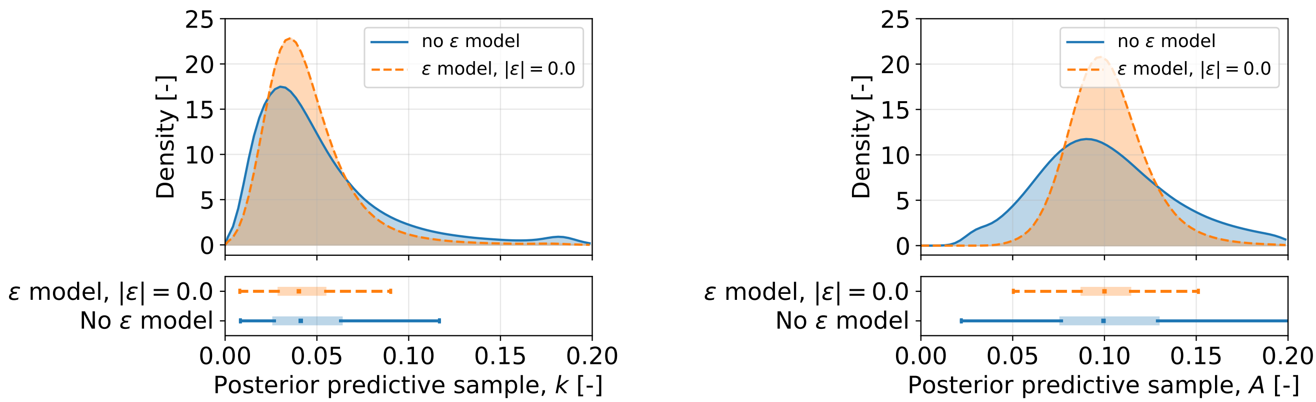
The posterior summaries of α_{ϵ} indicate that the residual mismatch indicator contributes to the observation-scale variability for both wake models. The inferred median values are approximately 0.13 for the Jensen wake model and 0.28 for the Gaussian TurbOPark model. This suggests that residual mismatch is associated with additional spread in the calibrated tuning parameters for both wake models. Because k and A are defined on different tuning-parameter scales, the absolute values of α_{ϵ} should not be interpreted as directly comparable between the two wake models. Together with the reduction in σ_{intra}, this indicates that the residual-mismatch-aware model separates part of the residual scatter into an ϵ-dependent observation-scale term.

The inferred degrees-of-freedom parameter ν indicates that the log-Student-t likelihood retains heavy-tailed residual behaviour for both wake models. For the Jensen model, ν has median values of approximately 8.6 in the reference model and 9.9 in the residual-mismatch-aware model. This indicates moderately heavy-tailed residual behaviour, but less extreme than for the Gaussian TurbOPark reference model. For Gaussian TurbOPark, ν increases from approximately 4.1 in the reference model to 8.5 in the residual-mismatch-aware model. This stronger increase suggests that, for Gaussian TurbOPark, a larger part of the heavy-tailed residual behaviour is accounted for by the ϵ-dependent observation-scale contribution. Overall, the increase in ν when including the residual mismatch term indicates that part of the tail behaviour in the reference model is explained by residual-mismatch-dependent uncertainty, although the posterior still supports heavy-tailed residuals.

Figure 6 shows the posterior predictive distributions of future calibrated tuning-parameter values for a new wind farm, $\hat{\omega}_{\mathrm{new},i}$. In contrast to ω_{new}, this quantity additionally includes residual intra-farm variability and therefore represents the range of values that could be obtained from future 10-minute calibration instances. Consistent with the observation model, these predictive distributions are generated using a log-Student-t residual structure, so that predicted tuning-parameter values

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

remain positive while retaining heavy-tailed residual behaviour. The figures compare the reference model without the residual mismatch term to the residual-mismatch-aware model evaluated at $\epsilon_{\text{new}} = 0$.



(a) Jensen wake model, $\hat{k}_{\text{new},i}$.

(b) Gaussian TurbOPark model, $\hat{A}_{\text{new},i}$.

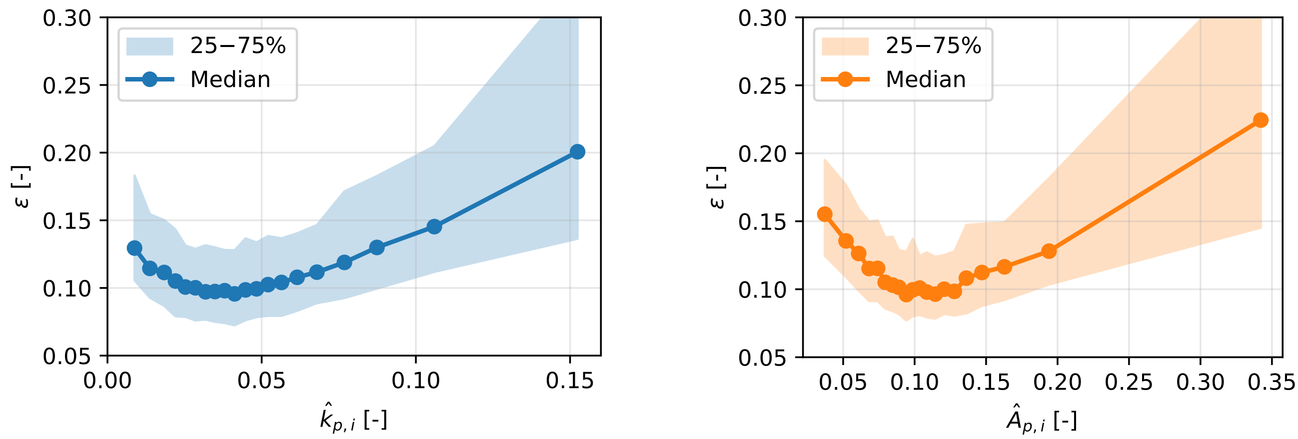
Figure 6. Posterior predictive distributions of future calibrated tuning-parameter values for a wind farm not included in the calibration dataset, for the Jensen wake model (a) and the Gaussian TurbOPark model (b). Blue indicates the reference model without the residual mismatch term, while orange indicates the residual-mismatch-aware model evaluated at $\epsilon_{\text{new}} = 0$. The upper and lower pannels show the smoothed predictive density and corresponding boxplot, respectively.

The reduction in posterior predictive width for the residual-mismatch-aware model is consistent with the posterior summaries in Table 1. When the ϵ-dependent term is included, part of the residual scatter is attributed to residual-mismatch-dependent observation-scale variability through $\alpha_\epsilon \epsilon_{p,i}$, resulting in a reduced common intra-farm residual scale σ_{intra}. When the posterior predictive distribution is evaluated at $\epsilon_{\text{new}} = 0$, this additional residual-mismatch-dependent contribution is absent, and the predictive spread is governed by the reduced σ_{intra}.

This narrower distribution should not be interpreted as a generally more certain wake-model parameter. Instead, it represents a conditional low-mismatch calibration scenario in which no additional residual-mismatch contribution is included in the posterior predictive distribution. The role of $\epsilon_{p,i}$ is to identify calibration instances with larger post-calibration power mismatch, which are assigned larger observation uncertainty in the residual-mismatch-aware model. For future calibration instances with non-zero residual mismatch, the predicted spread increases through the additional $\alpha_\epsilon \epsilon_{\text{new}}$ contribution.

To support the interpretation that $\epsilon_{p,i}$ provides observation-specific information, the relationship between the residual mismatch indicator and the calibrated tuning parameter is analysed. Figure 7 shows $\epsilon_{p,i}$ as a function of the calibrated tuning parameters $\hat{k}_{p,i}$ and $\hat{A}_{p,i}$. The residual mismatch is smallest near the central range of both calibrated parameters and increases toward the lower and upper tails. This suggests that calibration instances yielding parameter values in the tails of the distribution are associated with poorer agreement between the calibrated wake model and the SCADA observations.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026



(a) Jensen wake model.

(b) Gaussian TurbOPark model.

Figure 7. Residual mismatch indicator $\epsilon_{p,i}$ as a function of the calibrated tuning parameter, shown for the Jensen wake model (a) and the Gaussian TurbOPark model (b). The binned median trend shows that residual mismatch is smallest near the central range of calibrated parameter values and increases toward the parameter tails. This suggests that extreme calibrated tuning parameters are more often associated with calibration instances where the wake model retains larger residual power mismatch after calibration.

As a complementary diagnostic, calibration instances are grouped into low-, medium-, and high-ϵ subsets. For each subset, the relative deviation of the calibrated tuning parameter from the posterior mean farm-level central parameter is evaluated as $|\hat{\omega}_{p,i} - \bar{\omega}_p|\bar{\omega}_p^{-1}$, where $\bar{\omega}_p$ denotes the posterior mean of the inferred farm-level central parameter ω_p.

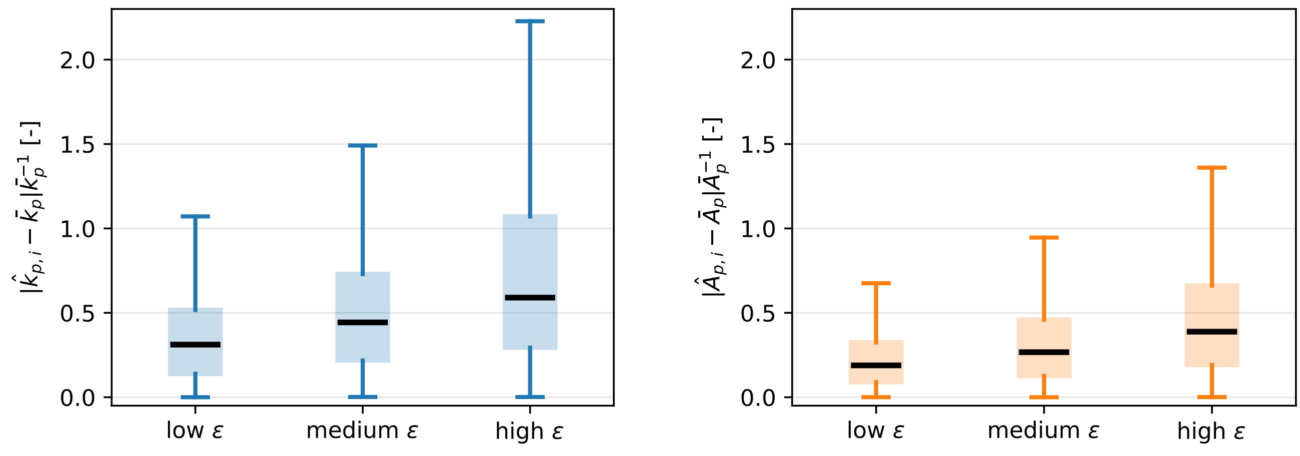
Figure 8 shows that the median and spread of this relative deviation increase with ϵ. This supports the interpretation of $\epsilon_{p,i}$ as a heteroscedastic observation-scale diagnostic, where calibration instances with larger residual mismatch are assigned larger uncertainty in the residual-mismatch-aware model.

Together, these diagnostics indicate that the residual mismatch indicator provides observation-specific information. The ϵ-dependent term does not simply act as an approximately constant additional scale contribution. Instead, it assigns larger uncertainty to calibration instances that are empirically associated with larger deviations from the inferred farm-level central parameter. This supports the use of $\epsilon_{p,i}$ as a residual-mismatch-dependent uncertainty contribution in the posterior predictive model.

The results above describe the generalised tuning parameters while accounting for inter-farm variability and residual uncertainty, but without explicitly conditioning the parameters on atmospheric stability. The next section investigates whether atmospheric stability explains systematic variation in the calibrated tuning parameters beyond these generalised estimates by modelling the central tuning parameter as a function of the stability proxies Ri_b and L^{-1}.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

Open Access
WIND ENERGY SCIENCE
DISCUSSIONS



(a) Jensen wake model.

(b) Gaussian TurbOPark model.

Figure 8. Relative deviation of the calibrated tuning parameter from the posterior mean farm-level central parameter for low-, medium-, and high-ϵ calibration instances. The groups are defined using tertiles of $\epsilon_{p,i}$. The increase in relative deviation with increasing ϵ indicates that high-mismatch calibration instances are associated with larger deviations from the latent farm-level parameter.

5 Atmospheric stability

The generalised tuning-parameter results in the previous section describe the central wake-model parameters without explicitly conditioning them on atmospheric stability. However, atmospheric stability is expected to influence wake recovery through its effect on turbulence, vertical mixing, and momentum exchange in the atmospheric boundary layer (Porté-Agel et al., 2019; Meyers et al., 2022). Before relating the calibrated tuning parameters to stability, the stability conditions present in the case-study area must therefore be characterised.

This section introduces the two atmospheric stability proxies used in this work. The bulk Richardson number, Ri_b, is calculated from meteorological mast measurements and provides an observation-based estimate of near-surface stability. The inverse Monin–Obukhov length, L^{-1}, is derived from ERA5 reanalysis data and provides a complementary stability proxy available over the full offshore-cluster domain. First, the two proxies are compared directly to assess their consistency and to identify systematic differences between the mast-based and reanalysis-based stability estimates. The site-specific distributions of the stability proxies are then analysed, followed by their seasonal and diurnal variability and their dependence on wind speed and direction.

The results of this section provide the atmospheric-stability reference used in Section 6, where the calibrated wake-model tuning parameters are analysed as functions of Ri_b and L^{-1}.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

5.1 Measuring site-specific atmospheric stability

For each meteorological mast, the bulk Richardson number, Ri_b, is computed as

$$Ri_b = \frac{g\,z\,(T_\mathrm{a} - T_\mathrm{s})}{(273.15 + T_\mathrm{a})\,u^2}, \tag{31}$$

under the assumption that the virtual potential temperature is close to the air temperature and that only the horizontal component of wind speed is observed (Albornoz et al., 2022). Here, g is the gravitational acceleration $[\mathrm{m\,s^{-2}}]$, z is the reference height above the sea surface $[\mathrm{m}]$, T_a is the air temperature $[^\circ\mathrm{C}]$, T_s is the sea surface temperature $[^\circ\mathrm{C}]$, and u is the horizontal wind speed $[\mathrm{m\,s^{-1}}]$, as described in Bahamonde and Litrán (2019) and Albornoz et al. (2022). Values of $Ri_b \approx 0$ indicate a near-neutral or shear-dominated boundary layer. Positive values, $Ri_b > 0$, indicate stable stratification, while negative values, $Ri_b < 0$, indicate unstable stratification. The magnitude of Ri_b describes the relative importance of buoyancy effects compared with mechanical shear.

In addition, the inverse Monin–Obukhov length is calculated from ERA5 reanalysis data (Stull, 1988; ERA5; Hersbach, 2020) at the meteorological-mast coordinates using bilinear interpolation as

$$L^{-1} = \frac{\kappa\,g\,T_{v*}}{T_v\,u_*^2}, \tag{32}$$

where κ is the von Kármán constant, taken as 0.4, T_{v*} is the turbulent virtual-temperature scale, T_v is the virtual temperature, and u_* is the friction velocity. For interpretation, $L^{-1} \approx 0$ indicates near-neutral atmospheric stability, $L^{-1} > 0$ indicates stable stratification, and $L^{-1} < 0$ indicates unstable stratification. The magnitude of L^{-1} describes the strength of atmospheric stability per unit length.

5.2 Atmospheric stability classification

In the literature, a range of stability-classification schemes is used to quantify atmospheric stability. For example, Cantero et al. (2022) define seven stability classes for the bulk Richardson number and nine stability classes for the Monin–Obukhov stability parameter, adopted from Sorbjan and Grachev (2010). Additional studies commonly adopt five (Van Wijk et al., 1990; Porchetta et al., 2019) or seven (A. Sathe, 2010) classes for the inverse Monin–Obukhov length. Because the resulting statistics can be sensitive to the number and placement of class boundaries, conclusions may depend strongly on the chosen classification scheme. Therefore, this work presents the relationship between atmospheric stability and the calibrated tuning parameters both as continuous trends and after grouping the stability proxies into predefined stability classes.

For L^{-1}, classification is performed according to Porchetta et al. (2019), as presented in Table 2, while for Ri_b, classification is performed according to Cantero et al. (2022), as presented in Table 3. In Section 5.3, L^{-1} and Ri_b are first compared with each other, followed by an analysis of their temporal variability and their dependence on wind speed and wind direction.

5.3 Comparing atmospheric stability metrics

For both Ri_b and L^{-1}, the sign of the stability proxy determines the stability regime, while its magnitude reflects the strength of the stratification, with values near zero corresponding to near-neutral conditions. Agreement in sign between Ri_b and

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

Table 2. Atmospheric stability classes expressed in terms of inverse Monin–Obukhov length, L^{-1}, following Porchetta et al. (2019).

Stability class	Range of L^{-1} [m^{-1}]
Very unstable	$L^{-1} < -0.005$
Unstable	$-0.005 \leq L^{-1} < -0.001$
Near neutral	$-0.001 \leq L^{-1} < 0.001$
Stable	$0.001 \leq L^{-1} < 0.005$
Very stable	$L^{-1} \geq 0.005$

Table 3. Atmospheric stability classes expressed in terms of bulk Richardson number, Ri_b, following Cantero et al. (2022).

Stability class	Range of Ri_b [–]
Very unstable	$Ri_b < -0.023$
Unstable	$-0.023 \leq Ri_b < -0.011$
Weakly unstable	$-0.011 \leq Ri_b < -0.0036$
Near neutral	$-0.0036 \leq Ri_b < 0.0072$
Weakly stable	$0.0072 \leq Ri_b < 0.042$
Stable	$0.042 \leq Ri_b < 0.084$
Very stable	$Ri_b \geq 0.084$

L^{-1}, where stable and unstable regimes are consistently identified by both proxies, therefore increases confidence in their interpretation. Under near-neutral conditions, an approximately linear relationship between the two stability metrics may be expected for near-surface offshore measurements (Grachev and Fairall, 1997).

For met mast 1, a sign agreement of 81.8% is observed between Ri_b and L^{-1}, indicating strong consistency in the classification of stable and unstable conditions by the two stability proxies. Figure 9 shows the probability density functions (PDFs) of Ri_b and L^{-1}, shown in blue and orange, respectively. Both distributions are strongly concentrated around zero, indicating that near-neutral conditions dominate the dataset.

This is further supported by the cumulative distribution functions (CDFs) shown in Figure 10. The CDFs show that 78% of the Ri_b values and 70% of the L^{-1} values are below zero, indicating that unstable conditions occur more frequently than stable conditions. This predominance of unstable conditions is consistent with Sathe et al. (2011).

Finally, Figure 11 shows the joint distribution of Ri_b from met mast 1 and L^{-1} from ERA5 using probability-density contours from 10% to 100% in 10% increments. The overlaid least-squares linear fit highlights the central tendency of the relationship between the two stability proxies.

In addition to the direct comparison between Ri_b and L^{-1}, both stability proxies are grouped into the stability classes defined in Subsection 5.2. The resulting stability statistics are analysed as a function of hour of day, month of the year, wind speed, and wind direction. The diurnal variability is shown in Figures 12 and 13, the seasonal variability in Figures 14 and 15, the dependence on wind speed in Figures 16 and 17, and the dependence on wind direction in Figures 18 and 19.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

WIND ENERGY SCIENCE DISCUSSIONS
Open Access

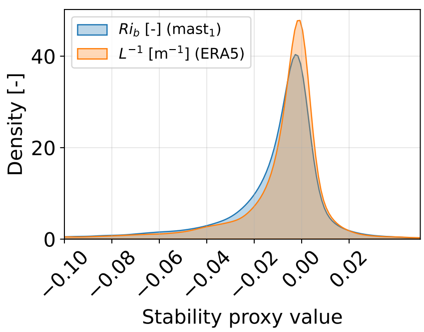


Figure 9. PDF of Ri_b and L^{-1} in blue and orange, respectively.

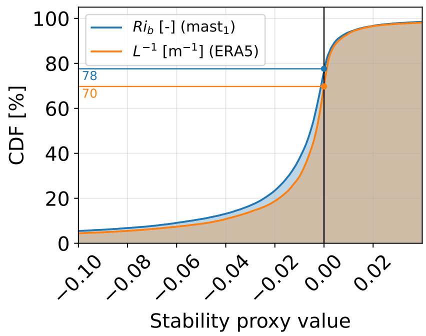


Figure 10. CDF of Ri_b and L^{-1} in blue and orange, respectively.

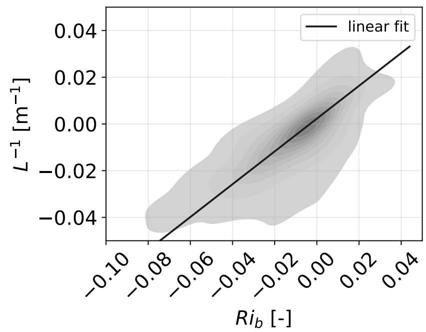


Figure 11. 2D kernel-density probability-density contours from 10% to 100% in 10% increments.

Figures 12 and 13 show that the diurnal cycle is weak for Ri_b, whereas a more pronounced diurnal cycle is present in L^{-1}. This difference is relevant because diurnal variability is generally weaker offshore than over land. Over land, daytime solar heating rapidly warms the surface, producing positive buoyancy and a convective boundary layer with enhanced turbulent mixing. Over water, the larger heat capacity and vertical mixing of the ocean reduce the surface-temperature response to radiative forcing, thereby limiting the diurnal temperature cycle (Stull, 1988).

The weak diurnal signal in Ri_b is therefore consistent with its calculation from local near-surface gradients between the met mast measurements and the sea surface. These gradients are directly constrained by the relatively weak offshore air–sea temperature cycle. In contrast, L^{-1} is derived from ERA5 as a model-diagnosed surface-layer stability parameter and depends on turbulent surface fluxes. In coastal regions, ERA5 may also reflect the influence of nearby land through horizontal advection and grid-scale land–sea contrasts. This can amplify the diurnal variability in L^{-1} relative to the offshore measurement-based Ri_b.

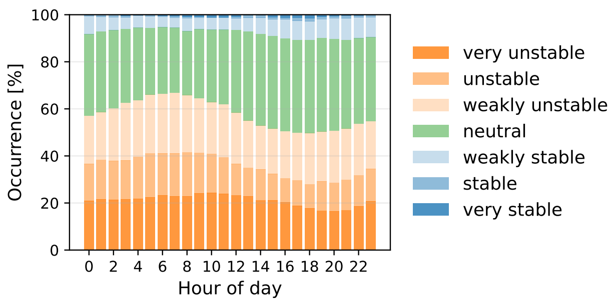


Figure 12. Hourly stability occurrence for Ri_b.

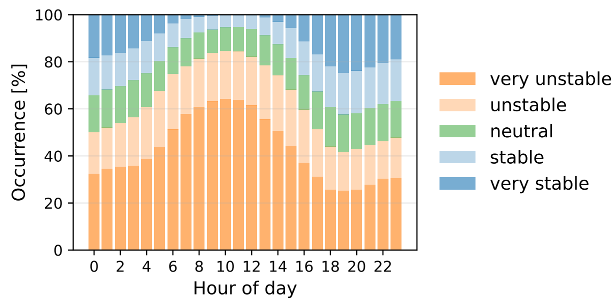


Figure 13. Hourly stability occurrence for L^{-1}.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

Offshore atmospheric stability often exhibits seasonal variability, primarily driven by the annual cycle in sea surface temperature and air temperature (Holtslag, 2016). These changes affect the air–sea temperature difference and thereby the stability of the marine boundary layer. Figures 14 and 15 show that the stability proxies vary over the year; however, the observed monthly variability does not follow a clear seasonal pattern consistent with the expected annual cycle. Since these figures are based on two years of data, the observed variability should therefore not be interpreted as a robust climatological seasonal signal. A multi-year dataset would be required to determine whether a consistent seasonal trend is present at the site.

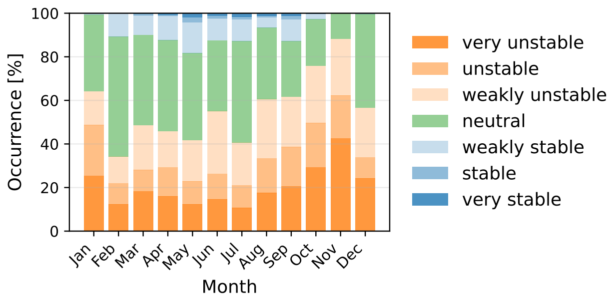


Figure 14. Monthly stability occurrence for Ri_b.

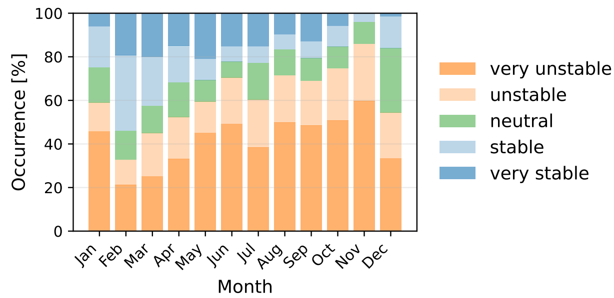


Figure 15. Monthly stability occurrence for L^{-1}.

When the stability proxies are analysed as a function of wind speed, the distribution of stability values is expected to shift towards near-neutral conditions with increasing wind speed (Rodrigo et al., 2015). This reflects the increasing role of mechanically generated turbulence, which enhances vertical mixing and weakens thermal stratification. Consequently, the magnitude of both Ri_b and L^{-1} is expected to decrease with increasing wind speed, consistent with their dependence on wind speed and friction velocity, respectively. This behaviour is observed in Figures 16 and 17, and is particularly evident for Ri_b.

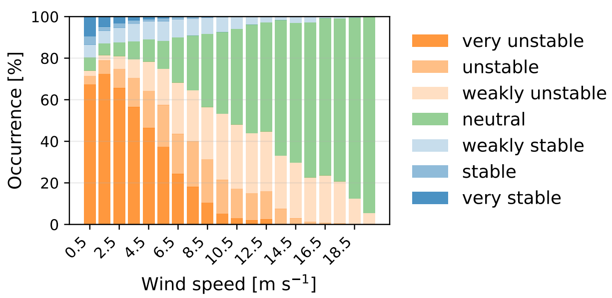


Figure 16. Ri_b against wind speed.

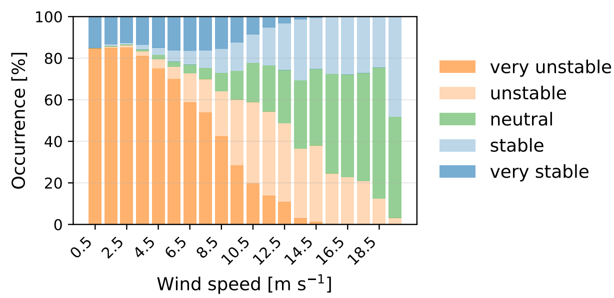


Figure 17. L^{-1} against wind speed.

Figures 18 and 19 show the stability classification for Ri_b and L^{-1}, respectively, as a function of wind direction using 10° bins. These results should be interpreted together with the wind climate, shown by the wind rose in Figure 20, and the land–sea sector classification in Figure 1. The wind rose shows that the prevailing wind direction is in the southwest sector (SW–WSW),

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

particularly at higher wind speeds. According to Figure 1, this sector corresponds to wind originating from sea, whereas wind directions from northeast to south-southwest (NE–SSW) are associated with flow from land.

Based on the wind-speed dependence of stability and the expected influence of coastal land–sea contrasts, two patterns are expected. First, near-neutral stability conditions are expected to occur more frequently for winds from the dominant southwest sea sector, because this sector is associated with higher wind speeds. Second, stable and unstable conditions are expected to occur more frequently for winds from land, where stronger surface-temperature contrasts and diurnal variability can enhance thermal stratification.

These expectations are largely supported by the results. In Figure 18, near-neutral conditions are most prominent within the dominant southwest sea sector, while stable conditions occur more frequently for wind directions associated with land. A similar, but less pronounced, directional pattern is observed for L^{-1} in Figure 19. The weaker directional dependence of L^{-1} may partly reflect differences in the stability-classification thresholds, but also the different data sources and formulations of the two stability proxies. While L^{-1} captures diurnal variability more strongly, Ri_b appears to show a clearer dependence on wind direction. This is consistent with previous studies showing that boundary-layer stability proxies can differ in their sensitivity to wind shear, surface heating, and temperature-driven stratification (Albornoz et al., 2022; Kosović et al., 2026).

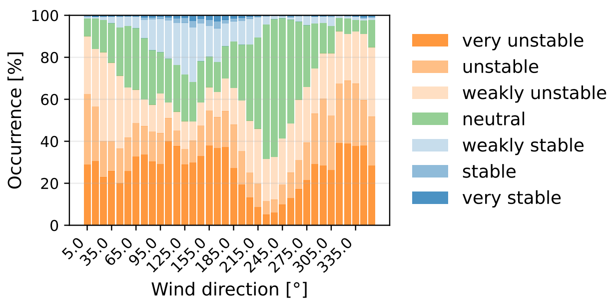


Figure 18. Ri_b against wind direction.

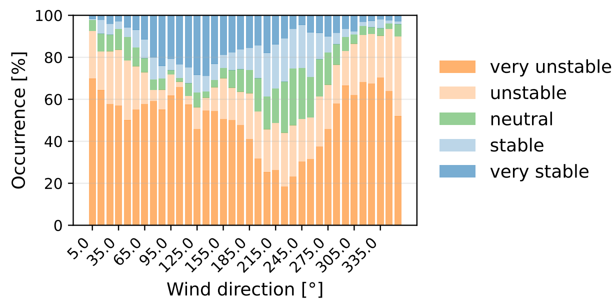


Figure 19. L^{-1} against wind direction.

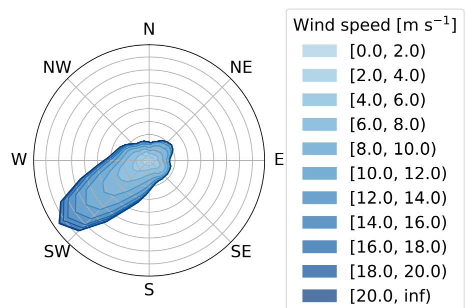


Figure 20. Wind rose measured from met mast 1.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

6 Effect of atmospheric stability on wake-model tuning parameters

The previous sections established generalised wake-model tuning-parameter estimates without explicitly conditioning the parameters on atmospheric stability, and characterised the stability conditions present in the case-study area. This section combines both parts by investigating whether atmospheric stability explains systematic variation in the calibrated wake-model tuning parameters.

The analysis is carried out for two stability proxies: the bulk Richardson number, Ri_b, derived from meteorological mast data, and the inverse Monin–Obukhov length, L^{-1}, derived from ERA5 reanalysis data. For both the Jensen wake model and the Gaussian TurbOPark model, the stability-dependent central tuning parameter is inferred and compared across the observed range of stability conditions.

The results are first presented as smooth stability-dependent trends. The calibrated tuning parameters are then grouped into stability regimes to assess the effect of using stability-specific parameter values in wake-loss and yield calculations.

6.1 Stability-dependent extension of the generalisation framework

To model the effect of atmospheric stability, the generalised Bayesian framework from Section 4 is extended by replacing the single central tuning parameter with a stability-dependent central parameter. The scalar tuning parameter is denoted by ω, with $\omega = k$ for the Jensen wake model and $\omega = A$ for the Gaussian TurbOPark model. For a stability proxy s, with $s = Ri_b$ or $s = L^{-1}$, the shared stability-dependent central parameter is denoted by $\omega(s)$.

Each stability proxy is discretised into $N_b = 25$ quantile-based bins, so that each bin contains approximately the same number of calibration instances. As a result, the outer bins cover wider stability ranges where observations are less frequent. For each 10-minute calibration instance i at wind farm p, the corresponding stability value $s_{p,i}$ is assigned to a bin $b(p,i)$.

The calibrated tuning parameter $\hat{\omega}_{p,i}$ is modelled using the same log-Student-t observation model and uncertainty decomposition as defined in Subsection 4.2.2. The only change is that the generalised farm-level central parameter ω_p is replaced by a wind-farm- and stability-bin-specific central parameter, $\omega_{p,b(p,i)}$. Therefore, the observation model becomes

$$\log \hat{\omega}_{p,i} \sim \text{Student-}t\left(\nu,\, \log \omega_{p,b(p,i)},\, \sigma_{\log,p,i}\right), \tag{33}$$

where $\sigma_{\log,p,i}$ is obtained by converting the same original-scale uncertainty contributions used in the generalised model to the log scale using the local central parameter $\omega_{p,b(p,i)}$.

The stability-dependent model should not be interpreted as estimating unrelated parameters for independent stability bins. Adjacent stability bins are expected to have related values, because small changes in atmospheric stability should not generally produce discontinuous changes in the wake-model tuning parameter. To represent this smoothness assumption, the shared stability-dependent trend is modelled using a latent effect that is smoothed across adjacent bins. The shared central parameter in bin b is written as

$$\omega_b = \omega_{\text{ref,stab}} + g_b, \tag{34}$$

where $\omega_{\mathrm{ref,stab}}$ is the reference level of the stability-dependent curve and g_b is the shared stability-dependent deviation for bin b. This additive formulation keeps the stability-dependent deviations on the original tuning-parameter scale. Since the additive formulation does not by itself guarantee positivity, the resulting central tuning parameters are constrained to remain positive in the implementation.

A first-order random-walk prior is used for the latent stability effect. Changes between neighbouring stability bins are modelled as Gaussian increments,

$$g_b - g_{b-1} \sim \mathcal{N}\left(0, \sigma_{\mathrm{stab}}^2\right), \qquad b = 2, \ldots, N_b, \tag{35}$$

where σ_{stab} controls the smoothness of the stability dependence. Small values of σ_{stab} imply a slowly varying relationship between the tuning parameter and atmospheric stability, while larger values allow stronger bin-to-bin variation. To ensure that the stability effect only describes deviations around the reference level, the latent effect g_b is shifted to have zero mean across the stability bins.

By analogy with the generalised model, individual wind farms are allowed to deviate from the shared cluster-level relationship. In the stability-dependent model, this deviation is defined for each wind farm and stability bin. The wind-farm- and stability-bin-specific central tuning parameter is written as

$$\omega_{p,b} = \omega_b + \sigma_{\mathrm{park,stab}}\, z_{p,b}, \qquad z_{p,b} \sim \mathcal{N}(0,1). \tag{36}$$

Here, $\sigma_{\mathrm{park,stab}}$ controls the magnitude of farm-specific variation around the shared stability-dependent curve. These farm-specific terms allow individual wind farms to differ from the shared stability-dependent trend across stability bins.

The random-walk prior with Gaussian increments is the smoothing component of the model. It does not imply that the stability proxies themselves are assumed to be Gaussian. Rather, it expresses the modelling assumption that neighbouring stability bins should have similar latent tuning parameters, unless the calibrated data provide evidence for a stronger change.

The posterior distribution of the shared stability-dependent central parameter $\omega(s)$ is represented by the posterior samples of the shared bin-specific parameters ω_b, evaluated across the stability bin centres. The posterior median is used as the central estimate, while the 5–95% posterior interval is used to visualise uncertainty. This yields a stability-dependent estimate of the central wake-model tuning parameter for each combination of wake model and stability proxy.

Posterior predictive distributions for individual future calibrated tuning-parameter values additionally include farm-specific or new-wind-farm deviations, together with residual observation-scale variability, through the same log-Student-t predictive structure introduced in the generalised model.

6.2 Stability-dependent tuning-parameter results

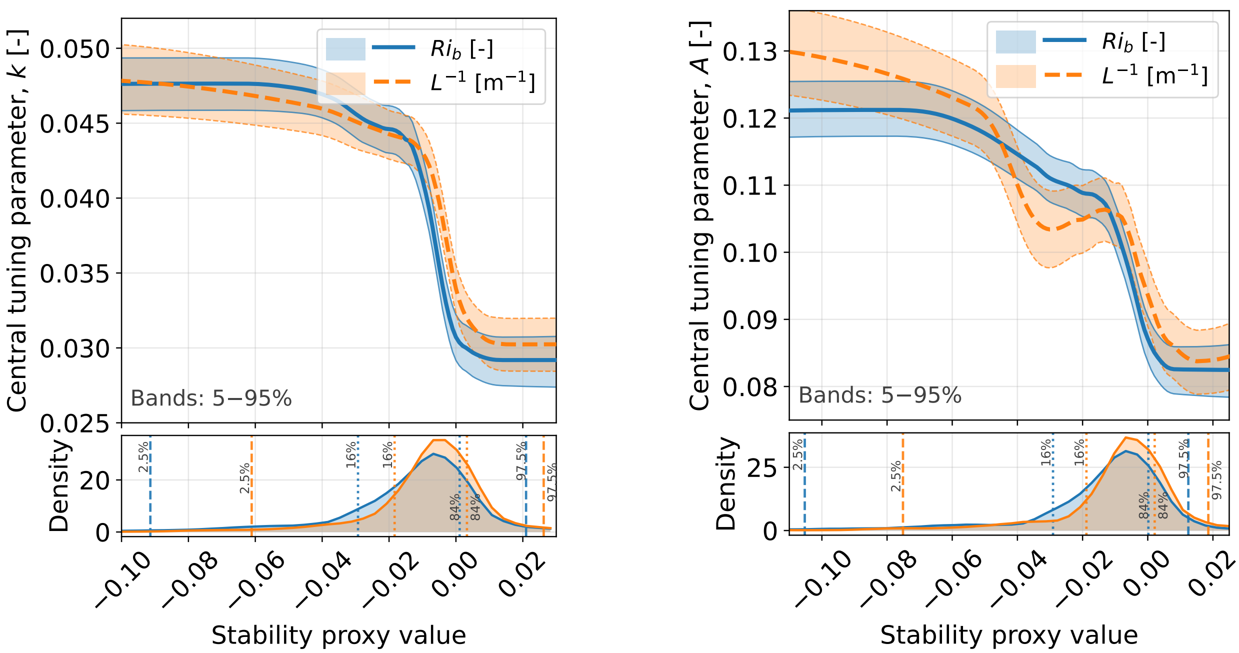
Figure 21 shows the inferred shared stability-dependent central tuning parameter for the Jensen and Gaussian TurbOPark wake models. The blue curve represents the posterior median of the central tuning parameter as a function of Ri_b, while the orange curve represents the corresponding relationship as a function of L^{-1}. The shaded regions indicate the 5–95% posterior intervals.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

The figure shows how the central tuning parameter changes across the observed range of atmospheric stability conditions. The generalised model presented in the previous section provides a single cluster-level central value, whereas the stability-dependent model allows this central value to vary with the stability proxy. Deviations from a horizontal trend therefore indicate that atmospheric stability contains information relevant to the calibrated wake-model parameter. A clear trend is observed, where lower stability proxy values, corresponding to unstable conditions, are associated with larger tuning parameters, implying stronger wake expansion and faster wake recovery.

The uncertainty bands reflect uncertainty in the inferred stability-dependent relationship and the amount of retained calibration data available for each wake model. Wider intervals are expected when fewer calibration instances are available to constrain the stability-dependent trend. This is relevant when comparing the Jensen and Gaussian TurbOPark results, because the retained calibration datasets are not identical for the two wake models. Consequently, differences in posterior spread should not be interpreted as resulting solely from the wake-model formulation.

Within each model, trends in the central part of the stability range should generally be interpreted with greater confidence than trends near the extremes, where quantile-based bins span wider stability ranges. The change in wake expansion appears strongest around near-neutral and slightly unstable conditions. The trend is smoother for the Jensen wake model, while the Gaussian TurbOPark trend shows somewhat stronger local variability.



(a) Jensen wake model with tuning parameter k.

(b) Gaussian TurbOPark model with tuning parameter A.

Figure 21. Shared stability-dependent central tuning parameter as a function of atmospheric stability, shown for the Jensen wake model (a) and the Gaussian TurbOPark model (b). In each case, the blue curve shows the posterior median as a function of Ri_b, while the orange curve shows the posterior median as a function of L^{-1}. The shaded regions indicate the 5–95% posterior intervals.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Table 4 reports posterior summaries of the main scale and tail parameters in the stability-dependent model. The parameter σ_{stab} controls the smoothness of the shared stability-dependent trend, $\sigma_{\mathrm{park,stab}}$ controls the magnitude of farm-specific deviations around this shared trend, σ_{intra} describes residual intra-farm variability around the stability-dependent central parameter, α_ϵ quantifies the residual-mismatch-dependent observation-scale contribution, and ν controls the tail heaviness of the log-Student-t likelihood. For comparison with the stability-independent formulation, the corresponding uncertainty parameters of the generalised model are reported in Table 1.

Table 4. Posterior summaries of the main scale and tail parameters in the stability-dependent tuning-parameter model. Values are reported as the posterior median with the 5–95% posterior interval in brackets.

Wake model	Proxy	σ_{stab}	$\sigma_{\mathrm{park,stab}}$	σ_{intra}	α_ϵ	ν
Jensen	Ri_b	0.0016 [0.0011, 0.0022]	0.002 [0.002, 0.003]	0.014 [0.014, 0.015]	0.12 [0.12, 0.12]	9.4 [8.5, 10.4]
Jensen	L^{-1}	0.0015 [0.0010, 0.0022]	0.003 [0.002, 0.003]	0.014 [0.014, 0.015]	0.13 [0.12, 0.13]	9.3 [8.4, 10.4]
Gaussian TurbOPark	Ri_b	0.0036 [0.0025, 0.0052]	0.004 [0.003, 0.005]	0.015 [0.013, 0.016]	0.26 [0.25, 0.27]	6.6 [5.8, 7.5]
Gaussian TurbOPark	L^{-1}	0.0062 [0.0042, 0.0088]	0.006 [0.005, 0.007]	0.015 [0.014, 0.017]	0.26 [0.25, 0.27]	6.2 [5.6, 7.1]

The posterior summaries in Table 4 should be interpreted together with the stability-independent summaries in Table 1. The posterior values of σ_{stab} indicate that the inferred stability-dependent trends are relatively smooth on their respective tuning-parameter scales. For the Jensen model, the median smoothness scale is approximately 0.0015–0.0016, while for the Gaussian TurbOPark model it is approximately 0.0036 for Ri_b and 0.0062 for L^{-1}. These values describe the typical bin-to-bin variation in the latent stability-dependent central parameter. Larger changes across the full stability range can still occur through the accumulation of local bin-to-bin deviations, as seen in Figures 21a and 21b.

The farm-level deviation scale, $\sigma_{\mathrm{park,stab}}$, remains non-negligible for both wake models. For the Jensen model, the posterior medians are approximately 0.002–0.003, while for the Gaussian TurbOPark model they are approximately 0.004–0.006. This indicates that farm-specific differences remain present even after conditioning the central tuning parameter on atmospheric stability. In other words, atmospheric stability explains part of the systematic variation in the calibrated parameters, but does not remove farm-specific variability.

The residual intra-farm scale, σ_{intra}, is approximately 0.014–0.015 for the Jensen model and 0.015 for the Gaussian TurbOPark model. This term contributes to the observation uncertainty around the stability-dependent central parameter and is converted, together with the other uncertainty contributions, to the log-scale residual width used in the log-Student-t likelihood. Compared with the stability-independent residual-mismatch-aware formulation, where σ_{intra} is approximately 0.016 for Jensen and 0.018 for Gaussian TurbOPark, the residual intra-farm scale is slightly reduced when atmospheric stability is included. This indicates that part of the residual scatter previously represented by the common intra-farm scale is captured by the stability-dependent trend, although the reduction is modest and residual observation-level variability remains present.

The residual-mismatch coefficient α_ϵ remains clearly positive for both wake models, with values around 0.12–0.13 for Jensen and around 0.26 for Gaussian TurbOPark. This indicates that post-calibration residual power mismatch continues to

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

provide observation-specific information about the spread of calibrated tuning parameters after conditioning the central tuning parameter on atmospheric stability.

Finally, the inferred degrees-of-freedom parameter ν indicates that the log-Student-t likelihood retains heavy-tailed residual behaviour. For the Jensen model, ν is approximately 9.3–9.4, while for the Gaussian TurbOPark model it is approximately 6.2–6.6. Compared with the stability-independent residual-mismatch-aware formulation, where ν is approximately 9.9 for Jensen and 8.5 for Gaussian TurbOPark, the inferred tail behaviour remains broadly similar for Jensen but becomes somewhat heavier-tailed for Gaussian TurbOPark. This indicates that occasional large multiplicative deviations in the calibrated tuning parameters remain present even after accounting for atmospheric stability and residual mismatch.

In addition to the posterior distribution of the central stability-dependent parameter, posterior predictive distributions are constructed for the calibrated tuning parameter as a function of atmospheric stability. These distributions include uncertainty in the latent central relationship $\omega(s)$, wind-farm-level variation around this relationship, and residual variability in the calibrated values. They therefore represent the range of tuning-parameter values predicted for individual future calibration instances under a given stability condition.

For the residual-mismatch-aware model, the posterior predictive distributions shown here are evaluated at $\epsilon_{\text{new}} = 0$. They should therefore be interpreted as conditional low-mismatch predictions, for which the additional residual-mismatch-dependent scale contribution is absent. The remaining predictive spread is still governed by uncertainty in the latent stability-dependent relationship, wind-farm-level variability, and the common residual intra-farm scale.

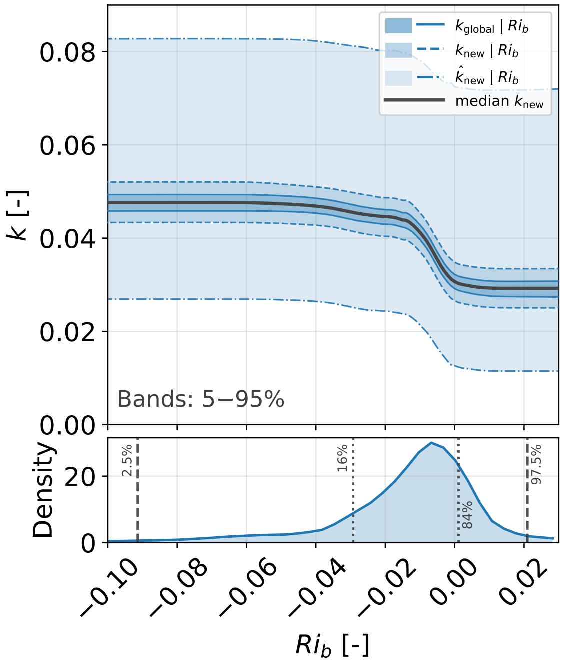
Figures 22 and 23 show the stability-dependent posterior and posterior predictive distributions for the Jensen and Gaussian TurbOPark wake models, respectively. For each wake model, the distributions are shown as a function of both Ri_b and L^{-1}. These figures provide a complementary interpretation of the central-parameter curves in Figure 21: ω_{global} describes posterior uncertainty in the global latent central parameter, ω_{new} describes the posterior predictive distribution of the latent central parameter for a wind farm outside the calibration dataset, and $\hat{\omega}_{\text{new}}$ describes the posterior predictive distribution of individual calibrated values.

This distinction is important because the calibrated 10-minute tuning-parameter estimates contain variability from changing flow conditions, measurement and data-processing effects, residual model mismatch, and observation-level scatter not captured by the central stability-dependent relationship. As a result, the interval for ω_{new} is wider than that of ω_{global} because it includes wind-farm-level variability, while the interval for $\hat{\omega}_{\text{new}}$ is wider still because it additionally includes residual variability in individual calibrated values.

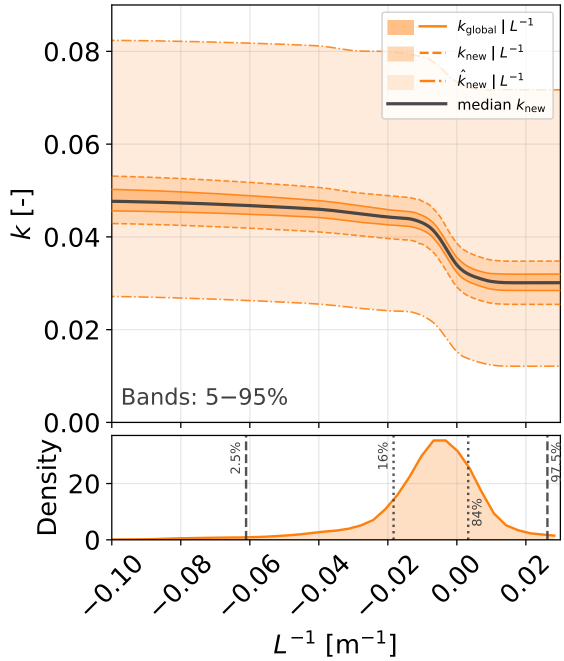
To connect the smooth stability-dependent relationships to the wake-loss analysis, representative tuning parameters are also derived for discrete atmospheric stability classes. The stability classes are defined using the thresholds introduced in Section 5. For each stability class, the inferred stability-dependent central tuning parameter, $\omega(s)$, is summarised over the corresponding stability range. This yields class-specific central tuning parameters that provide a compact representation of the smooth stability dependence.

The resulting class-specific posterior medians are shown in Tables 5 and 6 for L^{-1} and Ri_b, respectively. These values are used in the subsequent wake-loss analysis. Since the stability classes discretise an underlying continuous relationship, the

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026



(a) As a function of Ri_b.



(b) As a function of L^{-1}.

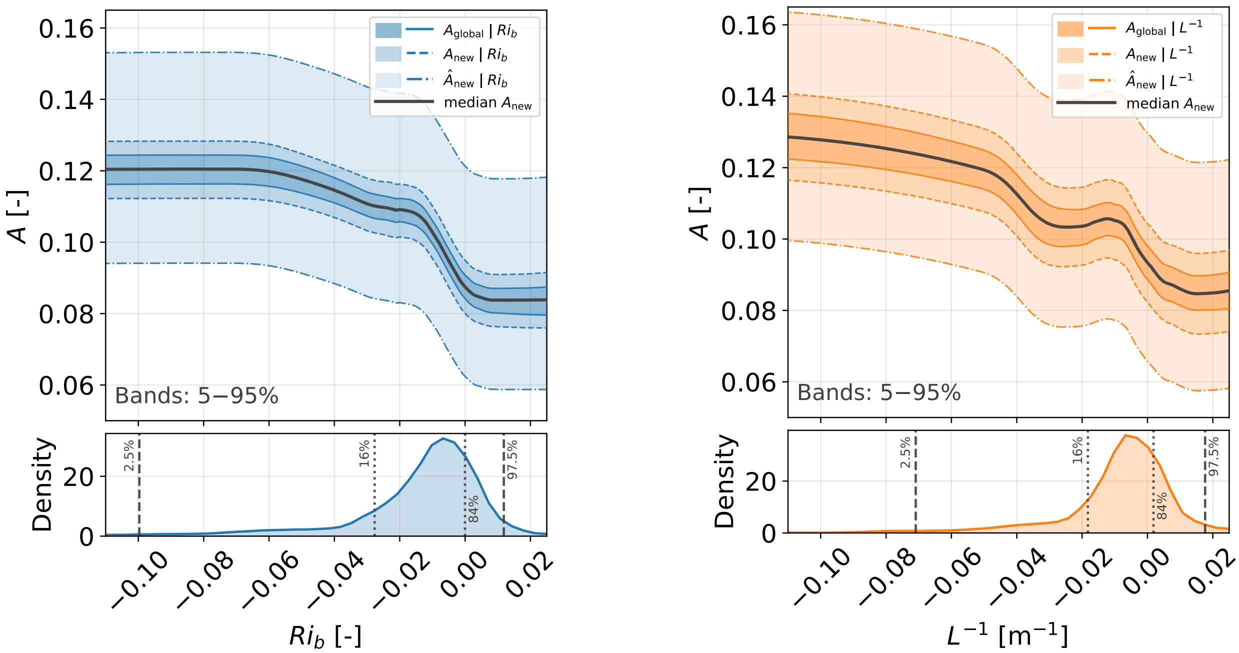
Figure 22. Stability-dependent posterior and posterior predictive distributions of the Jensen wake expansion parameter k, shown as a function of Ri_b (a) and L^{-1} (b). The upper panels show the posterior distribution of the global latent central parameter k_{global}, the posterior predictive distribution of the latent central parameter for a wind farm outside the calibration dataset k_{new}, and the posterior predictive distribution of the corresponding calibrated value $\hat{k}_{\mathrm{new}}$. In the upper panels, shaded bands indicate 5–95 % intervals, and the black line denotes the posterior median of k_{new}. The lower panels show the empirical density of the atmospheric-stability proxy values used for conditioning.

755 reported values should be interpreted as representative central parameters for broad stability regimes, rather than as sharp physical transitions at the class boundaries.

Table 5. Atmospheric stability classes expressed in terms of inverse Monin–Obukhov length, L^{-1}, following Porchetta et al. (2019). Reported values are posterior medians of the class-specific central tuning parameters derived from the stability-dependent model.

Stability class	Range of L^{-1} [m^{-1}]	Jensen, k [–]	Gaussian TurbOPark, A [–]
Very unstable	$L^{-1} < -0.005$	0.047	0.125
Unstable	$-0.005 \leq L^{-1} < -0.001$	0.038	0.098
Near neutral	$-0.001 \leq L^{-1} < 0.001$	0.034	0.094
Stable	$0.001 \leq L^{-1} < 0.005$	0.032	0.090
Very stable	$L^{-1} \geq 0.005$	0.030	0.086

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026



(a) As a function of Ri_b.

(b) As a function of L^{-1}.

Figure 23. Stability-dependent posterior and posterior predictive distributions of the Gaussian TurbOPark tuning parameter A, shown as a function of Ri_b (a) and L^{-1} (b). The upper panels show the posterior distribution of the global latent central parameter A_{global}, the posterior predictive distribution of the latent central parameter for a wind farm outside the calibration dataset A_{new}, and the posterior predictive distribution of the corresponding calibrated value $\hat{A}_{\mathrm{new}}$. In the upper panels, shaded bands indicate 5–95 % intervals, and the black line denotes the posterior median of A_{new}. The lower panels show the empirical density of the atmospheric-stability proxy values used for conditioning.

Table 6. Atmospheric stability classes expressed in terms of bulk Richardson number, Ri_b, following Cantero et al. (2022). Reported values are posterior medians of the class-specific central tuning parameters derived from the stability-dependent model. Dashes indicate stability classes for which no representative value is reported because too few retained calibration instances are available.

Stability class	Range of Ri_b [–]	Jensen, k [–]	Gaussian TurbOPark, A [–]
Very unstable	$Ri_b < -0.023$	0.048	0.120
Unstable	$-0.023 \leq Ri_b < -0.011$	0.044	0.109
Weakly unstable	$-0.011 \leq Ri_b < -0.0036$	0.039	0.099
Near neutral	$-0.0036 \leq Ri_b < 0.0072$	0.030	0.086
Weakly stable	$0.0072 \leq Ri_b < 0.042$	0.029	0.084
Stable	$0.042 \leq Ri_b < 0.084$	–	–
Very stable	$Ri_b \geq 0.084$	–	–

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

6.3 Propagation to wake-loss estimates

The stability-dependent tuning parameters are subsequently compared with the stability-independent generalised reference case to evaluate how the inferred stability dependence affects predicted wake losses. This analysis is intended as a sensitivity study under idealised fixed-stability conditions, rather than as a realistic annual energy production calculation. To isolate the effect of the tuning-parameter choice, the same long-term joint wind-speed and wind-direction distribution derived from ERA5 is used for all wake-loss calculations, while only the wake-model tuning parameter is varied.

For each wake model, three representative tuning-parameter settings are considered: an unstable case, a stability-independent generalised case, and a stable case. The stability-dependent settings are selected as rounded representative values from the class-specific posterior medians reported in Tables 5 and 6. For the Jensen wake model, $k = 0.048$, $k = 0.04$, and $k = 0.03$ are used for the unstable, generalised, and stable cases, respectively. For the Gaussian TurbOPark model, the corresponding values are $A = 0.12$, $A = 0.10$, and $A = 0.085$. The unstable and stable values are derived from the stability-dependent model, whereas the generalised values correspond to the stability-independent central tuning parameters.

The generalised tuning parameter should not be interpreted as a neutral-stability value. It is inferred by pooling over all stability conditions represented in the calibration dataset and therefore reflects the stability distribution of the available observations. If the calibration dataset is dominated by unstable or stable conditions, the stability-independent central parameter may be shifted toward the corresponding part of the stability-dependent relationship. The generalised case is therefore interpreted as a stability-averaged reference case for the calibration dataset, rather than as a neutral-stability case.

This distinction is relevant when comparing the inferred generalised parameters with reference values from idealised simulations or literature studies. Large-eddy simulations and model-development studies may consider idealised neutral atmospheric conditions, whereas the generalised parameter inferred here reflects the stability mix present at the case-study offshore site. Differences between neutral-condition reference values and the values presented in this work can therefore arise partly from differences in the atmospheric stability conditions represented by the datasets.

Three wake-loss quantities are evaluated for wind farms in the Belgian–Dutch offshore concession zone. Internal wake losses are calculated by considering wakes generated by turbines within the same wind farm only. Total wake losses are calculated by including both internal wakes and wakes from the surrounding wind farms in the Belgian–Dutch offshore concession zone. External wake losses are then obtained as the difference between total and internal wake losses. The resulting distributions across wind farms are summarised using boxplots in Figure 24.

For internal wake losses, the unstable, generalised, and stable tuning-parameter assumptions lead to clearly different loss estimates, showing that the stability-dependent tuning parameter has a substantial impact on predicted internal wake losses. Larger tuning parameters correspond to faster wake expansion and recovery, and therefore generally reduce the predicted internal wake-loss magnitude. However, for a given stability assumption, the Jensen and Gaussian TurbOPark models give broadly similar internal wake-loss estimates. This suggests that, after calibration and generalisation, the choice between the two wake-model formulations has a limited effect on internal wake-loss predictions for the considered case-study farms.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

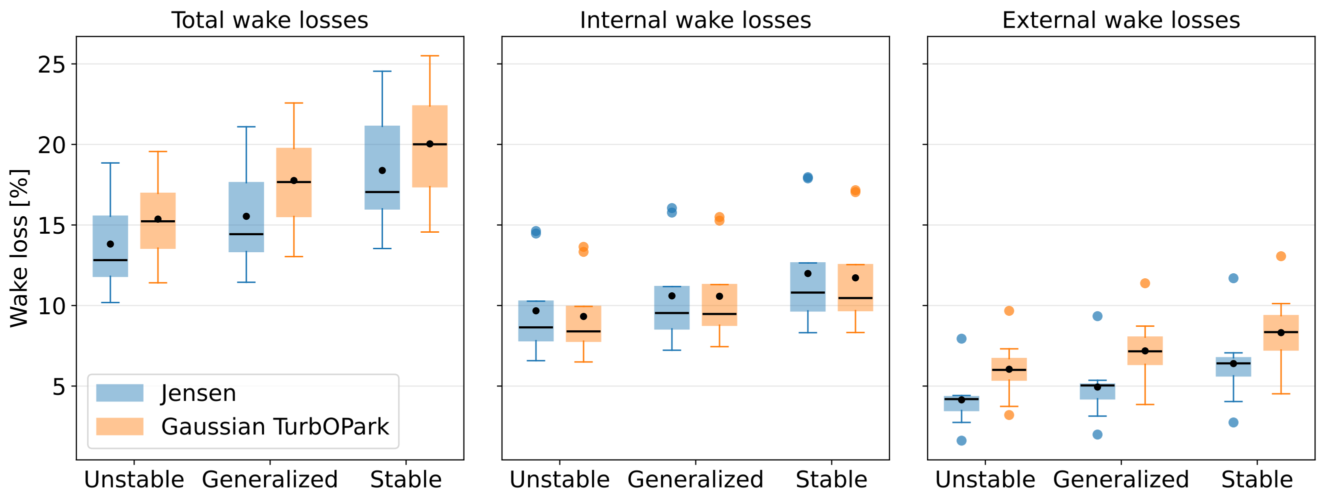


Figure 24. Wake-loss estimates obtained using representative tuning parameters for unstable, generalised, and stable conditions for the Jensen and Gaussian TurbOPark wake models. Wake losses are normalised by the unwaked reference production. Internal wake losses include wakes generated within each wind farm only, total wake losses include surrounding wind farms in the Belgian–Dutch offshore concession zone, and external wake losses are computed as the difference between total and internal wake losses.

790 Larger differences between the wake models appear when the surrounding wind farms are included. For both wake models, the total wake-loss estimates increase when moving from the unstable to the generalised and stable tuning-parameter assumptions, consistent with the reduced wake expansion and slower wake recovery associated with smaller tuning-parameter values. In addition, a clearer distinction between the Jensen and Gaussian TurbOPark models is visible for total wake losses than for internal wake losses alone. This indicates that the difference between the two wake-model formulations is mainly driven by the 795 external wake-loss component.

The stronger external wake losses predicted by the Gaussian TurbOPark model are consistent with differences in the downstream wake-recovery formulation of the two models. The Jensen model represents wake recovery through a linear wake-expansion assumption, whereas the Gaussian TurbOPark formulation includes a TI-dependent expansion model. Under the assumed turbulence conditions, this leads to stronger wake deficits farther downstream, and therefore stronger external wake 800 losses when surrounding wind farms are included. As a result, differences between the Jensen and Gaussian TurbOPark models become more pronounced for total and external wake losses than for internal wake losses alone.

The spread within each boxplot should be interpreted separately from the difference between wake models. This spread reflects variability between wind farms, arising from both site-specific factors, such as farm location within the offshore cluster and exposure to wakes from neighbouring wind farms, and layout-specific factors, such as internal layout topology and turbine 805 spacing. Therefore, while the difference between the Jensen and Gaussian TurbOPark results is mainly related to the wake-model formulation, the spread across wind farms reflects differences in site-specific and layout-specific wake exposure.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

7 Conclusions

This study presented a Bayesian framework for generalising calibrated analytical wake-model tuning parameters across offshore wind farms. The framework was applied to SCADA-based calibration results for the Jensen wake model and the Gaussian TurbOPark model, with the objective of quantifying cluster-level, farm-specific, and new-farm tuning-parameter uncertainty. Rather than treating the calibrated tuning parameter as a single deterministic value, the proposed approach represents it through posterior distributions that account for calibration repeatability, residual intra-farm variability, inter-farm variability, and residual model-data mismatch.

The generalised results show that the posterior distribution of the cluster-level central tuning parameter is narrower than the posterior predictive distribution for a new wind farm. This distinction is important for model transferability: the cluster-level distribution describes uncertainty in the central parameter inferred from the observed farms, whereas the new-farm distribution additionally includes inter-farm variability. The residual-mismatch-aware formulation further shows that part of the scatter in the calibrated tuning parameters is associated with calibration instances that retain larger post-calibration power mismatch. Including this term reduces the common residual intra-farm scale, indicating that residual mismatch provides observation-specific information about calibrated-parameter uncertainty.

The framework was then extended to investigate whether atmospheric stability explains systematic variation around the generalised tuning parameter. Two stability proxies were considered: the bulk Richardson number, Ri_b, derived from meteorological mast measurements, and the inverse Monin–Obukhov length, L^{-1}, derived from ERA5 reanalysis data. For both wake models and both stability proxies, unstable conditions were associated with larger tuning parameters, corresponding to faster wake expansion and recovery, while stable conditions were associated with smaller tuning parameters and more persistent wakes. This demonstrates that atmospheric stability explains part of the systematic variation in the calibrated wake-model parameters.

The stability-dependent results also show that the stability-independent generalised parameter should not be interpreted as a neutral-stability parameter. Instead, it represents a pooled estimate over the stability conditions present at the case-study offshore site. Consequently, differences between neutral-condition reference values from idealised simulations or literature studies and the values inferred in this work can arise partly from differences in the atmospheric stability conditions represented by the datasets.

Finally, representative stability-class-specific tuning parameters were propagated to wake-loss estimates. The results show that the choice between unstable, generalised, and stable tuning parameters has a substantial impact on predicted wake losses. For internal wake losses, the calibrated Jensen and Gaussian TurbOPark models give broadly similar results for a given stability assumption. Larger differences appear when wakes from surrounding wind farms are included, indicating that the wake-model formulation becomes more influential for external and cluster-scale wake losses. The spread in wake-loss estimates across wind farms reflects both site-specific exposure to neighbouring wakes and layout-specific characteristics such as farm topology and turbine spacing.

Overall, the results demonstrate that Bayesian generalisation provides an uncertainty-aware basis for transferring calibrated analytical wake-model parameters across offshore wind farms. The proposed framework separates cluster-level, farm-specific, and new-farm uncertainty, incorporates residual model-data mismatch, and can be extended to condition the tuning parameter on atmospheric stability. This makes it possible to quantify a generalised calibrated wake-model parameter, identify systematic stability-dependent variation around the generalised estimate, and assess the uncertainty associated with transferring calibrated parameters to other offshore wind farms.

Code and data availability. The wake model framework used in this study is PyWake (Pedersen et al., 2023). PyWake is available from DTU's GitLab repository https://gitlab.windenergy.dtu.dk/TOPFARM/PyWake. The software is archived on Zenodo with DOI 10.5281/zenodo.6806136. The accompanying documentation is available at https://topfarm.pages.windenergy.dtu.dk/PyWake/.

The wake-model calibration was performed using Optuna (Akiba et al., 2019). The source code is available from the Optuna GitHub repository https://github.com/optuna/optuna, with accompanying documentation available at https://optuna.readthedocs.io/en/stable/.

The Bayesian hierarchical modelling and posterior inference were implemented using NumPyro (Phan et al., 2019). The source code is available from the NumPyro GitHub repository https://github.com/pyro-ppl/numpyro, with documentation available at https://num.pyro.ai/.

The Python implementation developed for this work is available from the corresponding author upon reasonable request. The SCADA data used in this work is not openly available due to confidentiality constraints. Meteorological mast data can be obtained from Meetnet Vlaamse Banken, while ERA5 reanalysis data is openly available. The data and methodological requirements for computing the inverse Monin–Obukhov length are available through the ECMWF ERA5 Obukhov-length guidance.

Author contributions. DvB was responsible for conceptualisation, data curation, formal analysis, investigation, methodology, project administration, software development, validation, visualisation, writing of the original draft, implementation of co-author comments, and final manuscript editing. PJD and TV contributed to conceptualisation, methodology, supervision, and manuscript review. KV contributed to data curation using Meetnet Vlaamse Banken data. ARN contributed to final-phase manuscript review. JH contributed to conceptualisation, funding acquisition, supervision, and final-phase manuscript review.

Competing interests. ARN is member of the editorial board of the Wind Energy Science journal. The authors have no other competing interests to declare.

Acknowledgements. The authors would like to acknowledge support from VLAIO through De Blauwe Cluster for the Supersized 5.0 and Cloud4Wake projects. In addition, the authors acknowledge support from the ACCEMO project, funded by the Research Foundation – Flanders (FWO), and from the INSPIRE project, funded through the Sustainable Blue Economy Partnership.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

References

IEC61400-1: Wind Energy Generation Systems - Part 1: Design Requirements, 2019.

A. Sathe, J. Mann, J. G. M. C.: Estimating the systematic errors in turbulence sensed by wind lidars, Tech. rep., DTU, 2010.

Abkar, M. and Porté-Agel, F.: Influence of atmospheric stability on wind-turbine wakes: A large-eddy simulation study, Physics of Fluids, 27, https://doi.org/10.1063/1.4913695, 2015.

Aerts, F., Lanzilao, L., and Meyers, J.: Bayesian uncertainty quantification framework for wake model calibration and validation with historical wind farm power data, Wind Energy, 26, 786–802, https://doi.org/10.1002/we.2841, 2023.

Aerts, F., Devesse, K., and Meyers, J.: Bayesian uncertainty quantification of engineering models for wind farm–atmosphere interaction, Wind Energy Science, 11, 1205–1225, https://doi.org/10.5194/wes-11-1205-2026, 2026.

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M.: Optuna: A Next-generation Hyperparameter Optimization Framework, https://doi.org/10.48550/ARXIV.1907.10902, 2019.

Albornoz, C. P., Soberanis, M. E., Rivera, V. R., and Rivero, M.: Review of atmospheric stability estimations for wind power applications, Renewable and Sustainable Energy Reviews, 163, 112 505, https://doi.org/10.1016/j.rser.2022.112505, 2022.

Bahamonde, M. I. and Litrán, S. P.: Study of the energy production of a wind turbine in the open sea considering the continuous variations of the atmospheric stability and the sea surface roughness, Renewable Energy, 135, 163–175, https://doi.org/10.1016/j.renene.2018.11.075, 2019.

Barthelmie, R. J., Hansen, K., Frandsen, S. T., Rathmann, O., Schepers, J. G., Schlez, W., Phillips, J., Rados, K., Zervos, A., Politis, E. S., and Chaviaropoulos, P. K.: Modelling and measuring flow and wind turbine wakes in large wind farms offshore, Wind Energy, 12, 431–444, https://doi.org/10.1002/we.348, 2009.

Barthelmie, R. J., Pryor, S. C., Frandsen, S. T., Hansen, K. S., Schepers, J. G., Rados, K., Schlez, W., Neubert, A., Jensen, L. E., and Neckelmann, S.: Quantifying the Impact of Wind Turbine Wakes on Power Output at Offshore Wind Farms, Journal of Atmospheric and Oceanic Technology, 27, 1302–1317, https://doi.org/10.1175/2010jtecha1398.1, 2010.

Barthelmie, R. J., Churchfield, M. J., Moriarty, P. J., Lundquist, J. K., Oxley, G. S., Hahn, S., and Pryor, S. C.: The role of atmospheric stability/turbulence on wakes at the Egmond aan Zee offshore wind farm, Journal of Physics: Conference Series, 625, 012 002, https://doi.org/10.1088/1742-6596/625/1/012002, 2015.

Bastankhah, M. and Porté-Agel, F.: A new analytical model for wind-turbine wakes, Renewable Energy, 70, 116–123, https://doi.org/10.1016/j.renene.2014.01.002, 2014.

Bastankhah, M., Welch, B. L., Martínez-Tossas, L. A., King, J., and Fleming, P.: Analytical solution for the cumulative wake of wind turbines in wind farms, Journal of Fluid Mechanics, 911, https://doi.org/10.1017/jfm.2020.1037, 2021.

Bergstra, J. and Bengio, Y.: Random Search for Hyper-Parameter Optimization, Journal of Machine Learning Research, 13, 281–305, http://jmlr.org/papers/v13/bergstra12a.html, 2012.

Bergstra, J., Yamins, D., and Cox, D. D.: Making a Science of Model Search, https://doi.org/10.48550/ARXIV.1209.5111, 2012.

Binsbergen, D. v., Daems, P.-J., Verstraeten, T., Nejad, A., and Helsen, J.: Scalable SCADA-Based Calibration for Analytical Wake Models Across an Offshore Cluster, Journal of Physics: Conference Series, 2745, 012 014, https://doi.org/10.1088/1742-6596/2745/1/012014, 2024.

Campagnolo, F., Imširović, L., Braunbehrens, R., and Bottasso, C. L.: Further calibration and validation of FLORIS with wind tunnel data, Journal of Physics: Conference Series, 2265, 022 019, https://doi.org/10.1088/1742-6596/2265/2/022019, 2022.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

Cantero, E., Sanz, J., Borbón, F., Paredes, D., and García, A.: On the measurement of stability parameter over complex mountainous terrain, Wind Energy Science, 7, 221–235, https://doi.org/10.5194/wes-7-221-2022, 2022.

Daems, P.-J., Binsbergen, D. V., Sterckx, J., Verstraeten, T., and Helsen, J.: A comparison between scanning LiDAR and scada-based hyperparameter tuning of analytical wake models, Journal of Physics: Conference Series, 3025, 012 002, https://doi.org/10.1088/1742-6596/3025/1/012002, 2025.

Doekemeijer, B. M., Van Wingerden, J.-W., and Fleming, P. A.: A tutorial on the synthesis and validation of a closed-loop wind farm controller using a steady-state surrogate model, in: 2019 American Control Conference (ACC), IEEE, https://doi.org/10.23919/acc.2019.8815126, 2019.

Dörenkämper, M., Witha, B., Steinfeld, G., Heinemann, D., and Kühn, M.: The impact of stable atmospheric boundary layers on wind-turbine wakes within offshore wind farms, Journal of Wind Engineering and Industrial Aerodynamics, 144, 146–153, https://doi.org/10.1016/j.jweia.2014.12.011, 2015.

ERA5, E.: ERA5: How to calculate Obukhov Length, https://confluence.ecmwf.int/display/CKB/ERA5

Gaumond, M., Réthoré, P., Ott, S., Peña, A., Bechmann, A., and Hansen, K. S.: Evaluation of the wind direction uncertainty and its impact on wake modeling at the Horns Rev offshore wind farm, Wind Energy, 17, 1169–1178, https://doi.org/10.1002/we.1625, 2013.

Grachev, A. A. and Fairall, C. W.: Dependence of the Monin-Obukhov Stability Parameter on the Bulk Richardson Number over the Ocean, Journal of Applied Meteorology, 36, 406–414, https://doi.org/10.1175/1520-0450(1997)036<0406:DOTMOS>2.0.CO;2, 1997.

Göçmen, T., Campagnolo, F., Duc, T., Eguinoa, I., Andersen, S. J., Petrović, V., Imširović, L., Braunbehrens, R., Liew, J., Baungaard, M., van der Laan, M. P., Qian, G., Aparicio-Sanchez, M., González-Lope, R., Dighe, V. V., Becker, M., van den Broek, M. J., van Wingerden, J.-W., Stock, A., Cole, M., Ruisi, R., Bossanyi, E., Requate, N., Strnad, S., Schmidt, J., Vollmer, L., Sood, I., and Meyers, J.: FarmConners wind farm flow control benchmark – Part 1: Blind test results, Wind Energy Science, 7, 1791–1825, https://doi.org/10.5194/wes-7-1791-2022, 2022.

Hansen, K. S., Barthelmie, R. J., Jensen, L. E., and Sommer, A.: The impact of turbulence intensity and atmospheric stability on power deficits due to wind turbine wakes at Horns Rev wind farm, Wind Energy, 15, 183–196, https://doi.org/10.1002/we.512, 2011.

Hastings, W. K.: Monte Carlo Sampling Methods Using Markov Chains and Their Applications, Biometrika, 57, 97–109, https://doi.org/10.1093/biomet/57.1.97, 1970.

Hersbach, H.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hoffman, M. D. and Gelman, A.: The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, Journal of Machine Learning Research, 15, 1593–1623, https://jmlr.org/papers/v15/hoffman14a.html, 2014.

Holtslag, M.: Far offshore wind conditions in scope of wind energy, Ph.D. thesis, https://doi.org/10.4233/UUID:3C66F401-6CFF-4273-AA49-DF4274BA767F, 2016.

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, https://doi.org/10.1016/j.jweia.2018.04.010, 2018.

J G Pedersen, E Svensson, L. P. and Nygaard, N. G.: Turbulence Optimized Park model with Gaussian wake profile, Journal of Physics: Conference Series, 2265, 022 063, https://doi.org/10.1088/1742-6596/2265/2/022063, 2022.

Jensen, N.: A note on wind generator interaction, 1983.

Katić, I., Højstrup, J., and Jensen, N.: A Simple Model for Cluster Efficiency, 1987.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

Kosović, B., Basu, S., Berg, J., Berg, L. K., Haupt, S. E., Larsén, X. G., Peinke, J., Stevens, R. J. A. M., Veers, P., and Watson, S.: Impact of atmospheric turbulence on performance and loads of wind turbines: knowledge gaps and research challenges, Wind Energy Science, 11, 509–555, https://doi.org/10.5194/wes-11-509-2026, 2026.

Lissaman, P. B. S.: Energy Effectiveness of Arbitrary Arrays of Wind Turbines, Journal of Energy, 3, 323–328, https://doi.org/10.2514/3.62441, 1979.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E.: Equation of State Calculations by Fast Computing Machines, The Journal of Chemical Physics, 21, 1087–1092, https://doi.org/10.1063/1.1699114, 1953.

Meyers, J., Bottasso, C., 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.

Niayifar, A. and Porté-Agel, F.: Analytical Modeling of Wind Farms: A New Approach for Power Prediction, Energies, 9, 741, https://doi.org/10.3390/en9090741, 2016.

Nygaard, N. G., Steen, S. T., Poulsen, L., and Pedersen, J. G.: Modelling cluster wakes and wind farm blockage, Journal of Physics: Conference Series, 1618, 062 072, https://doi.org/10.1088/1742-6596/1618/6/062072, 2020.

Nygaard, N. G., Poulsen, L., Svensson, E., and Pedersen, J. G.: Large-scale benchmarking of wake models for offshore wind farms, Journal of Physics: Conference Series, 2265, 022 008, https://doi.org/10.1088/1742-6596/2265/2/022008, 2022.

Pedersen, M. M. and Larsen, G. C.: Integrated wind farm layout and control optimization, Wind Energy Science, 5, 1551–1566, https://doi.org/10.5194/wes-5-1551-2020, 2020.

Pedersen, M. M., Meyer Forsting, A., van der Laan, P., Riva, R., Alcayaga Romàn, L. A., Criado Risco, J., Friis-Møller, M., Quick, J., Christiansen, J. P. S., Rodrigues, R. V., Olsen, B. T., and Réthoré, P.-E.: PyWake 2.5.0: An open-source wind farm simulation tool, https://doi.org/10.5281/zenodo.6806136, 2023.

Peña, A., Réthoré, P.-E., and Rathmann, O.: Modeling large offshore wind farms under different atmospheric stability regimes with the Park wake model, Renewable Energy, 70, 164–171, https://doi.org/10.1016/j.renene.2014.02.019, 2014.

Phan, D., Pradhan, N., and Jankowiak, M.: Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro, https://doi.org/10.48550/arXiv.1912.11554, 2019.

Platis, A., Hundhausen, M., Lampert, A., Emeis, S., and Bange, J.: The Role of Atmospheric Stability and Turbulence in Offshore Wind-Farm Wakes in the German Bight, Boundary-Layer Meteorology, 182, 441–469, https://doi.org/10.1007/s10546-021-00668-4, 2021.

Porchetta, S., Temel, O., Muñoz-Esparza, D., Reuder, J., Monbaliu, J., van Beeck, J., and van Lipzig, N.: A new roughness length parameterization accounting for wind–wave (mis)alignment, Atmospheric Chemistry and Physics, 19, 6681–6700, https://doi.org/10.5194/acp-19-6681-2019, 2019.

Porté-Agel, F., Bastankhah, M., and Shamsoddin, S.: Wind-Turbine and Wind-Farm Flows: A Review, Boundary-Layer Meteorology, 174, 1–59, https://doi.org/10.1007/s10546-019-00473-0, 2019.

Rodrigo, J. S., Cantero, E., García, B., Borbón, F., Irigoyen, U., Lozano, S., Fernande, P. M., and Chávez, R. A.: Atmospheric stability assessment for the characterization of offshore wind conditions, Journal of Physics: Conference Series, 625, 012 044, https://doi.org/10.1088/1742-6596/625/1/012044, 2015.

Sathe, A., Gryning, S., and Peña, A.: Comparison of the atmospheric stability and wind profiles at two wind farm sites over a long marine fetch in the North Sea, Wind Energy, 14, 767–780, https://doi.org/10.1002/we.456, 2011.

Sorbjan, Z. and Grachev, A. A.: An Evaluation of the Flux–Gradient Relationship in the Stable Boundary Layer, Boundary-Layer Meteorology, 135, 385–405, https://doi.org/10.1007/s10546-010-9482-3, 2010.

https://doi.org/10.5194/wes-2026-120
Preprint. Discussion started: 16 July 2026

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Springer Netherlands, ISBN 9789400930278, https://doi.org/10.1007/978-94-009-3027-8, 1988.

Sukhman, D., Kingston, R., Bø, O. I., and Bartl, J.: An offshore wind farm's annual energy production and associated wake losses at Vestavind F, Journal of Physics: Conference Series, 3131, 012 036, https://doi.org/10.1088/1742-6596/3131/1/012036, 2025.

Syed, A. H., Mann, J., Platis, A., and Bange, J.: Turbulence structures and entrainment length scales in large offshore wind farms, Wind Energy Science, 8, 125–139, https://doi.org/10.5194/wes-8-125-2023, 2023.

van Beek, M. T., Viré, A., and Andersen, S. J.: Sensitivity and Uncertainty of the FLORIS Model Applied on the Lillgrund Wind Farm, Energies, 14, 1293, https://doi.org/10.3390/en14051293, 2021.

van Binsbergen, D.: Modeling and analysis of offshore wind farm wake effects on wind turbine components and power production, Ph.D. thesis, VUB, NTNU, 2025.

Van Binsbergen, D., Daems, P.-J., Verstraeten, T., Nejad, A., and Helsen, J.: Performance comparison of analytical wake models calibrated on a large offshore wind cluster, Journal of Physics: Conference Series, 2767, 092 059, https://doi.org/10.1088/1742-6596/2767/9/092059, 2024.

van Binsbergen, D., Daems, P.-J., Verstraeten, T., Nejad, A. R., and Helsen, J.: Hyperparameter tuning framework for calibrating analytical wake models using SCADA data of an offshore wind farm, Wind Energy Science, 9, 1507–1526, https://doi.org/10.5194/wes-9-1507-2024, 2024.

Van Wijk, A., Beljaars, A., Holtslag, A., and Turkenburg, W.: Evaluation of stability corrections in wind speed profiles over the North Sea, Journal of Wind Engineering and Industrial Aerodynamics, 33, 551–566, https://doi.org/10.1016/0167-6105(90)90007-y, 1990.

Zhan, L., Letizia, S., and Iungo, G. V.: Optimal tuning of engineering wake models through lidar measurements, Wind Energy Science, 5, 1601–1622, https://doi.org/10.5194/wes-5-1601-2020, 2020.

Zong, H. and Porté-Agel, F.: A momentum-conserving wake superposition method for wind farm power prediction, Journal of Fluid Mechanics, 889, https://doi.org/10.1017/jfm.2020.77, 2020.

Ørsted: Ørsted presents update on its long-term financial targets, 2019.